Fermilab

DESI Spectroscopy of HETDEX Emission-line Candidates I: Line Discrimination Validation

FERMILAB-PUB-25-0186-PPD

arXiv:2503.02229

Draft version March 5, 2025
Typeset using LaTeX **twocolumn** style in AASTeX631

DESI Spectroscopy of HETDEX Emission-line Candidates I: Line Discrimination Validation

Martin Landriau,[1] Erin Mentuch Cooper,[2,3] Dustin Davis,[2] Karl Gebhardt,[2] Robin Ciardullo,[4,5] Éric Armengaud,[6] Arjun Dey,[7] Anand Raichoor,[1] David J. Schlegel,[1] Michael Wilson,[8] J. Aguilar,[1] S. Ahlen,[9] D. Bianchi,[10,11] D. Brooks,[12] T. Claybaugh,[1] A. de la Macorra,[13] S. Ferraro,[1,14] J. E. Forero-Romero,[15,16] E. Gaztañaga,[17,18,19] S. Gontcho A Gontcho,[1] G. Gutierrez,[20] C. Hahn,[21] K. Honscheid,[22,23] C. Howlett,[24] M. Ishak,[25] S. Juneau,[7] R. Kehoe,[26] T. Kisner,[1] Anthony Kremin,[1] L. Le Guillou,[27] Michael E. Levi,[1] M. Manera,[28,29] A. Meisner,[7] R. Miquel,[30,29] J. Moustakas,[31] S. Nadathur,[18] I. Pérez-Ràfols,[32] C. Poppett,[1,33] F. Prada,[34] G. Rossi,[35] E. Sanchez,[36] M. Schubnell,[37] D. Sprayberry,[7] G. Tarlé,[38] B. A. Weaver,[7] R. Zhou,[1] H. Zou,[39] Daniel J. Farrow,[40,41] Gary J. Hill,[2,3] Donghui Jeong,[4,5] Chenxu Liu,[42] Shun Saito,[43,44] and Donald P. Schneider[4,5]

[1] *Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*
[2] *Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA*
[3] *McDonald Observatory, The University of Texas at Austin, 2515 Speedway Boulevard, Austin, TX 78712, USA*
[4] *Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA*
[5] *Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA*
[6] *Saclay - Commissariat à l'énergie atomique et aux énergies alternatives, IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*
[7] *NSF NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, USA*
[8] *Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*
[9] *Physics Dept., Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA*
[10] *Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Via Celoria 16, I-20133 Milano, Italy",0000-0001-9712-0006*
[11] *INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy*
[12] *Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK*
[13] *Instituto de Física, Universidad Nacional Autónoma de México, Circuito de la Investigación Científica, Ciudad Universitaria, Cd. de México C. P. 04510, México*
[14] *University of California, Berkeley, 110 Sproul Hall #5800 Berkeley, CA 94720, USA*
[15] *Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia*
[16] *Observatorio Astronómico, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia*
[17] *Institut d'Estudis Espacials de Catalunya (IEEC), c/ Esteve Terradas 1, Edifici RDIT, Campus PMT-UPC, 08860 Castelldefels, Spain*
[18] *Institute of Cosmology and Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth, PO1 3FX, UK*
[19] *Institute of Space Sciences, ICE-CSIC, Campus UAB, Carrer de Can Magrans s/n, 08913 Bellaterra, Barcelona, Spain*
[20] *Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA*
[21] *Steward Observatory, University of Arizona, 933 N, Cherry Ave, Tucson, AZ 85721, USA*
[22] *Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[23] *Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[24] *School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*
[25] *Department of Physics, The University of Texas at Dallas, 800 W. Campbell Rd., Richardson, TX 75080, USA*
[26] *Department of Physics, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75275, USA*
[27] *Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), FR-75005 Paris, France*
[28] *Departament de Física, Serra Húnter, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain*
[29] *Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Edifici Cn, Campus UAB, 08193, Bellaterra (Barcelona), Spain*
[30] *Institució Catalana de Recerca i Estudis Avançats, Passeig de Lluís Companys, 23, 08010 Barcelona, Spain*
[31] *Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA*
[32] *Departament de Física, EEBE, Universitat Politècnica de Catalunya, c/Eduard Maristany 10, 08930 Barcelona, Spain*
[33] *Space Sciences Laboratory, University of California, Berkeley, 7 Gauss Way, Berkeley, CA 94720, USA*
[34] *Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, s/n, E-18008 Granada, Spain*
[35] *Department of Physics and Astronomy, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea*

Corresponding author: Martin Landriau
mlandriau@lbl.gov

[36] CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain
[37] Department of Physics, University of Michigan, 450 Church Street, Ann Arbor, MI 48109, USA
[38] University of Michigan, 500 S. State Street, Ann Arbor, MI 48109, USA
[39] National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Rd., Chaoyang District, Beijing, 100012, P.R. China
[40] Centre of Excellence for Data Science, Artificial Intelligence & Modelling (DAIM), University of Hull, Cottingham Road, Hull, HU6 7RX, UK
[41] E. A. Milne Centre for Astrophysics, University of Hull, Cottingham Road, Hull, HU6 7RX, UK
[42] South-Western Institute for Astronomy Research, Yunnan University, Kunming, Yunnan, 650500, People's Republic of China
[43] Institute for Multi-messenger Astrophysics and Cosmology, Department of Physics, Missouri University of Science and Technology, 1315 N Pine St, Rolla, MO 65409
[44] 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

(Received March 4, 2025; Revised; Accepted)

ABSTRACT

The Hobby-Eberly Dark Energy Experiment (HETDEX) is an untargeted spectroscopic galaxy survey that uses Lyα emitting galaxies (LAEs) as tracers of $1.9 < z < 3.5$ large scale structure. Most detections consist of a single emission line, whose identity is inferred via a Bayesian analysis of ancillary data. To determine the accuracy of these line identifications, HETDEX detections were observed with the Dark Energy Spectroscopic Instrument (DESI). In two DESI pointings, high confidence spectroscopic redshifts are obtained for 1157 sources, including 982 LAEs. The DESI spectra are used to evaluate the accuracy of the HETDEX object classifications, and tune the methodology to achieve the HETDEX science requirement of $\lesssim 2\%$ contamination of the LAE sample by low-redshift emission-line galaxies, while still assigning 96% of the true Lyα emission sample with the correct spectroscopic redshift. We compare emission line measurements between the two experiments assuming a simple Gaussian line fitting model. Fitted values for the central wavelength of the emission line, the measured line flux and line widths are consistent between the surveys within uncertainties. Derived spectroscopic redshifts, from the two classification pipelines, when both agree as an LAE classification, are consistent to within $\langle \Delta z/(1+z) \rangle = 6.9 \times 10^{-5}$ with an rms scatter of 3.3×10^{-4}.

Data to be released concurrently with DESI DR1.

Keywords: methods: observational — surveys

1. INTRODUCTION

HETDEX (Hill et al. 2021; Gebhardt et al. 2021) is an untargeted galaxy survey that uses Lyα emitting galaxies (LAEs) as tracers of large-scale structure in the redshift range $1.88 < z < 3.52$. The survey instrument, the Visible Integral-Field Replicable Unit Spectrograph (VIRUS) consists of 78 integral field units (IFUs), each with 448 1.″5- diameter fibers, feeding 156 spectrographs with a spectral resolution of 5.6 Å ($R \sim$800) in the wavelength range 3500 - 5500 Å. The goal of HETDEX is to measure the power spectrum of LAEs and thereby constrain dark energy at $z \sim 2$ with as little contamination as possible from foreground objects.

The challenge of HETDEX is that normal LAEs, i.e. those without AGN emission, are single emission-line sources in HETDEX spectra. However, because of the limited wavelength range of the VIRUS spectrographs, foreground [O II] emitting galaxies between $0.13 < z < 0.47$ also appear as single-line objects. Since the resolution of VIRUS is insufficient to resolve the [O II] doublet, these objects can be mistaken for LAEs. Misclassifications are therefore a serious issue, since to be successful, HETDEX needs to ensure that contamination of the LAE sample by [O II] galaxies is $\lesssim 2\%$ (Gebhardt et al. 2021).

During a pilot survey, Adams et al. (2011) demonstrated that a equivalent width (EW) cuts can be used to separate LAEs from most [O II] galaxies. (See also Gawiser et al. 2007.) By supplementing the EW selection using a modified Bayesian approach based on prior knowledge of the galaxy populations, (Leung et al. 2017; Davis et al. 2023) show that that the misclassification of [O II] galaxies in the HETDEX LAE sample may be as low as 1.5%. However, their conclusion was based on a comparison sample of spectroscopically confirmed

objects that was biased towards brighter high-redshift galaxies.

The Dark Energy Spectroscopic Instrument (DESI) (Levi et al. 2013; DESI Collaboration et al. 2016a,b, 2022; Silber et al. 2023; Miller et al. 2024; Poppett et al. 2024), survey is a program of targeted spectroscopy aimed at building the largest 3D map of the Universe to date, and thereby inferring the expansion history of the Universe up to redshift $z = 3.5$. The instrument consists of 5000 robot-positioned fibers feeding 10 3-arm spectrographs covering the wavelength range 3600-9800 Å. With its wider wavelength coverage and higher spectral resolution (R = 2,000 - 3,200) relative to HETDEX, DESI spectra have a clear advantage over HETDEX for the classification of faint emission-line objects. Not only can DESI resolve the [O II] doublet, but its longer exposure times and wider wavelength range can detect continuum emission and reveal additional spectral lines such as Hα for lower-redshift sources, and C III] λ1909 and C IV in higher-redshift objects.

The focus of this paper is to determine whether the methodology of Leung et al. (2017) as implemented by Davis et al. (2023) does indeed meet the HETDEX science requirements for low-redshift contamination. A second goal of the analysis is to be part of a pilot survey for DESI-2, a proposed survey targeting high-z galaxies using DESI; this aspect will be studied in a companion paper.

Section 2 describes our target selection process and the DESI observations. Section 3 presents our analysis and the results of our visual classification and redshift determinations. In Section 4 the DESI spectroscopic sample is used to measure HETDEX's contamination rate and accuracy. Finally, Section 5 presents a comparison of measured line fluxes between the two experiments; and Section 6 states our conclusions. All wavelengths stated in this paper are as measured in air unless otherwise noted. No dust corrections are applied, although a measure of visual dust extinction, A_V is included in the catalog provided with this paper to be applied as desired. A full description of the data accompanying the release is provided in the Appendix.

2. TARGET SELECTION AND OBSERVATIONS

During the DESI survey validation phase (SV) (DESI Collaboration et al. 2024a), a number of secondary target projects were carried out, one of them being a follow up of HETDEX detections. Figure 1 show the HETDEX main survey footprint and the completion status at the time of designing the target selection. The circles show the two DESI pointings. At the time of target selection, HETDEX was in its early stages and these fields had the most complete coverage within the area of the DESI focal plane and contained a number of interesting objects.

The parent sample of HETDEX detections are from HDR2 (internal data release version 2.1.2). A subset of this dataset is presented in the HETDEX Public Source Catalog 1 (Mentuch Cooper et al. 2023), the difference being that the public release was limited to emission lines detections with a signal-to-noise ratio (S/N) above 5.5 (the internal sample is limited to sources with $S/N \geq 4.8$) and contain some omissions due to quality assessment. In this paper, we cross-match the input sample to those in a later HETDEX internal release (HDR3) to better calibrate with the latest HETDEX reduction pipeline. Note: separations between the input coordinates (listed by 'TARGET_RA', 'TARGET_DEC') and the HETDEX positions are provided in the column 'SEP' in the data table released with this paper. The initial target coordinates are consistent with the later catalog release (DR3) within 0.″3.

The DESI sample was sub-selected from the main HETDEX emission line galaxy catalog based on two criteria:

1. *LAE candidates*: A $g_{\rm HETDEX}$ value (a pseudo g-magnitude measured from the HETDEX spectrum itself) fainter than 22.5 mag (AB) and a P(Lyα) likelihood value greater than 0.5 (Leung et al. 2017). Line emitters with P(Lyα) greater than 0.5 are assigned a redshift assuming the detected emission line is Lyα. Conversely, detections with P(Lyα) less than 0.5 are generally assigned a redshift assuming the emission line is due to [O II] λ3727. (Other observables, such as the detection of a second emission line, are also incorporated into the probability value; for more details, see Davis et al. 2023.)

2. *Borderline LAE/[O II] emitters:* Faint galaxies with $g_{\rm HETDEX}$ value between 23 and 25 mag as well as a P(Lyα) likelihood value between 0.3 and 0.5 in the second date release HETDEX catalog. This sub-sample represents galaxies which are either faint [O II] emitters or bright (LBG-like) LAEs and will be most difficult to classify based on a single emission line. This selection does bias the sample towards slightly brighter targets. As a result the contamination that is measured from low redshift emission line interlopers is expected to be biased slightly higher compared to the full HETDEX sample.

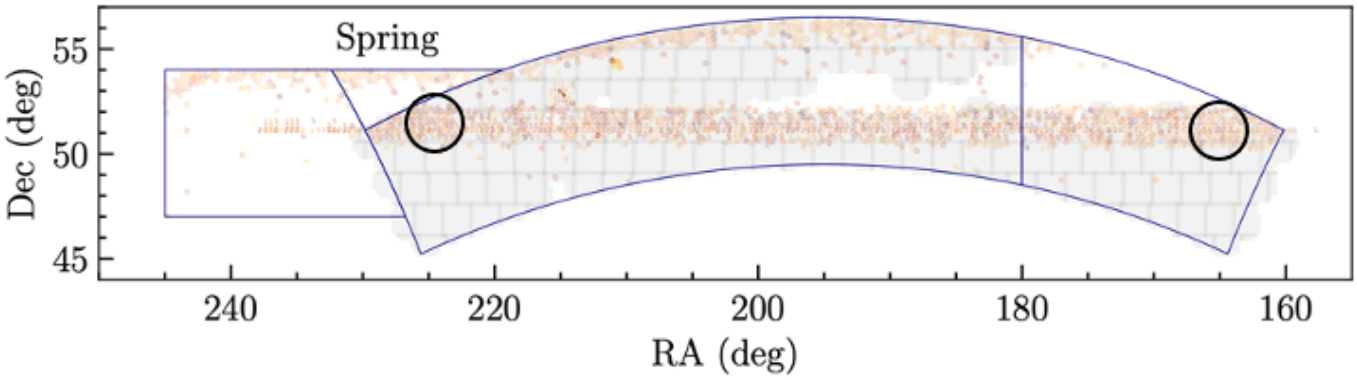

Figure 1. HETDEX main (a.k.a "spring") field outlined in blue, with completion status as indicated by orange tiles as of January 2021 when the target list for the sample was selected (the grey tiles were planned but yet to be observed). The two circles show the position of the two DESI tiles chosen for the follow-up.

Based on these criteria, HETDEX provided DESI with 10,780 possible targets. DESI then positioned fibers on 1633 (out of 5335) in Pointing 1 and 1516 (out of 5448) objects in Pointing 2, for a total of 3149 targets. This selection was based on fiber positioning requirements and fiber availability. The distribution of selected targets with valid fibers for the two pointings is shown in Figure 2.

The data were taken between 13 March and 13 May 2021 and are summarized in Table 2. Note that the total effective exposure times of 18944.0 sec for tile 80869 and 11306.4 sec for tile 80870 are considerably longer than the nominal 1000 seconds used for regular DESI dark time tiles (though some targets have less accumulated time, due to fiber placement inaccuracies in some exposures). The reason for the cumulative exposure for tile 80869 was $\sim 60\%$ longer than that for tile 80870 is that time was initially allocated for only one of the two fields; towards the end of the survey validation period (SV; DESI Collaboration et al. 2024a), it was decided to expand the secondary target programs as much as time allowed before starting the main survey. The spectra obtained through these observations are available as part of the DESI early data release (DESI Collaboration et al. 2024b), and the forthcoming first data release (DR1) (DESI Collaboration 2025; DESI Collaboration et al. 2024c,d,e,f,g,h).

3. VISUAL INSPECTION

The raw data were processed using the DESI spectroscopic pipeline (Schlafly et al. 2023; Guy et al. 2023) and the spectra were visually inspected using the DESI VI tool `prospect`[1]. Sources were classified into five categories, numbered 0 to 4, based on the confidence for which the redshift could be determined, with 4 being the highest confidence (for details of the DESI visual in-

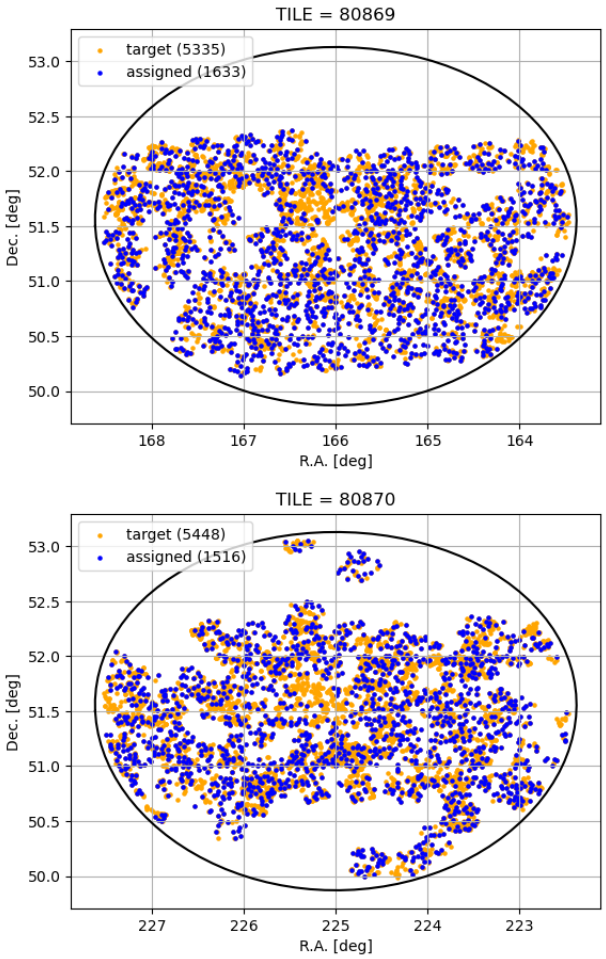

Figure 2. Parent and assigned target distributions for the two pointings. The numbers 80869 and 80870 refer to the DESI tileID, which is unique for the combination of pointing and fiber assignments.

[1] https://github.com/desihub/prospect/

DESI tileID	RA (deg)	DEC (deg)	$N_{\rm targets}$	$N_{\rm nights}$	Exposures (sec)	$\rm ExpTime_{eff}$ (sec)
80869	165	51.5	1959	4	2×60 9×900 7×1220	18944.0
80870	225	51.5	1826	3	1×900 2×1800 2×1820	11306.4

Table 1. Summary of exposures. The exposures column lists the number and duration of the exposures on the respective fields, and $\rm ExpTime_{eff}$ denotes the total effective exposure time.

spection procedure and results from SV, see Lan et al. (2023) and Alexander et al. (2023)).

The analysis presented here used the data reductions of Data Release 1. Earlier reductions of the spectra are available as part of the DESI Early Data Release (EDR, DESI Collaboration et al. 2024b)

3.1. *Preliminary cuts*

After excluding spectra due to fiber-positioning errors, several additional cuts were made before determining the accuracy of the HETDEX line classification algorithm.

Some HETDEX sources had their assumed Lyα line too close to the blue end of the DESI spectral range for an effective confirmation. We therefore excluded those HETDEX objects whose emission line wavelength was $\lambda \leq 3610$ Å. A further restriction was imposed to require HETDEX sources to have good IFU coverage as there is evidence for an increase in the HETDEX false positive rate for objects located near the edges of an IFU (`APCOR_DEX` > 0.6, where `APCOR_DEX` is the fraction of area covered by VIRUS fibers within a $3\farcs5$ aperture radius centered on the detection). We also restricted the sample to sources that were in both the final HETDEX second data release and in the third release, which used new, improved reduction algorithms. After these criteria were applied, our final sample contained 2374 sources, of which 1157 (49.0%) were detected with confidence by DESI, i.e., with `VI_QUALITY` ≥ 3. In addition, the DESI spectra of several high-redshift sources show peaks of emission at the same wavelength as in the HETDEX spectra, however the S/N of these objects is very poor, and the line is confirmed at lower `VI_QUALITY` (1 or 2).

3.2. *Visual Inspection*

The DESI automated redshift assignment algorithm proved to be quite poor for most HETDEX emission-line objects. Since the majority of the sample does not have significant continuum emission, many of the spectral template fits either failed or found spurious emission lines in the red. The result was generally an erroneous redshift, even when it was returned with a high confidence. Thus, redshifts are assigned based on visual inspection. If multiple line identifications are visible, a highly confident redshift can be assigned to the source. This is the case for low-redshift objects, where multiple spectral lines are available in the DESI spectral range. An example of such a source can be seen in the top panel of Figure 3. In addition, for low-redshift, emission-line sources, a zoom in on the [O II] doublet reveals a double peak in the DESI spectra providing a high confidence redshift (`VI_QUALITY`=4).

In contrast, the majority of HETDEX LAE candidates are single emission line sources with no continua. An example is given in the bottom panel of Figure 3. Even with the increased sensitivity of the DESI spectrograph, the continuum level of the LAEs is not detected. No absorption features are available to secure the redshift and, in most cases, no other emission lines are available. If no other (lower redshift) solution fits the spectrum, we assume the emission line is due to Lyα, and we tune the redshift accordingly. In the case of a single emission line source, we assign a `VI_QUALITY` of 3. If other emission lines are detected, such as C IV or C III], then a `VI_QUALITY` of 4 is assigned. In the case of a very noisy spectrum in which a peak in emission matches the HETDEX redshift, we tune the Lyα line to this peak and assign a lower `VI_QUALITY` of 1 or 2 depending on the relative strength of the line to the noise. Emission line galaxies blueward of 3800 Å are difficult to confirm with DESI spectroscopy due to a number of spiky noise features.

Two independent classifications were made on each member of the sample. When a disagreement in classification occurred, a re-assessment was made on the source such that both classifiers were in agreement. Out of over 2500 classifications, only 15 re-assessments were made.

3.3. *Results*

The results of our HETDEX-DESI comparison are summarized in Table 2. Objects are separated into different groups, labeled `SOURCE_TYPE` in the catalog accompanying this paper, based on the spectroscopic red-

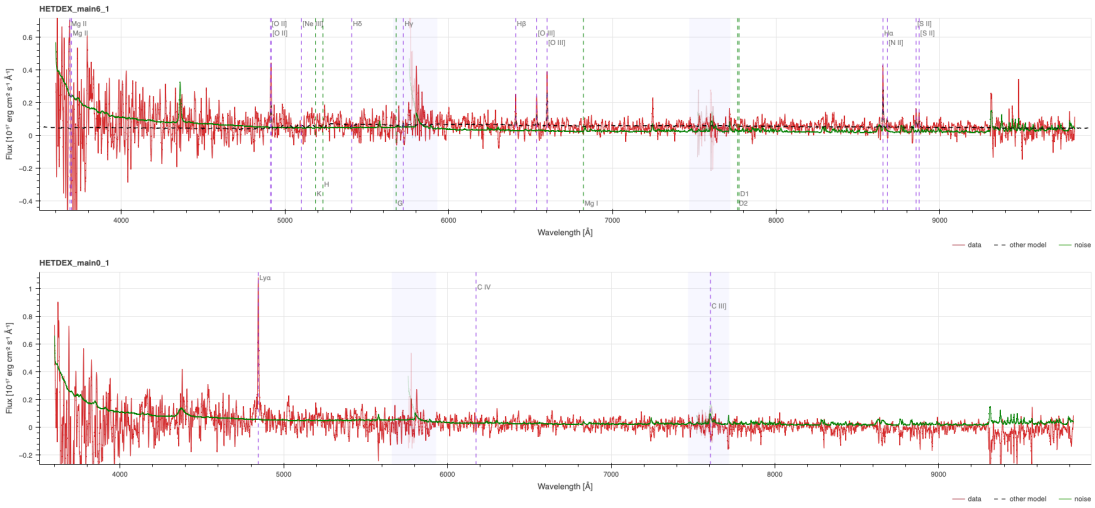

Figure 3. Example DESI spectra of a faint [O II] galaxy at $z = 0.318$ (top) and an LAE at $z = 3.006$ (bottom) confirmed by visual inspection.

Type	VI quality≥0	VI quality≥ 3	VI quality≥4
ALL	2374	1157	303
LAE	2077	982	137
[O II]	241	131	131
AGN	36	32	24
Other	17	12	11

Table 2. Summary of results of visual inspection

shifts obtained from Visual Inspection (`VI_Z`). For objects without spectroscopic redshifts from VI (`VI_Z`=-1, `VI_QUALITY`=0), the HETDEX catalog indicates the source type as described in Mentuch Cooper et al. (2023). The object groups are as follows:

- LAE: A `VI_Z` > 1.9 emission-line source not previously identified as an AGN in the HETDEX AGN catalog (Liu et al. 2022).
- AGN: An emission-line source previously identified as an AGN candidate in Liu et al. (2022).
- OII: An [O II] emission-line galaxy at $z < 0.5$.
- Other: A confirmed redshift of a non-emission line source or an intermediate redshift object with emission lines beyond the HETDEX spectral range of 3500–5500 Å. These are usually blended sources that do not confirm the HETDEX emission-line detection. Often, the continuum in these sources would not be detected in HETDEX.

All [O II] sources are confirmed with high confidence (`VI_QUALITY`=4). In all spectra, other emission lines (such as Hα, Hβ, and [O III]) confirm the spectroscopic redshift and the [O II] doublet at 3727.092 Å and 3729.875 Å is resolved. For the AGN sample, 24/32 sources are confirmed with multiple emission lines. Those that are confirmed with only one broad line and do not show other spectral features and are assigned `VI_QUALITY`=3. These are candidate Lyman Alpha Blobs (LABs) that were misidentified as AGN due to the broad nature of the line. They exhibit extended Lyα emission and little to no continua (more information on LABs in HETDEX is presented in Mentuch Cooper et al., in prep).

A number of LAE candidates (137/982) are confirmed with high confidence based on the presence of additional spectral lines (generally C IV and C III]). Although not previously identified as AGN, these sources exhibit signatures of an AGN contribution. Either the additional lines are relatively faint or they fall outside the spectral window of the VIRUS spectrographs.

Figure 4 shows the distribution of sources from the input target sample (grey) and the recovered sample (blue) as a function of S/N and $g_{\rm HETDEX}$ magnitude on the top and bottom panels, respectively. In the top panel, we include the cumulative recovery fraction of sources (confirmed with `VI_QUALITY`≥ 3 above a given S/N threshold). In grey, we plot the cummulative recovery without any sub-selection of the sample, while in dashed-black, we select only those targets with emission lines redward of 3800 Å which are not flagged by HETDEX for various reasons (i.e., we exclude internally flagged detections, `DEX_FLAG`= 0, objects).

The distribution of visual mangitude is provided in the bottom panel of Figure 4. The targeted sample of LAEs is biased towards slightly higher $g_{\rm HETDEX}$ values compared to the LAE sample presented in the HETDEX public catalog in Mentuch Cooper et al. (2023).

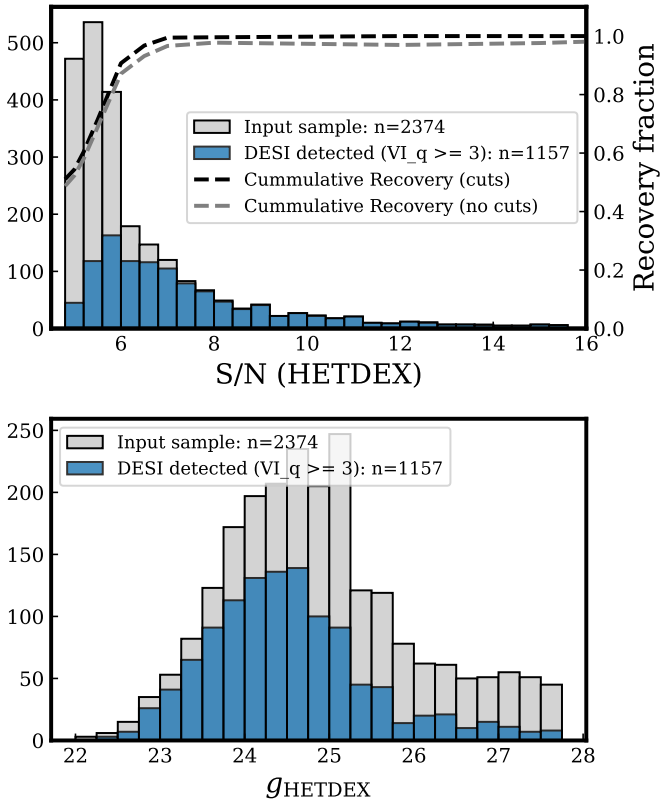

Figure 4. The top panel shows the distribution of emission-line detection S/N for HETDEX targets. The grey histogram indicates the range in the input target sample, while the blue shows the recovered galaxies with VI_quality$\geq$ 3. The cumulative recovery fraction as a function of signal-to-noise ratio is shown in dashed grey for the full sample and in dashed black line where some cuts are applied. Only well positioned sources with emission lines $\lambda > 3800$ are included in this sample and those that have not been excluded from the sample in later HETDEX catalogs (ie. we only include those emission lines with `DEX_FLAG==1`). Sources at low S/N are missed due to astrometric uncertainties for low S/N sources and possible false positives. The bottom panels presents the distribution in $g_{\rm HETDEX}$ magnitudes of the input and recovered sample, in grey and blue respectively. Broadband continuum estimates fainter than $g_{\rm HETDEX}\sim$ 24.5 are below the sensitivity of HETDEX sources and are unreliable.

A sub-sample of brighter LAE candidates were selected in this study to test the HETDEX classification method on LBG-like Lyα emitting galaxies that can potentially be confused with fainter, low-redshift [O II]-emitters.

3.4. *Missing sources*

At high HETDEX signal-to-noise ($S/N > 6.5$), 97% (448/463) of the sample is recovered when we account for only those sources in the latest HETDEX catalog (`DEX_FLAG==1`), whose TARGET fiber position is well matched to the HDR3 detection (`SEP`$< 0\farcs3$) and are redder than 3800 Å. After visual inspection, roughly half the sources are likely false positives in the HETDEX sample. Most of these are higher line width objects that are excluded in most HETDEX analyses.

However, four of these objects are missed by DESI simply because they happen to be related to a brighter resolved object, but not centered on that object. The HETDEX detection search algorithm described in Gebhardt et al. (2021) is designed to search for point sources, but in the case of extended objects, multiple detections can arise for the same source. Bright point sources such as stars and AGN also result in multiple HETDEX detections. When the target list was created, detection grouping to merge a set of multiple redundant detections into single HETDEX source (as described in Mentuch Cooper et al. 2023), was not yet developed. For this reason, some DESI fibers were not centered on their HETDEX target. With its single fiber and better image quality, DESI missed the object.

At lower HETDEX S/N the number of missing sources increases. Specifically, at a S/N threshold of 6, 93% of HETDEX sources are recovered by DESI, but by S/N=5.5, the cumulative recovery rate drops to 75% and to 56% at S/N=5.0. Simulations show that due to the large ($1\farcs5$-diameter) size of the HETDEX fibers and fiber-to-fiber throughput variations, the computed centroid of a low S/N source may be offset from its true position on the sky by over $1''$. Specifically, the simulations described in Gebhardt et al. (2021), as well as comparison to independent emission line samples, show that at $S/N > 5.5$, the median centering accuracy of a HETDEX source is about $0\farcs4$ with up to 10% of sources recovered at over $1''$ away. Moreover, at lower S/N, this astrometric error quickly increases, so that at the limit of HETDEX detections, the median centering accuracy is $0\farcs6$, and over 20% of sources are located more than $1''$ away from their input position. As the DESI spectrograph uses single $1\farcs5$ fibers and has better image quality than HETDEX, it is likely that if positioned off center from the true position of an LAE, the signal would be lost completely.

False positive contamination in the HETDEX sample also contributes to the increase in missing sources at lower HETDEX S/N. At S/N values above 5.5, false detections in the HETDEX database are less than 10% (Mentuch Cooper et al. 2023). However, when integrated over all S/N, the false positive rate is likely between 15–30%(Gebhardt et al. 2021). Unfortunately, due to the centering uncertainties described above, the DESI spectroscopy cannot be used to better define this value. More generally, DESI cannot say anything about

these missing sources and are not used in the subsequent analysis.

4. HETDEX CLASSIFICATION

The details of the spectral classification methodologies employed by HETDEX are presented in Davis et al. (2023) and Mentuch Cooper et al. (2023), and only a brief overview is provide here.

For the HETDEX wavelength window and depth, often only a single emission line is visible: generally Lyα ($1.9 < z < 3.5$) or [O II] ($0.13 < z < 0.5$). Moreover, because the resolution of the VIRUS spectrograph is only $\lambda/\Delta\lambda \sim 800$, the [O II] doublet is not resolvable in the HETDEX data (Gebhardt et al. 2021; Davis et al. 2023). It is therefore necessary to use on statistical methods to distinguish these lines.

The primary classification code that performs this analysis is the Emission Line eXplorer (Davis et al. 2023, ELiXer) which builds on the Bayesian analysis in Leung et al. (2017) with modifications in Farrow et al. (2021). ELiXer is a significant improvement over the commonly used Lyα-restframe 20 Å equivalent width threshold for Lyα and [O II] emission lines (Gronwall et al. 2007; Adams et al. 2011), as it takes advantage of additional information, both in the HETDEX spectra and, when available, in external photometric and spectroscopic catalogs. In brief, ELiXer enhances the simple 20 Å $W_{Ly\alpha}$ cut with a weighted voting scheme. This scheme incorporates various "votes" for or against individual or sets of possible classifications based on different elements. Each vote is assigned a weight derived from the frequency with which it aligns with the correct classification, given a training set of spectra with known redshifts. The votes are summed with their weights to determine the likely classification in reference to a configurable P(Lyα) threshold value, a normalized value between 0 and 1 (where 0 = not likely Lyα and 1 = likely Lyα). In addition, a set of specialized conditions are evaluated and may modify the redshift or enhance the classification (Section 3 of Davis et al. 2023).

As one purpose of this work is to evaluate the performance of the HETDEX spectral classifications on their own merit, the use of all external phot-z and spec-z catalogs is disabled. Imaging cutouts from external photometric archives are still used for bandpass aperture fluxes, as are object sizes as measured by the ELiXer code. These are the same testing conditions imposed in Davis et al. (2023).

The contamination and recovery fractions of the LAE sample as a function of the P(Lyα) threshold are presented in Table 3 and Figure 5. The data clearly show the trade-off between the two parameters. Here, we define contamination as the number of [O II] emission lines that were incorrectly classified by ELiXer as Lyα divided by the total number of Lyα lines (as classified by ELiXer). [O II] λ3727 is the most likely contaminant in the HETDEX Lyα sample and therefore has the most impact on the measurement of the $z \sim 2.5$ galaxy clustering signal (Gebhardt et al. 2021; Davis et al. 2023). The recovery fraction (represented as a percent in the figure) is the number of correctly classified Lyα lines, again divided by the total number of Lyα lines in the sample. As described above, this test sample is limited to the visually vetted subset of objects with `VI_QUALITY` ≥ 3.

The results shown in Figure 5 are consistent with those in Davis et al. (2023), and supports the efficacy and robustness of the HETDEX classification mechanism. In particular, the distribution of the objects in the earlier work: (*1*) skews towards much brighter objects (see Figure 7 in Davis et al. (2023) and Figure 4 in this work) and (*2*) has a smaller fraction of Lyα sources (24% in Davis et al. (2023) vs. 85% here).

The default P(Lyα) setting of 0.5 yields a good recovery of Lyα sources with a contamination fraction due to low-redshift sources mis-classified as Lyα below the 2% science requirement specified by HETDEX (Gebhardt et al. 2021). With some additional decontamination methodologies, HETDEX may be able to sustain larger contamination, perhaps even above 5% (Farrow et al. 2021). If so, there is opportunity to move to a lower P(Lyα) threshold and recover additional Lyα sources at the expense of a modest increase in contamination.

Figure 6 shows a comparison between redshifts based on the HETDEX classification code using the P(Lyα)=0.5 threshold and redshifts determined from the visual inspection of DESI spectra. This value of P(Lyα) represents a trade-off between [O II] contamination and Lyα recovery: lowering its value increases the rate of LAE recovery (i.e., it reduces the number of points in the bottom-right region of the figure) at the expense of greater contamination (adding points in the upper left). Of the single broad-line AGN classified by ELiXer as Lyα, 8 were shown by DESI spectra to be C IV λ1549 based on the presence of secondary lines such as C III] λ1909 and Mg II λ2800. Both these lines fell outside the HETDEX wavelength range.

5. EMISSION LINE COMPARISONS

In this section, we compare HETDEX and DESI measurements of emission-line flux and line width of the confirmed LAEs and [O II] emitters with `VI_QUALITY` ≥ 3. Any emission lines blueward of 3800 Å are excluded due to lower S/N in the DESI spectra.

P(Lyα) Threshold	[O II] Contamination	Lyα Recovery
0.9	0.97%	91.4%
0.8	0.97%	91.5%
0.7	0.97%	91.6%
0.6	0.97%	91.7%
0.5	**1.15%**	**94.4%**
0.4	2.13%	95.8%
0.3	2.71%	96.2%
0.2	3.72%	97.6%
0.1	6.67%	98.6%

Table 3. Summary of the results of HETDEX ELiXer classification of visually vetted spectra, see Figure 5. The Contamination and Recovery columns for each row are independently computed using that row's P(Lyα) Threshold. The row in **boldface** shows the default ELiXer classification threshold configuration. The results are consistent with those in Davis et al. (2023) despite differences in the distributions of the magnitudes and emission line identifications.

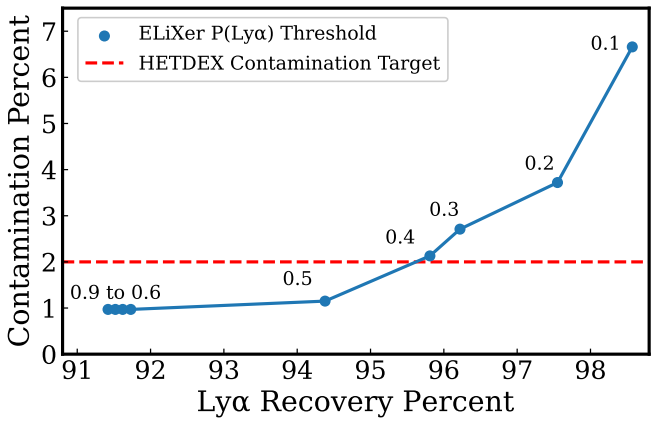

Figure 5. The results of our comparison of ELiXer classifications of HETDEX spectra with the high confidence visually vetted DESI spectra. The y-axis gives the percent contamination of the HETDEX LAE sample by misidentified [O II] galaxies, while the percentage of LAEs recovered by the classifier is on the x-axis. The labeled points show the corresponding P(Lyα) threshold used to configure the ELiXer classifier. The red horizontal dashed line represents the requirements of the HETDEX survey (Gebhardt et al. 2021), though higher contamination rates may be acceptable with the use of additional decontamination methodologies (Farrow et al. 2021). Despite differences in the redshift and magnitude distributions with the Spectrocopic-z Assessment Sample (SzAS) in Davis et al. (2023), the curve is very similar.

To minimize systematic errors in our line measurements, we use the same algorithms for both groups of spectra. The input to the line-fitting algorithm are the

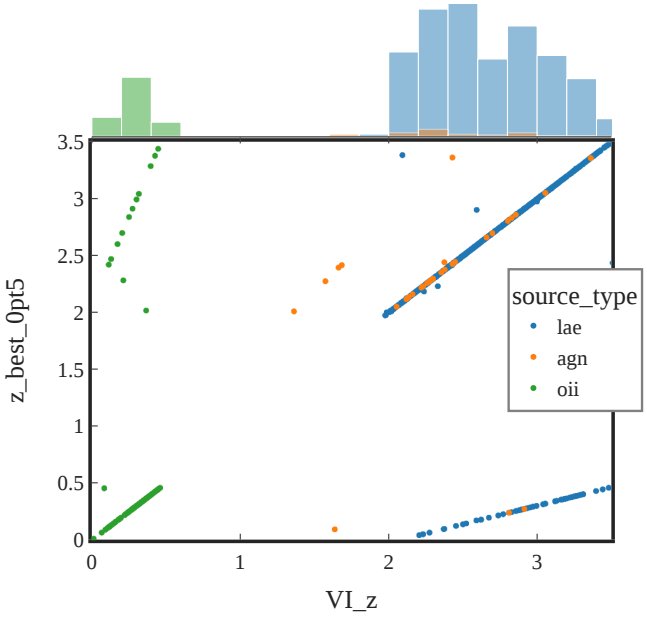

Figure 6. Comparison of redshifts obtained from visual inspection, `VI_Z`, of DESI spectra (x-axis) and those derived from HETDEX data (y-axis) using ELiXer. The plot assumes a P(Lyα)=0.5 threshold for the HETDEX line classifications. As discussed in Section 4, this value leads to a 1.15% contamination of the LAE sample by [O II] galaxies, and a loss of 5.6% true LAEs from the HETDEX sample.

DESI and HETDEX spectra, reduced by their respective collaboration pipelines; for HETDEX the pipeline used was that associated with the internal HDR3. After the reduction, the analyses for both spectra were identical.

We model the line emission using the software package `mcmc_gauss.py`, developed within HETDEX's software repository, `hetdex-api`[2]. In this method, a single Gaussian model is fit to the spectral line using the publicly available `emcee`[3](Foreman-Mackey et al. 2013), which implements Goodman & Weare's Affine Invariant Markov chain Monte Carlo (MCMC) Ensemble sampler. For the initial guess, we use the wavelength, flux, and line width as determined by the HETDEX pipeline. For determining the line properties of the DESI spectra, we limit the range to ±50 Å of the known line position. This is sufficient to obtain reliable fits for the majority of targets.

Figure 7 shows an example of four fitted lines of LAEs. The left column shows the higher spectral resolution and more sensitive DESI spectroscopy; the right column displays the corresponding spectra obtained by HETDEX.

[2] https://github.com/HETDEX/hetdex_api/blob/master/hetdex_tools/mcmc_gauss.py

[3] https://emcee.readthedocs.io/

Table 4. Classification and line fit parameters. The first 10 rows of the 2374 row catalog associated with this paper are printed for demonstrative purposes. The full version can be found online. In the Appendix, a full description of the FITS file released with this paper can be found found in Table 5 with a description of each column described in Table 6. All wavelength values are reported in air, in units of Å. Line width values, σ, are reported in Å. The HETDEX values are not corrected for instrumental broadening. The flux values reported are in 10^{-17} erg s^{-1} cm^{-2} Å^{-1}.

TARGETID	TARGET_RA	TARGET_DEC	TILEID	RA_HETDEX	DEC_HETDEX	SEP	DETECTID
103606502031361	224.490280	50.115120	80870	224.4905	50.1151	0.4960	3007922753
103606506225670	224.752151	50.046459	80870	224.7521	50.0464	0.3578	3007918479
103606506225674	224.588257	50.011024	80870	224.5883	50.0110	0.1413	3007918678
103606506225675	224.637772	50.085388	80870	224.6378	50.0854	0.2018	3007919753
103606506225676	224.721085	50.040512	80870	224.7212	50.0406	0.2334	3007918756
103606506225680	224.657257	50.085510	80870	224.6575	50.0855	0.4738	3007919755
103606506225682	224.762589	50.021271	80870	224.7626	50.0212	0.2984	3007918514
103606506225685	224.833771	50.057350	80870	224.8338	50.0574	0.0017	3007918908
103606506225686	224.667603	50.055523	80870	224.6678	50.0555	0.4585	3007919332
103606506225687	224.779495	50.102844	80870	224.7795	50.1028	0.0045	3007921868

VI_Z	VI_QUALITY	SOURCE_TYPE	DEX_FLAG	Z_BEST_0PT3	Z_BEST_0PT4	Z_BEST_0PT5	AC	GMAG	SN_HETDEX
-1.0000	0	lae	1	2.2106	2.2106	2.2106	0.0469	25.8217	5.9500
3.4182	3	lae	1	3.4182	3.4182	3.4182	0.0518	24.8484	5.3700
-1.0000	0	lae	1	2.3794	2.3794	2.3794	0.0526	25.0488	5.1200
-1.0000	0	lae	0	2.6206	2.6206	2.6206	0.0530	24.0520	5.3300
2.9116	3	lae	1	0.2756	0.2756	0.2756	0.0528	22.9939	6.8700
-1.0000	0	lae	1	2.2066	2.2066	2.2066	0.0536	25.4402	4.8692
3.0000	3	lae	1	2.9997	2.9997	2.9997	0.0514	26.0349	7.6200
-1.0000	0	oii	1	0.4678	0.4678	0.4678	0.0499	23.6870	5.2800
-1.0000	0	lae	1	2.1348	2.1348	2.1348	0.0539	24.8588	5.8300
2.7621	4	lae	1	2.7623	2.7623	2.7623	0.0496	26.4235	6.2300

WAVE_HETDEX	WAVE_ERR_HETDEX	FLUX_HETDEX	FLUX_ERR_HETDEX	SIGMA_HETDEX	SIGMA_ERR_HETDEX	CONT_HETDEX	CONT_ERR_HETDEX
3901.9700	0.6000	13.9400	3.1300	1.8600	0.5800	-0.1400	0.1800
5369.7100	0.5100	7.8600	1.4600	1.9700	0.5100	0.0050	0.1000
4107.1602	0.8300	14.5900	2.9500	3.5400	0.7600	0.0450	0.1500
4400.2500	0.4900	8.3600	2.1800	1.6500	0.6500	0.0650	0.1300
4755.1802	0.6500	15.4400	2.1400	4.4000	0.7400	0.1850	0.1000
3897.0801	0.5600	12.1500	2.8200	2.0200	0.7000	-0.1300	0.1800
4861.0498	0.3900	16.0500	1.4600	3.8000	0.4000	0.0050	0.0800
5471.7100	0.4000	8.1600	1.2100	2.0800	0.4000	0.0650	0.0800
3809.8501	0.2900	15.3600	1.9700	1.8100	0.2700	-0.1500	0.1500
4572.5400	0.5600	11.6800	1.7800	2.9600	0.5200	-0.0300	0.1000

WAVE_DESI	WAVE_ERR_DESI	FLUX_DESI	FLUX_ERR_DESI	SIGMA_DESI	SIGMA_ERR_DESI	CONT_DESI	CONT_ERR_DESI
3901.0369	3.7251	0.9410	0.8899	6.0968	3.4219	-0.0795	0.0239
5369.4976	0.1637	8.3260	0.5096	1.9146	0.1712	0.0192	0.0108
4104.7891	3.3367	0.8758	0.7595	5.7900	3.1882	0.0095	0.0174
4398.3242	3.5279	0.7357	0.6811	5.8236	3.2706	-0.0131	0.0155
4754.4004	0.2762	10.2795	0.7797	2.7820	0.3651	0.1109	0.0127
3895.1643	3.6269	0.6176	0.6168	5.4959	3.5032	0.0786	0.0236
4861.7876	0.3043	12.7067	0.7819	3.7228	0.3496	-0.0088	0.0113
5470.2783	3.6197	0.6026	0.5033	5.9930	3.1566	-0.0087	0.0111
3806.3154	3.5183	1.4814	1.4153	6.2903	3.2750	-0.1125	0.0326
4572.7666	0.4182	7.2106	0.9677	2.6546	0.6306	0.0425	0.0152

The y-axis is fixed to the same scale limits to demonstrate differences in absolute flux. For extended sources, fluxes can vary significantly, as the HETDEX spectra are derived from an IFU while DESI only captures light from a single 1.″5 fiber, and no aperture corrections for extended emission are applied.

The bottom example shows a double-peaked Lyα line profile. With the higher spectral resolution of DESI, a single Gaussian model is a poor fit to the total emission. In contrast, the HETDEX spectrum with its higher noise, fits both peaks to a single Gaussian and consequently obtains a higher line width, $\sigma_{\rm HETDEX}$.

5.1. *Line Fluxes*

Figure 8 shows the ensemble comparison of Lyα line fluxes between the two surveys. In the bottom panel,

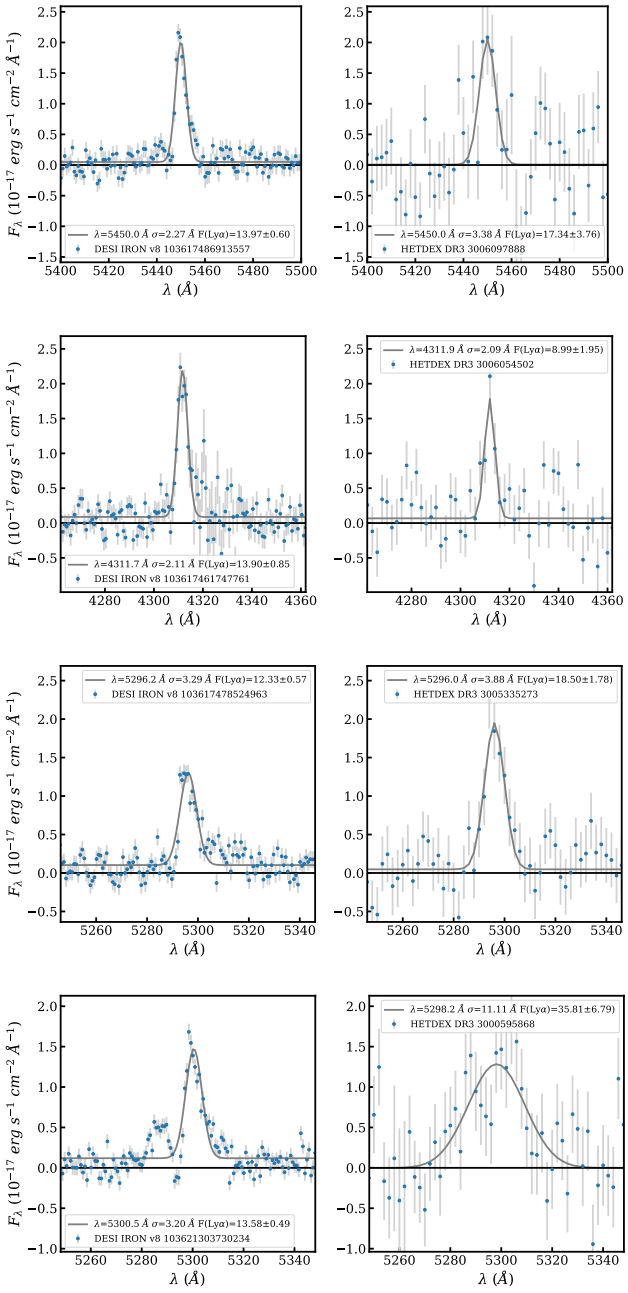

Figure 7. Examples of fitted Lyα lines for DESI on left and HETDEX on right. Each row represents a single LAE. The Lyα flux values in the legends are reported in units of 10^{-17} erg s^{-1} cm^{-2} Å^{-1}. In the bottom example, the line is double peaked; as a result, the fitted linewidth is higher for HETDEX than for the DESI line fit which only fits the brightest component of the line.

we show error-weighted difference between the two measures as a function of HETDEX flux values ($(F_{DESI} - F_{HETDEX})/\sqrt{\sigma^2_{DESI} + \sigma^2_{HETDEX}}$). Generally, HETDEX flux values are found to be higher than the DESI values, with the difference growing larger with increased

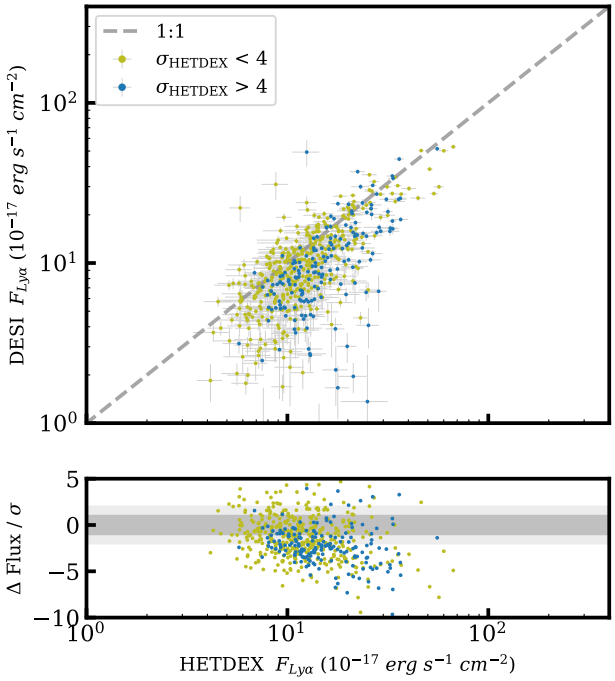

Figure 8. Line flux comparison for the 982 LAEs detected by HETDEX confirmed by DESI (`VI_QUALITY`$\geq$ 3) and the error–weighted difference is shown in the bottom panel. Data points are separated based on HETDEX line width values, `SIGMA_HETDEX`. As demonstrated in the bottom panel of Figure 7, asymmetric and double-peaked line profiles are fit with higher line widths in the HETDEX spectra resulting in higher Lyα flux values.

flux. Since no aperture corrections are applied to the DESI data, this offset is not surprising. At higher line fluxes, many LAEs are extended and flux will be lost when observed with a single fiber. In addition, in some cases where the line is double-peaked, the higher spectral resolution DESI data only fits one peak and a significant fraction of the Lyα line flux is lost. This latter effect is illustrated in the bottom panel of Figure 7.

A significant contribution to the scatter also comes from centering uncertainty. As discussed in Section 3.4, simulations show that the ability for the HETDEX detection pipeline to find the true position of a source decreases in accuracy as S/N decreases. For robust emission line sources detected at $S/N > 5.5$, up to 10% of sources are found more than 1″ away from the input position. Lower S/N emission lines are much more difficult to locate their precise center. Over 20% of simulated sources have output positions over 1″ away from their input position. With just a single DESI fiber placed on the source, the position uncertainty translates to a flux loss in the target DESI fiber.

Positional uncertainty contributes to the scatter in the measured line fluxes between the two surveys and it also

contributes to large scatter in the measured spectral continuum. Using spectral data from both surveys, a common pseudo-magnitude value can be derived for both datasets by convolving a g-band filter curve with the spectral data. On average the HETDEX magnitudes are brighter ($\langle$ $g_{\rm HETDEX}$ - $g_{\rm DESI}\rangle$ = -0.66 mag) and exhibit an rms scatter of 1.15 mag. Generally, the continuum flux of HETDEX LAEs are too faint to be detected by VIRUS which reaches a sensitivity of $g_{\rm HETDEX}\sim$ 25 mag so in most cases the HETDEX continuum measures are an upper limit.

5.2. *Wavelengths and Redshifts*

A comparison between the best-fit central wavelengths calculated for each experiment is shown in Figure 9. DESI spectral data are converted to wavelength measurements in air to match those of HETDEX. Values are derived from line fitting each spectrum using a single Gaussian model. The two values agree very well between the two surveys with an average offset of $\langle\lambda\rangle$ = 0.15 Å and a rms scatter of 1.17Å.

We can also consider a comparison to the best-fit visual wavelength. During the VI process, the classifier tunes the location of the Lyα line. The classifier must decide whether they are fitting to the peak of the Lyα emission or instead to the middle of the emission when an asymmetric profile shape is clearly seen. This decision will lead in some scatter relative to the wavelength determined from automated fitting and consequently some scatter in the spectroscopic redshift determined for the LAE. The middle panel in Figure 9 converts the visual spectroscopic redshift, `VI_Z`, to a Lyα wavelength value in the observe frame and compares to the best-fit DESI wavelength. These values are consistent within the scatter of 1.37 Å. The dispersion in the determined wavelength is higher in this case relative to the above comparison between the automated values from the two surveys.

The bottom panel presents a comparison between HETDEX spectroscopic redshifts and DESI VI spectroscopic redshifts. The HETDEX redshifts are the result of the classifcation pipeline described in Section 4 and come from fitting a single Gaussian model to the emission line. The DESI redshifts are from visually tuning to the peak emission wavelength using the DESI VI tool. Ignoring gross mis-classifications ($\Delta z > 0.1$), the median difference between the two values is $\langle\Delta z/(1+z)\rangle = 6.87\times10^{-5}$ with an rms scatter of 3.32×10^{-4}.

5.3. *Linewidths*

Figure 10 presents a comparison between the Gaussian line widths measured by each survey. The HET-

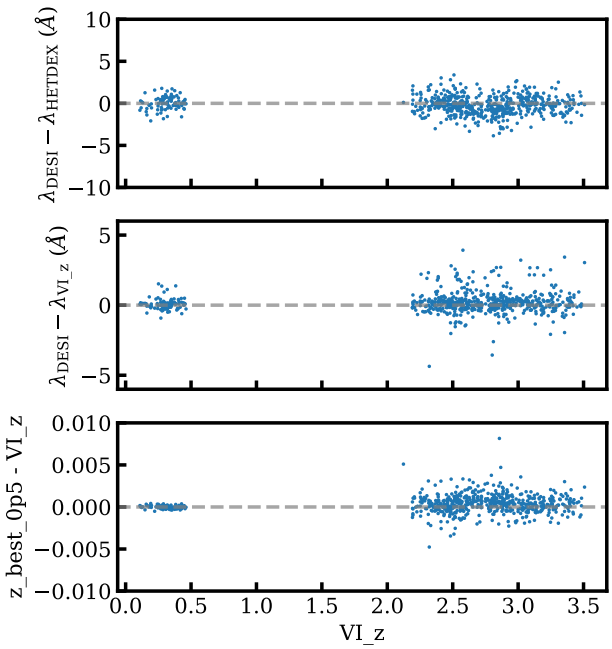

Figure 9. A comparison between the best-fit central wavelengths calculated for each experiment. Both use a single Gaussian model and are fit with their wavelengths as measured in air. The middle panel converts the visual spectroscopic redshift, `VI_Z`, to a Lyα wavelength value redshifted to the observed frame and compares to the best fit wavelength in the DESI data. The bottom panel presents a comparison between HETDEX spectroscopic redshifts and DESI VI spectroscopic redshifts.

DEX value, $\sigma_{\rm HETDEX}$, is corrected for the VIRUS instrument's low spectral resolution ($R\sim800$). The values reported in the data table of this paper contain the measured, uncorrected values. Once corrected, the line width values agree.

While there is some scatter, the two projects agree well about the basic properties of the detected emission lines. The flux and line width ratios are within 1-σ of 1 and the error weighted differences are within 1 standard deviation of 0.

$$\begin{aligned}
\frac{flux(DESI))}{flux(HETDEX)} &= 0.853\pm0.391\\
\frac{flux(DESI)-flux(HETDEX)}{\sqrt{\sigma^2_{f,\ DESI}+\sigma^2_{f,\ HETDEX}}} &= -1.023\pm2.441\\
\frac{width(DESI)}{width(HETDEX)} &= 0.859\pm0.550\\
\frac{width(DESI)-width(HETDEX)}{\sqrt{\sigma^2_{w,\ DESI}+\sigma^2_{w,\ HETDEX}}} &= -0.71\pm1.79
\end{aligned} \quad (1)$$

6. CONCLUSION

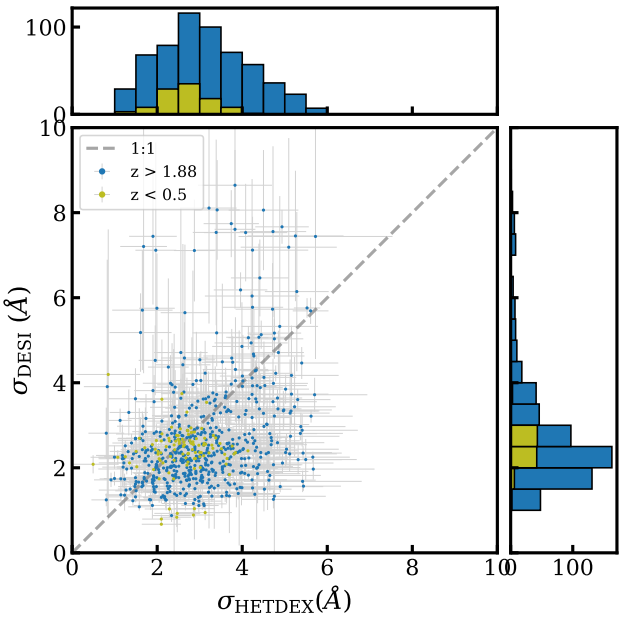

Figure 10. Comparison of Gaussian line width, σ, values between DESI (y-axis) and HETDEX (x-axis). Both line profiles are modeled using a single Gaussian model. HETDEX values are corrected for the lower instrumental resolution of the VIRUS spectrograph compared to the DESI spectrograph. The values show large scatter at higher line width where the difference in spectral resolution can lead to fitting of only one peak in the double-peaked and asymmetric lines found in the LAE sample.

Observations of HETDEX LAE candidates with DESI has enabled the validation of the HETDEX line classification method in separating Lyα from [O II] emitters and demonstrates that the HETDEX survey achieves the science requirement of less than 2% interloper contamination. For those sources that are confidently detected in DESI, the derived central wavelength of the emission line, their line fluxes and line widths are in good agreement between the two experiments. The spectroscopic redshifts obtained from each survey, when the classifications agree, are in very good agreement with only a small median offset between the two surveys of $\langle \Delta z/(1+z) \rangle = 6.87 \times 10^{-5}$ with an rms scatter of 3.32×10^{-4}.

ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy (DOE), Office of Science, Office of High-Energy Physics, under Contract No. DE–AC02–05CH11231, and by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract. Additional support for DESI was provided by the U.S. National Science Foundation (NSF), Division of Astronomical Sciences under Contract No. AST-0950945 to the NSF's National Optical-Infrared Astronomy Research Laboratory; the Science and Technologies Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National Council of Science and Technology of Mexico (CONACYT); the Ministry of Science and Innovation of Spain (MICINN), and by the DESI Member Institutions: https://www.desi.lbl.gov/collaborating-institutions. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. National Science Foundation, the U.S. Department of Energy, or any of the listed funding agencies.

The authors are honored to be permitted to conduct scientific research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

HETDEX is led by the University of Texas at the 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, and the US Air Force (AFRL FA9451-04-2-0355) and receives generous support from private individuals and foundations.

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

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

REFERENCES

Adams, J. J., Blanc, G. A., Hill, G. J., et al. 2011, ApJS, 192, 5, doi: 10.1088/0067-0049/192/1/5

Adams, J. J., Blanc, G. A., Hill, G. J., et al. 2011, ApJS, 192, 5, doi: 10.1088/0067-0049/192/1/5

Alexander, D. M., Davis, T. M., Chaussidon, E., et al. 2023, AJ, 165, 124, doi: 10.3847/1538-3881/acacfc

Davis, D., Gebhardt, K., Cooper, E. M., et al. 2023, ApJ, 946, 86, doi: 10.3847/1538-4357/acb0ca

DESI Collaboration. 2025, Data Release 1 of the Dark Energy Spectroscopic Instrument

DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016a, arXiv e-prints, arXiv:1611.00036. https://arxiv.org/abs/1611.00036

—. 2016b, arXiv e-prints, arXiv:1611.00037. https://arxiv.org/abs/1611.00037

DESI Collaboration, Abareshi, B., Aguilar, J., et al. 2022, AJ, 164, 207, doi: 10.3847/1538-3881/ac882b

DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2024a, AJ, 167, 62, doi: 10.3847/1538-3881/ad0b08

—. 2024b, AJ, 168, 58, doi: 10.3847/1538-3881/ad3217

—. 2024c, arXiv e-prints, arXiv:2411.12020, doi: 10.48550/arXiv.2411.12020

—. 2024d, arXiv e-prints, arXiv:2404.03000, doi: 10.48550/arXiv.2404.03000

—. 2024e, arXiv e-prints, arXiv:2404.03001, doi: 10.48550/arXiv.2404.03001

—. 2024f, arXiv e-prints, arXiv:2411.12021, doi: 10.48550/arXiv.2411.12021

—. 2024g, arXiv e-prints, arXiv:2404.03002, doi: 10.48550/arXiv.2404.03002

—. 2024h, arXiv e-prints, arXiv:2411.12022, doi: 10.48550/arXiv.2411.12022

Farrow, D. J., Sánchez, A. G., Ciardullo, R., et al. 2021, Monthly Notices of the Royal Astronomical Society, 507, 3187–3206, doi: 10.1093/mnras/stab1986

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

Gawiser, E., Francke, H., Lai, K., et al. 2007, ApJ, 671, 278, doi: 10.1086/522955

Gebhardt, K., Cooper, E. M., Ciardullo, R., et al. 2021, ApJ, 923, 217, doi: 10.3847/1538-4357/ac2e03

Gronwall, C., Ciardullo, R., Hickey, T., et al. 2007, ApJ, 667, 79, doi: 10.1086/520324

Guy, J., Bailey, S., Kremin, A., et al. 2023, AJ, 165, 144, doi: 10.3847/1538-3881/acb212

Hill, G. J., Lee, H., MacQueen, P. J., et al. 2021, AJ, 162, 298, doi: 10.3847/1538-3881/ac2c02

Lan, T.-W., Tojeiro, R., Armengaud, E., et al. 2023, ApJ, 943, 68, doi: 10.3847/1538-4357/aca5fa

Leung, A. S., Acquaviva, V., Gawiser, E., et al. 2017, ApJ, 843, 130, doi: 10.3847/1538-4357/aa71af

Levi, M., Bebek, C., Beers, T., et al. 2013, arXiv e-prints, arXiv:1308.0847. https://arxiv.org/abs/1308.0847

Liu, C., Gebhardt, K., Cooper, E. M., et al. 2022, ApJS, 261, 24, doi: 10.3847/1538-4365/ac6ba6

Mentuch Cooper, E., Gebhardt, K., Davis, D., et al. 2023, ApJ, 943, 177, doi: 10.3847/1538-4357/aca962

Miller, T. N., Doel, P., Gutierrez, G., et al. 2024, AJ, 168, 95, doi: 10.3847/1538-3881/ad45fe

Poppett, C., Tyas, L., Aguilar, J., et al. 2024, AJ, 168, 245, doi: 10.3847/1538-3881/ad76a4

Schlafly, E. F., Kirkby, D., Schlegel, D. J., et al. 2023, AJ, 166, 259, doi: 10.3847/1538-3881/ad0832

Silber, J. H., Fagrelius, P., Fanning, K., et al. 2023, AJ, 165, 9, doi: 10.3847/1538-3881/ac9ab1

Facility: KPNO:Mayall (DESI)

Facility: McDonald:HET (VIRUS)

APPENDIX

Table 5. FITS File HDU Information for `DESI_HETDEX_SPEC_v1.6.fits`

No.	Name	Type	Dimensions	Column Description
0	PRIMARY	PrimaryHDU	()	
1	INFO	BinTableHDU	2374R × 34C	Main Table. Columns described in Table 6
2	HETDEX_WAVE	ImageHDU	(1036)	Wavelength array, in air, for HETDEX data in Å
3	HETDEX_SPEC	ImageHDU	(1036, 2374)	Spectral flux array for HETDEX data in 1×10^{-17} erg Å^{-1} s^{-1} cm^{-2}
4	HETDEX_SPEC_ERR	ImageHDU	(1036, 2374)	Uncertainty in HETDEX spectral flux in 1×10^{-17} erg Å^{-1} s^{-1} cm^{-2}
5	DESI_WAVE	ImageHDU	(7781)	Wavelength array, in air, for DESI data in Å
6	DESI_WAVE_VACUUM	ImageHDU	(7781)	Wavelength array (in vacuum) for DESI data in Å
7	DESI_SPEC	ImageHDU	(7781, 2374)	Spectral flux array for DESI data in 1×10^{-17} erg Å^{-1} s^{-1} cm^{-2}
8	DESI_SPEC_ERR	ImageHDU	(7781, 2374)	Uncertainty in DESI spectral flux in 1×10^{-17} erg Å^{-1} s^{-1} cm^{-2}

Table 6. Column Description for INFO BinTable in the HETDEX-DESI Catalog

Column Name	Column Description	Data Type
TARGETID	DESI ID	int64
TARGET_RA	DESI fiber position right ascension (ICRS deg)	float64
TARGET_DEC	DESI fiber position declination (ICRS deg)	float64
TILEID	DESI Tile ID	int32
RA_HETDEX	HETDEX detectid right ascension in HDR3 (ICRS deg)	float32
DEC_HETDEX	HETDEX detectid declination in HDR3 (ICRS deg)	float32
SEP	separation between HETDEX detectid and the DESI fiber position	float32
DETECTID	HETDEX emission line detection ID	int64
VI_Z	Visual spectroscopic redshift	float32
VI_QUALITY	Quality in VI_z (4=multiple spectral features, 3=single emission line)	int32
SOURCE_TYPE	HETDEX source_type (lae, oii, agn, other)	str10
DEX_FLAG	HETDEX detections that are now flagged as compromised, ignored in analysis	int64
Z_BEST_0PT3	redshift determined from ELiXeR with threshold (P(Lyα)=0.3)	float32
Z_BEST_PT4	redshift determined from ELiXeR with threshold (P(Lyα)=0.4)	float32
Z_BEST_0PT5	redshift determined from ELiXeR with threshold (P(Lyα)=0.5)	float32
AV	dust correction in V band – not applied to fluxes	float32
GMAG	sdss-g magnitude measured in HETDEX spectrum	float32
SN_HETDEX	HETDEX signal-to-noise for line emission	float32
WAVE_HETDEX	central wavelength of line emission (Å, in air)	float32
WAVE_ERR_HETDEX	mcmc error in wave	float32
FLUX_HETDEX	observed line flux at WAVE_HETDEX (1×10^{-17} erg s^{-1} cm^{-2})	float32
FLUX_ERR_HETDEX	mcmc error in line FLUX_HETDEX (1×10^{-17} erg s^{-1} cm^{-2})	float32
SIGMA_HETDEX	sigma linewidth in gaussian line fit (Å)	float32
SIGMA_ERR_HETDEX	mcmc error in sigma linewidth (Å)	float32
CONT_HETDEX	local fitted observed continuum	float32
CONT_ERR_HETDEX	mcmc error in continuum	float32
WAVE_DESI	central wavelength of line emission (Å, in air)	float32
WAVE_ERR_DESI	mcmc error in wave	float32
FLUX_DESI	observed line flux at WAVE_DESI (1×10^{-17} erg s^{-1} cm^{-2})	float32
FLUX_ERR_DESI	mcmc error in line FLUX_DESI (1×10^{-17} erg s^{-1} cm^{-2})	float32
SIGMA_DESI	sigma linewidth in gaussian line fit (Å)	float32
SIGMA_ERR_DESI	mcmc error in sigma linewidth (Å)	float32
CONT_DESI	local fitted observed continuum	float32
CONT_ERR_DESI	mcmc error in continuum	float32