The Astrophysical Journal Supplement Series, 276:72 (12pp), 2025 February https://doi.org/10.3847/1538-4365/ada4a5


# The Hobby–Eberly Telescope Dark Energy Experiment Survey (HETDEX) Active Galactic Nuclei Catalog: The Fourth Data Release

Chenxu Liu[1], Karl Gebhardt[2], Erin Mentuch Cooper[2,3], Dustin Davis[2], Donald P. Schneider[4,5], Matt J. Jarvis[6,7], Daniel J. Farrow[8,9], Steven L. Finkelstein[2], and Óscar A. Chávez Ortiz[2,10]
(The HETDEX Collaboration)
[1] South-Western Institute for Astronomy Research, Yunnan University, Kunming 650500, People's Republic of China; cxliu@ynu.edu.cn
[2] Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Austin, TX 78712, USA
[3] McDonald Observatory, The University of Texas at Austin, 2515 Speedway, Austin, TX 78712, USA
[4] Department of Astronomy & 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] Astrophysics, Department of Physics, University of Oxford, Keble Road, Oxford, OX1 3RH, UK
[7] Department of Physics and Astronomy, University of the Western Cape, Robert Sobukwe Road, 7535 Bellville, Cape Town, South Africa
[8] Centre of Excellence for Data Science, Artificial Intelligence and Modelling, University of Hull, Cottingham Road, Hull, HU6 7RX, UK
[9] E.A. Milne Centre for Astrophysics, University of Hull, Cottingham Road, Hull, HU6 7RX, UK
*Received 2024 November 25; revised 2024 December 23; accepted 2024 December 24; published 2025 February 6*

## Abstract

We present the active galactic nuclei (AGN) catalog from the fourth data release (HDR4) of the Hobby–Eberly Telescope Dark Energy Experiment Survey (HETDEX). HETDEX is an untargeted spectroscopic survey. HDR4 contains 345,874 Integral Field Unit observations from 2017 January to 2023 August covering an effective area of 62.9 deg$^2$. With no imaging preselection, our spectroscopic confirmed AGN sample includes low-luminosity AGN, narrow-line AGN, and/or red AGN down to $g \sim 25$. This catalog has 15,940 AGN across the redshifts of $z = 0.1 \sim 4.6$, giving a raw AGN number density of 253.4 deg$^{-2}$. Among them, 10,499 (66%) have redshifts either confirmed by line pairs or matched to the Sloan Digital Sky Survey Quasar Catalog. For the remaining 5441 AGN, 2083 are single broad-line AGN candidates, while the remaining 3358 are single intermediate broad-line (full width at half-maximum, FWHM $\sim$1200 km s$^{-1}$) AGN candidates. A total of 4060 (39%) of the 10,499 redshift-confirmed AGN have emission-line regions $3\sigma$ more extended than the image quality, which could be strong outflows blowing into the outskirts of the host galaxies or ionized intergalactic medium.

*Unified Astronomy Thesaurus concepts:* Active galactic nuclei (16)

*Materials only available in the online version of record: tar.gz file*

## 1. Introduction

The Hobby–Eberly Telescope Dark Energy Experiment Survey[11] (HETDEX; K. Gebhardt et al. 2021) is a groundbreaking survey designed to map the large-scale structure of the Universe and constrain the nature of dark energy at cosmic noon through the observation of Ly$\alpha$ emitters (LAEs) and other extragalactic populations. However, beyond its cosmological objectives, HETDEX offers unique opportunities to probe active galactic nuclei (AGN) within its untargeted spectroscopic survey. Its integral field unit (IFU) configuration enables untargeted spectroscopic observations with no imaging preselections. This allows the identification of AGN, including populations that are typically underrepresented in traditional surveys, such as the low-luminosity AGN, narrow-line AGN, and/or red quasars at $z > 1$. These AGN, which are usually rejected by the imaging preselections in traditional surveys, can be efficiently revealed using IFU surveys (e.g., M. J. Jarvis et al. 2005; I. Gavignaud et al. 2006; C. Liu et al. 2022a, 2022c).

AGN are pivotal in understanding the coevolution of galaxies and their central supermassive black holes (SMBHs). They serve as laboratories for studying phenomena such as galaxy feedback, accretion physics, and the role of ionizing radiation in the intergalactic medium. HETDEX, with its large field of view and untargeted spectroscopic approach, is particularly suited to addressing challenges in AGN science, such as uncovering extended Ly$\alpha$ blobs around AGN. The spatially resolved spectra can help understand the origins of the extended emission-line regions around AGN and quantify the potential mass exchanges between the central SMBHs and their host galaxies.

The second HETDEX Data Release (HDR2) AGN catalog is discussed in C. Liu et al. (2022a, 2022b). In this paper, we release the 15,877 AGN identified from the fourth HETDEX Data Release (HDR4), which includes observations from 2017 January through 2023 August. Section 2 provides an overview of the HETDEX observations and data release history. Section 3 introduces the identification methods of AGN detections. Since HETDEX is an IFU survey, a single AGN can be detected with multiple fiber detections. Section 4 discusses how we deal with duplicate detections and generate the unique AGN catalog. Section 5 discusses the completeness of this catalog. Section 6 presents the statistics of the catalog (redshift and $g$-band magnitudes). Section 7 provides details of the data construction of the catalog file. Section 8 summarizes this catalog release.

[10] NASA FINESST Fellow.
[11] https://hetdex.org/

## 2. Observations, Data Reductions, and Data Release History

HETDEX is an untargeted spectroscopic survey aiming at mapping the $1.88 < z < 3.52$ Universe and measuring its expansion rate to 1% precision by providing the redshifts of one million LAEs over 540 deg$^2$. The untargeted spectroscopic observations with no imaging preselections are acquired by the Visible Integral field Replicable Unit Spectrograph (VIRUS; G. J. Hill et al. 2021). VIRUS has 78 $51'' \times 51''$ IFUs spanning the central 18′ field of view of the 22′ focal plane on the 10 m Hobby–Eberly Telescope (HET; L. W. Ramsey et al. 1998). The survey began from 2017 January with 16 IFUs mounted. IFUs were installed on the telescope gradually. All IFUs were completed with their installation in 2021 August. Each IFU bundle consists of 448 $1\rlap{.}''5$ diameter fibers. Upon completion, each HETDEX exposure collects 34,944 fiber spectra. The survey was recently completed with all observations in 2024 August. The spectral coverage is from 3500 to 5500 Å with the resolving power of $R \sim 800$.

The survey design, data reductions, and emission-line detection algorithms are detailed in K. Gebhardt et al. (2021). Each "shot" consists of three 6 minute dithered exposures to fill the space between fibers of individual IFU. This procedure yields a filling factor of 1/4.6 on the completion of all 78 IFUs. Sky subtraction is done with a single amplifier (112 fibers). A HETDEX $g$-band reconstructed image (integrated over 4400–5200 Å) is generated for each shot. Astrometry and flux of the reconstructed images are calibrated to a collection of stars in the Sloan Digital Sky Survey (SDSS) 15th Data Release (DR15) (D. G. York et al. 2000; K. N. Abazajian et al. 2009), Gaia DR2 (Gaia Collaboration et al. 2018), Panoramic Survey Telescope and Rapid Response System (PanSTARRS; K. C. Chambers et al. 2016; H. A. Flewelling et al. 2020), and United States Naval Observatory (S. Roeser et al. 2010) catalogs. All HETDEX flux-calibrated spectra in this paper are not corrected for foreground Galactic extinction.

Image quality and its error are calculated as the mean full width at half-maximum (FWHM) of the point spread functions (PSFs) of bright stars in the frame (typically ∼30 stars). The typical image quality is `FWHM_virus` $\sim 1\rlap{.}''8$. At the project's nominal depth, the PSF-weighted line flux sensitivity reaches a $5\sigma$ limit of $3.5 \times 10^{-17}$ erg cm$^{-2}$ s$^{-1}$ at $\lambda \sim 5000$ Å.

The first HETDEX Data Release (HDR1) contains observations through 2019 June and was commissioning so kept internal. HDR2 is through 2020 June. The HDR2 source catalog (200,000 sources with 51,863 $z \sim 2.5$ LAEs) is available in E. Mentuch Cooper et al. (2023). The HDR2 AGN catalog has 5322 AGN (C. Liu et al. 2022a), hereafter Liu22. The third HETDEX Data Release (HDR3) is through 2022 August. The first public data release (PDR1) based on HDR3 will contain ∼300 million fiber spectra and 500,000 sources (E. Mentuch Cooper et al. 2025, in preparation). Observations of HDR4 were completed in 2023 August with the full survey 72% completed. This paper describes the HETDEX HDR4 AGN catalog.

## 3. AGN Detection Identification

The most important sources that meet HETDEX's primary goal of measuring the cosmology at cosmic noon are the LAEs. The main source catalog is focused on the narrow emission-line detections down to signal-to-noise ratio (SNR) levels of $\sim 4.5\sigma$. The AGN whose lines are usually broad should be identified independently from the main source catalog.

Liu22 presents details of the AGN identification process in Section 4 and summarizes it in a flowchart (Figure 5). Here, we briefly review this process. We start from the direct detection products of the HETDEX pipeline (Section 3.1), narrow down the parent sample with quality control (Section 3.2), run the AGN search algorithms (Section 3.3), and collect the AGN detection candidates. These candidates are then visually inspected (Section 3.4).

### 3.1. HETDEX Source Detection

The HETDEX project focuses on the high completeness and the low false positive rate of detecting the LAEs. Continuum bright sources produce a high false positive rate of line detections because windows of continuums sometimes can be incorrectly fit with significant broad lines during the three-dimensional element search of line emission signals. However, broad emission lines are rare in the LAE population. Therefore, it is more efficient to separate the continuum bright sources (Section 3.1.1) and the line emission detections (Section 3.1.2) for different scientific purposes.

#### 3.1.1. The Continuum Catalog

We measure the counts in a 200 Å window in the blue (3700–3900 Å) and in the red (5100–5300 Å). A threshold is set to separate the continuum detections and the line emission detections. If either region of a fiber spectrum contains more than 50 counts per 2 Å pixel on average ($g \sim 22.5$), the position of the fiber is marked as a continuum source and searched with a $15 \times 15$ grid raster of $0\rlap{.}''1$ spatial bins to locate the position that has the lowest $\chi^2$ from the PSF fit. A PSF-weighted extraction spectrum is then generated for the continuum source, and no further wavelength-resolved element search is performed for this continuum source.

#### 3.1.2. The Line Catalog

All noncontinuum fiber spectra are examined for line emission signals in the three-dimensional space of the IFU data cubes. The initial search uses a $3 \times 3$ grid of $0\rlap{.}''5$ in the spatial direction and 8 Å element search in the wavelength direction. The large bins in the initial search resolution, for each IFU, is 448 fibers per dither × 3 dithers × 500 wavelength elements per fiber × 9 spatial elements per fiber ∼6 million element search. As of HDR4, there are 345,874 IFU observations in total, corresponding to $2 \times 10^{12}$ element analyses. We performed numerical simulations and determined that this initial grid search can effectively save computation time while guaranteeing the completeness to a 1% level compared to finer grids. Each resolution element is fit with a single Gaussian profile using the instrumental line width of 2 Å. Any Gaussian line candidates (SNR $> 4$, $\chi^2 < 3$) are examined with a ($5 \times 5$) grid of $0\rlap{.}''15$ raster to identify the line center. Each line detection is assigned parameters from the best single Gaussian fit (wavelength, line width, etc.) and a PSF-weighted extraction spectrum.

### 3.2. Parent Sample

AGN usually have strong emission lines, and a significant population of the AGN are continuum bright quasars.

Therefore, we start from a combination of the 600,000 continuum detections (Section 3.1.1) and the 28 million line detections (Section 3.1.2). Before running through the AGN detection algorithm, we made some initial constraints on the data. First, we require that a nominal throughput, assuming a 360 s exposure time, must be greater than 0.08. We then refine the catalog by removing known issues, such as bad amplifiers, hot pixels, and meteors. Finally, we remove some artifacts, mostly bleeding cosmic-ray hits, using $\chi^2_{\rm PSF} < 4$ from the 2D fit. After removal of suspicious detections with the three criteria, the 29 million raw catalog is reduced to 12 million detections (500,000 continuum detections and 11.5 million line detections).

### 3.3. AGN Identification Methods

We executed our two AGN search methods: the line-pair (LP) method (Section 3.3.1) and the single broad-line (sBL) search (Section 3.3.2), as detailed in Liu22.

#### 3.3.1. The LP Method

The LP search basically identifies emission LPs characteristic of AGN, especially when the highly ionized lines, such as C IV $\lambda$1549, are present (SNR > 5). This approach is applied for both the continuum catalog (Section 3.1.1) and the line catalog (Section 3.1.2). Considering the wavelength coverage of HETDEX (3500–5500 Å), the strongest AGN emission lines visible are O VI $\lambda$1034, Ly$\alpha$ $\lambda$1215, N V $\lambda$1241, C IV $\lambda$1549, C III] $\lambda$1909, Mg II] $\lambda$2799 from the redshift of $z = 4.32$ to $z = 0.25$ (see Figure 1 in Liu22).

For each line detection, we assume the wavelength suggested by the HETDEX pipeline's single Gaussian fit to be one of the six strongest lines and check if there are signals matched to at least another line emission. The other emission line is not limited to be one of the six lines, e.g., He II $\lambda$1640 can be used to confirm the C IV $\lambda$1549 emission. If the full wavelength range of the line detection spectrum is detected with the strongest and the second strongest lines at $>8\sigma$ and $>5\sigma$, this object is selected as an AGN candidate with a redshift indicated by the LP(s). Higher thresholds would result in missing real AGN, and lower thresholds would bring in too many false positives making it beyond our visual inspection ability (Section 3.4). We find that with the $>8\sigma$ and $>5\sigma$ thresholds no known AGN missed and the false positive rate is controlled at $\lesssim$50% in our visual inspection experiments on a few small test samples. We will eventually push to lower thresholds to achieve higher completeness by allowing higher false positive rates in our final data release.

For a continuum detection, there is no line wavelength suggested by the HETDEX pipeline in addition to a PSF-weighted extraction spectrum. We fit a power-law continuum, find the highest peak in the continuum subtracted spectrum, attempt to fit the peak as one of the six strong AGN lines, and see if any other emission line(s) match with at least another peak(s). If the spectrum of a continuum detection is detected with the two strongest emission lines at levels of $>8\sigma$ and $>5\sigma$, it is then identified as an AGN candidate. All AGN identified by the LP method are labeled with `sflag = 2` (column 8 described in Section 7 and provided in Table 2).

#### 3.3.2. The sBL Method

Because the wavelength coverage of HETDEX is only 2000 Å and the limited number of lines available in this wavelength range, many of the detections only have one line detected in their spectra. For example, at $0.96 < z < 1.26$, the only line among the six strong lines visible is C III] $\lambda$1909. Another possibility is that the second strongest line in our wavelength range is not significant enough to be identified ($<5\sigma$) so the spectrum also only shows a single line (see Figure 3 in Liu22 for some examples). We fit multi-Gaussian profiles to these single lines. Our code allows two emission-line components, one capturing the broad feature and the other for the narrow emission line, and at most four absorption components. If a line is detected with FWHM > 1200 km s$^{-1}$ at $>5\sigma$, the source is a broad-line AGN candidate. Again, the choice of $5\sigma$ is a result of the balance between the completeness and the false positive rate. When testing on several simulated test samples, we find that a $4.5\sigma$ threshold can marginally increase the completeness from 89% to 91%, while the false positive rate is dramatically increased from 50% to 90% compared to the $5\sigma$ threshold. $5\sigma$ is believed to be efficient in recovering real broad lines while reducing the amount of work in visual inspections.

Some of the sBL-identified objects are only AGN candidates, especially the ones with intermediate broad lines FWHM $\sim$1200 km s$^{-1}$; they can be AGN with intermediate massive black holes, or type 1.5–2 Seyferts whose narrow-line components are as strong or stronger than their broad-line components. They can also be star-bursting galaxies or galaxies undergoing other activities that can increase the gas velocity dispersion. These single intermediate broad lines (sIBLs) are labeled with `agn_flag = 0` (column 7 described in Section 7 and provided in Table 2) so that they can be easily removed from statistics.

For any sBL-identified AGN candidate (`sflag = 1`, column 8 described in Section 7 and provided in Table 2), an estimated redshift (flagged by `zflag = 0`, column 6 described in Section 7 and provided in Table 2) is provided based on the wavelength of the line detection and the line emission guess. The line is mainly identified as a specific emission according to their equivalent widths (EWs). Liu22 show in Figures 14, 18, and Table 2 that for the HETDEX AGN, different line emissions have different EW distributions. Within the HETDEX wavelength range, the most likely broad emission lines are (Ly$\alpha$ $\lambda$1215 +N V $\lambda$1241), C IV $\lambda$1549, C III] $\lambda$1909, Mg II] $\lambda$2799, and O VI $\lambda$1034 in a decreasing EW order. The (Ly$\alpha$ $\lambda$1215 +N V $\lambda$1241) emission is usually easily distinguished from other lines because of the high EW and the line profiles of two blended strong emissions. O VI $\lambda$1034 usually differs from the other emissions in their noisy line shape features due to the heavy Ly$\alpha$ $\lambda$1215 forest absorption lines. It is usually difficult to distinguish single broad C IV $\lambda$1549, C III] $\lambda$1909, and Mg II] $\lambda$2799 simply based on their EWs. We check if the stronger lines expected in the wavelength range are present or not to narrow down the possible line solutions. For example, assuming a sBL at 5000 Å as C IV $\lambda$1549, the Ly$\alpha$ $\lambda$1215 emission should be expected at 3921 Å. If no line emission is detected at 3921 Å, the sBL at 5000 Å is then rejected to be C IV $\lambda$1549. We checked the reliability of our redshift estimates by crossmatching the sBL-identified AGN with SDSS DR16Q (B. W. Lyke et al. 2020). About 90% of our redshift estimates are correct.

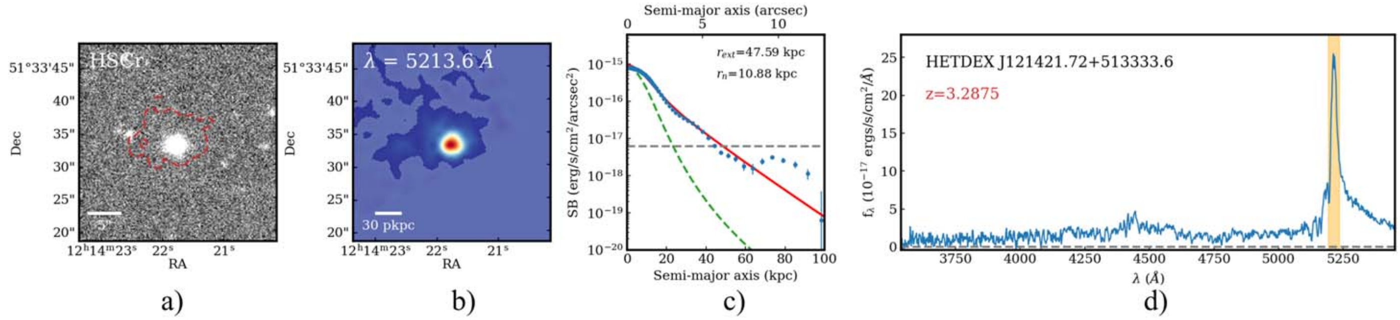


**Figure 1.** An example of an AGN (`agnid = 407`) detected with extended diffuse ionized gas. (a) The HSC $r$-band $30'' \times 30''$ image with the red Ly$\alpha$ $\lambda$1215 narrowband $2\sigma$ density contour. (b) The HETDEX Ly$\alpha$ $\lambda$1215 narrowband image. (c) The radial profile of the Ly$\alpha$ $\lambda$1215 flux density map. Blue data points are surface brightness variations. The red curve is a (PSF + exponential) fit to the blue points, giving a $r_{\rm ext} = 47.6$ kpc. The green dashed curve gives the PSF model using the image quality of this observation. The gray dashed horizontal line marks the standard deviation in the continuum subtracted background in the line flux map. (d) The spectrum extracted at the labeled coordinate in the observed frame. The yellow shaded region highlights the $\pm 21$ Å wavelength range making the Ly$\alpha$ $\lambda$1215 density map. We refer the readers to E. Mentuch Cooper et al. (2025, in preparation) for more detailed analyses of the extended diffuse ionized gas around LAEs.

### *3.3.3. Crossmatch with SDSS DR16Q*

Besides the two major AGN identification methods, the LP method described in Section 3.3.1 and the sBL method detailed in Section 3.3.2, we further crossmatch our parent sample (Section 3.2) with SDSS QSOs using the SDSS's 3″ fiber size as the matching radius. The HDR2 AGN catalog (Liu22) used SDSS DR14Q (I. Pâris et al. 2018). In this paper, the HDR4 AGN catalog is matched to SDSS DR16Q (B. W. Lyke et al. 2020).

A significant benefit of the crossmatches is that they confirm or correct the redshifts for AGN identified by the sBL method. If a guessed redshift is either confirmed or corrected by an SDSS quasar, the redshift flag is reset from `zflag = 0` to `zflag = 1` (column 6 in Table 2).

Besides redshift confirmations, the crossmatch with SDSS quasars can also identify some missed AGN (see Figure 4 in Liu22 for one example). For quasars with spatially extended emission-line regions, whose PSF-like continuum sources fall just outside our IFU edges, the edge fibers sometimes can catch their weak diffuse emission-line signals. For these cases, their spatially resolved information is important, while the lines are too weak and/or noisy to be identified by the LP method or the sBL method. Crossmatching with SDSS can help recover these AGN. Additionally, at higher redshifts ($z > 4.32$), even O VI $\lambda$1034 is beyond our wavelength coverage. The crossmatch can pick up the continuum spectra bluer than O VI $\lambda$1034 for these high-$z$ AGN (see the first spectrum in Figure 7 for one example). For broad-line AGN at low redshifts $z \gtrsim 0.1$, their broad H$\beta$ and H$\alpha$ lines lie beyond our red wavelength limit, the crossmatch can also recover these AGN (see the last spectrum in Figure 7 for one example).

### *3.4. Visual Identification*

Once the AGN detection candidates are identified from the parent sample as detailed in Section 3.3, visual inspections are made of each candidate. This step can remove artifacts, such as bleeding cosmic rays misidentified by the sBL search. Continuum features of stars or galaxies wrongly identified as broad lines by the sBL search can also be removed. There are some LPs whose weaker lines are not significant enough to be detected by the LP search ($\sim 4\sigma$) and their redshift estimates are given by the remaining sBL detections. During the visual inspections, their $\sim 3\sigma$–$4\sigma$ emission lines can provide additional information, leading to better redshift guesses. The visual inspections reduce the 130,000 AGN detection candidates into 63,348 visually confirmed AGN detections.

## 4. Duplicate Removal

As an IFU survey, a spatially extended source can illuminate multiple fibers and produce multiple detections in the HETDEX survey. This frequently occurs for the low-redshift big [O II] galaxies. At $z \sim 2$–3, about 5% of the HETDEX LAE sample are "Ly$\alpha$ $\lambda$1215 blobs" which are significantly more extended than the single PSF model (E. Mentuch Cooper et al. 2025, in preparation). There are also many duplicates in the 63,827 AGN detections identified by the processes in Section 3.3 and Section 3.4. We apply friend-of-friend grouping (FoF[12]) to remove duplicates from the sample. The challenge is the choice of the linking length to make the unique AGN sample: A large one would group real AGN overdensities into a single AGN, while an insufficient linking length would leave the sample with many duplicates. During our visual inspections, many of the HETDEX AGN are found to be surrounded by diffuse ionized gas whose associated emission lines are more significantly extended than the typical FWHM $= 1\farcs8$ PSF (e.g., C. Liu et al. 2022c). Figure 1 displays an extended AGN in our catalog. The PSF plus exponential fit ($r_{\rm ext} = 47.59$ kpc) can better reproduce the Ly$\alpha$ $\lambda$1215 density map than the PSF model alone ($r_{\rm n} = 10.88$ kpc).

There is no universal linking length suitable for the full AGN catalog. We carried out a set of FoF experiments with different linking lengths ($\delta r = 3\farcs0, 3\farcs5, 4\farcs0, 4\farcs5, 5\farcs0, 5\farcs5, 6\farcs0$) to determine the optimal choice. Liu22 demonstrates in Figure 16 that the broad-line AGN fraction is $\sim$90% of the AGN in the luminosity and redshift range we probe. Considering this high fraction of broad lines and the complex line profiles in the Ly$\alpha$ $\lambda$1215 regime, the 3D FoF is performed with $\delta z = 0.1$ in the redshift space. We also carried out 2D FoF experiments for comparison reasons. Figure 2 displays that the number of unique AGN decreases significantly as $\delta r$ increases at $\delta r \lesssim 5''$, reaching roughly a constant value at $n$(unique AGN) $\sim$ 16,000 at $\delta r \gtrsim 5''$. We randomly examine the line flux maps of some newly added distinct AGN in the ($\delta r$, $\delta z$) = ($4\farcs5$, 0.1) set in addition to the ($\delta r$, $\delta z$) = ($5\farcs0$, 0.1) set. The newly added

[12] https://github.com/HETDEX/hetdex_api/blob/master/hetdex_tools/fof_kdtree.py

**Table 1**
Summary of the Survey Areas of the HETDEX HDR4 AGN Catalog

| Field Name | Field Center (J2000, deg) | Area ($deg^2$) | $N_{AGN}$ | $n_{AGN}$ ($deg^{-2}$) | $N_{AGN,secure\ z}$ | $n_{AGN,secure\ z}$ ($deg^{-2}$) |
|---|---|---|---|---|---|---|
| DEX-spring | (201.7620, 52.2367) | 36.313 | 9733 | 268.0 | 6288 | 173.2 |
| DEX-fall | (23.0755, 0.3056) | 20.655 | 4819 | 233.3 | 3411 | 165.1 |
| NEP | (270.2904, 65.9958) | 4.238 | 981 | 231.5 | 548 | 129.3 |
| COSMOS | (150.1025, 2.2396) | 1.289 | 294 | 228.1 | 179 | 138.9 |
| SSA22 | (334.5131, 0.3211) | 0.362 | 83 | 229.3 | 59 | 163.0 |
| GOODS-N | (189.2030, 62.2461) | 0.052 | 30 | 546.9 | 14 | 269.2 |
| Total | ⋯ | 62.909 | 15,940 | 253.4 | 10,499 | 166.9 |

**Note.** DEX-spring and DEX-fall fields are HETDEX survey fields. COSMOS and GOODS-N fields are HETDEX science verification fields. The NEP field (Ó. A. Chávez Ortiz et al. 2023) and the SSA22 field are taken for our collaborators with their own scientific purposes.

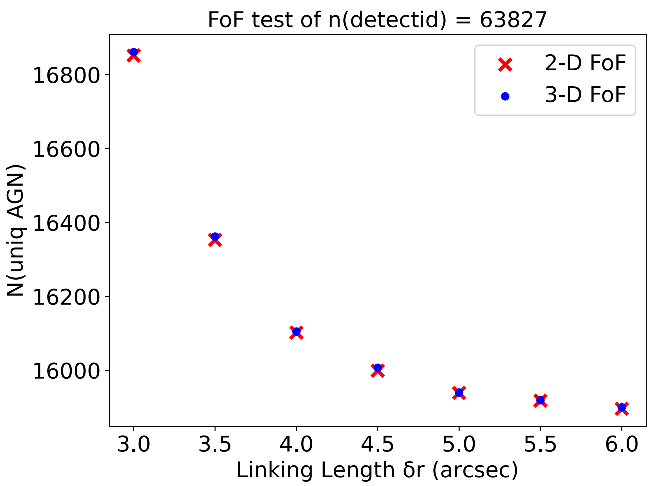


**Figure 2.** The FoF experiments with linking lengths of $\delta r = 3''.0, 3''.5, 4''.0, 4''.5, 5''.0, 5''.5, 6''.0$ for the 63,827 AGN detections. The 2D FoF tests are the red crosses. The 3D FoF tests ($\delta z = 0.1$) are the blue data points.

distinct AGN in the ($\delta r$, $\delta z$) = ($4''.5$, 0.1) set are just the extended emission of existing AGN in the ($\delta r$, $\delta z$) = ($5''.0$, 0.1) set. This suggests that $\delta r = 4''.5$ is not sufficient to remove duplicates. To avoid using a too large $\delta r$ and group real AGN pairs or groups into single ones, we decided to adopt $\delta r = 5''$. The 2D FoF ($\delta r = 5''.0$) yields the unique AGN number as $n$ (unique AGN) = 15,939, while the 3D FoF ($\delta r$, $\delta z$) = ($5''.0$, 0.1) gives $n$(unique AGN) = 15,940. The one additional AGN in the 3D FoF experiment is a real superposition of two separate AGN within $5''$. Our final choice of FoF parameter set to group the 63,827 AGN detections into a 15,940 unique AGN catalog is ($\delta r$, $\delta z$) = ($5''.0$, 0.1). Among the 15,940 unique AGN sample, 10,499 of them have confirmed redshifts.

The coordinate of each unique AGN is the emission-line flux-weighted center of all member detections in the FoF grouping. The member detection that is closest to the FoF center of a given AGN is then the "best" detection and marked as `detectid_best` (column 12 in Table 2). Figure 3 compares the separations between `detectid_best` (left) and their emission-line flux-weighted FoF centers (columns 2 and 3 in Table 2) to those of all AGN detections to their FoF centers (right). The median separations of the 15,940 `detectid_best` to their FoF center is $\delta D = 0''.12$, while the median separations of all AGN detections to their FoF center is $\delta D = 0''.63$. There are many AGN with member detections larger than $3''$ from their FoF centers. `agnid=407` (Figure 1) even has a $>5\sigma$ member detection $8''.2$ away from its FoF center. With the HETDEX resolving power (2 Å in the wavelength space, and $0''.15$ in the fiber space, Section 3.1), `agnid=407` shows no evidence of pair(s) or group(s) in the Ly$\alpha$ $\lambda$1215 flux density map and the 1D spectrum.

There are 4060 (39%) out of the 10,499 redshift-confirmed HETDEX HDR4 AGN with member detections ($>5\sigma$) more extended than their relevant PSF models `(FWHM_virus + 3 FWHM_virus_err)`/2. Detailed comparisons between a single PSF model fit and two components fit of the line flux maps can be found in E. Mentuch Cooper et al. (2025, in preparation). This high fraction of AGN with extended emission-line regions could be AGN experiencing strong outflows blowing to the circumgalactic medium or even the intergalactic medium. Alternatively, the extended emission-line regions could also be gaseous haloes lit up by the central AGN. Detailed two-dimensional analyses on the line shifts and the line widths would help understand the physics behind this (C. Liu et al. 2025, in preparation).

## 5. Completeness and Contamination

The HETDEX HDR4 AGN catalog consists of 15,940 AGN after grouping duplicates into unique objects (Section 4). Among them, 10,499 AGN (`zflag=1`, column 6 in Table 2) have redshifts either confirmed by the LP method or the crossmatch with SDSS DR16Q. The remaining 5441 AGN (`zflag=0`) are sBL-identified AGN with guessed redshifts only (see Section 3.3.2 for details). 3358 out of the 5441 AGN are only AGN candidates with sIBLs (FWHM ∼1200 km $s^{-1}$) whose emission lines might be broadened by activities other than AGN, such as starbursts. They are flagged with `agn_flag=0` (column 7 in Table 2). The (5441 − 3358 = 2083) sBL AGN combined with the 10,499 `zflag=1` redshift-confirmed AGN yields the 12,582 "secure AGN" catalog (`agn_flag=1`).

Table 1 summarizes the survey area, the number of AGN, and the raw AGN density in each field in HDR4. The high AGN number density in the GOODS-N fields can either originate from real cosmic variance or from small number statistics. Even only counting the redshift secured AGN, the HETDEX AGN density is 166.9 $deg^{-2}$, which is significantly higher than that of the SDSS quasars (56 $deg^{-2}$; I. Pâris et al. 2018). This highlights the effectiveness and the high spatial completeness in detecting optically faint AGN with the untargeted observational strategy of HETDEX.

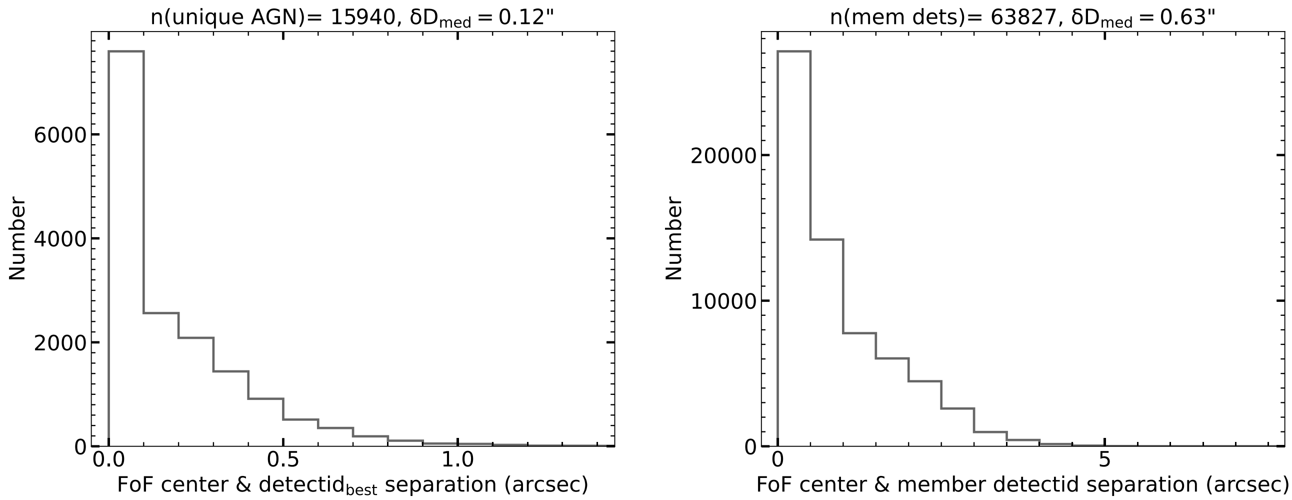


**Figure 3.** Left: distributions of the separations between the "best" detections (`detectid_best`) to their emission-line flux-weighted FoF centers. Right: distributions of the separations between all AGN detections to their emission-line flux-weighted FoF centers.

The HETDEX AGN identification is mostly regulated by the SNR of emission lines, as introduced in Sections 3.1 and 3.3. The detection completeness as a function of line SNR is quantitatively introduced in Figures 6 and 7 of Liu22. Additionally, in order to calculate the HETDEX AGN luminosity function, we carefully modeled the sample completeness in C. Liu et al. (2022b) against all relevant parameters, including the line wavelengths, line fluxes, line widths, the relevant locations on IFUs, and redshifts for the HETDEX pipeline and the two AGN identification methods (Sections 3.3.1 and 3.3.2). However, these parameters are not straightforward in comparing our completeness with other surveys. The broadband continuum level (HETDEX $g$-band magnitude) is then estimated for each detection by multiplying the 1D HETDEX spectrum with the SDSS $g$ filter's throughput curve using the software Emission Line eXplorer[13] (`ELiXer`; D. Davis et al. 2023).

Figure 4 presents the completeness estimation and the potential contamination of the HETDEX HDR4 AGN sample within each HETDEX $g$-band magnitude bin. We crossmatched all 500 million HETDEX HDR4 fibers with the 638,022 SDSS DR16Q quasar sample (quasars flagged by redshift warnings removed) to evaluate the completeness of both the HETDEX pipeline (Section 3.1) and our AGN identification methods (Section 3.3). There are 5872 SDSS quasars covered by HETDEX good fibers (hot fibers, fibers on the edge of the IFUs, and shots with low throughput excluded) within 1″. The weakness of using this crossmatch sample to evaluate the completeness is that the sample size is limited at $g \gtrsim 22.5$.

The recovery fraction of the SDSS DR16Q quasars brighter than $g = 22.5$ is 93% for the HETDEX pipeline (black squares) and 80% for our AGN sample identified either by the LP method (Section 3.3.1) or by the sBL method (Section 3.3.2) (blue circles). Considering all 5872 matched SDSS DR16Q quasars, these two fractions are 92% and 78%, respectively. The LP method (Section 3.3.1) requires the strongest emission line to be $>8\sigma$. We would eventually push our current SNR thresholds to lower values in our final AGN data release so that the blue circles can match the black squares. Although the LP method suffers from the HETDEX's short wavelength coverage (3500–5500 Å), it is the main contributor (65%) of the AGN sample. The sBL method is also very important (13%) in recovering AGN missed by the LP method. However, the sBL method also brings in AGN with unconfirmed redshifts (`zflag = 0`, column 6 in Table 2) and even unconfirmed AGN (`agn_flag = 0`, column 7 in Table 2) into our sample as detailed in Section 3.3.2.

For AGN brighter than $g = 22.5$, 9% of our AGN sample have unconfirmed redshifts, and 2% are unconfirmed AGN. When including fainter bins at $g > 22.5$, the two fractions are increased to 28% and 15%. We note here that these unconfirmed redshifts are not necessarily wrong. They are simply estimated based on sBL wavelengths. Similarly, the unconfirmed AGN are not necessarily non-AGN. They can be low-luminosity Seyferts. With our current data by hand, we do not have enough evidence to make further confirmations.

## 6. Statistics

Figure 5 shows the redshift distributions of the HETDEX HDR4 AGN sample. The major six strong lines (O VI $\lambda$1034, Ly$\alpha$ $\lambda$1215, N V $\lambda$1241, C IV $\lambda$1549, C III] $\lambda$1909, Mg II] $\lambda$2799) we used in our AGN search algorithm and the HETDEX wavelength range of 3500–5500 Å suggest a redshift range of $0.25 < z < 4.32$. The SDSS survey has a much broader wavelength range. By crossmatching with SDSS DR16Q, a few AGN can be identified at higher ($z > 4.32$) and lower ($z < 0.25$) redshifts, while the majority of the HETDEX HDR4 AGN population is at $z \sim 2.1$. The blue histogram is clearly more populous at $z = 2 \sim 3.5$ than the secure AGN subsample (`agn_flag=1`), suggesting that many of the sIBL AGN candidates could be non-AGN LAEs. SDSS quasars are more abundant than HETDEX AGN at $z \sim 1$; this result is because it is the redshift desert of single C III] $\lambda$1909 in the HETDEX wavelength range.

Figure 6 presents the $g$-band magnitude distributions of HETDEX HDR4 AGN. The typical flux density limit of HETDEX's 6 minute exposures is $\sim 5 \times 10^{-19}$ erg s$^{-1}$ cm$^{-2}$ Å$^{-1}$ ($g \lesssim 25$). HETDEX is better at detecting AGN fainter than $g \sim 22.5$ than SDSS. The second peak in the blue histogram at $g \sim 24$ is suspicious. This peak is dominated by the `agn_flag=0` sIBL AGN candidates. It would be interesting to observe these objects in deep narrowband images for potential

[13] https://github.com/HETDEX/elixer

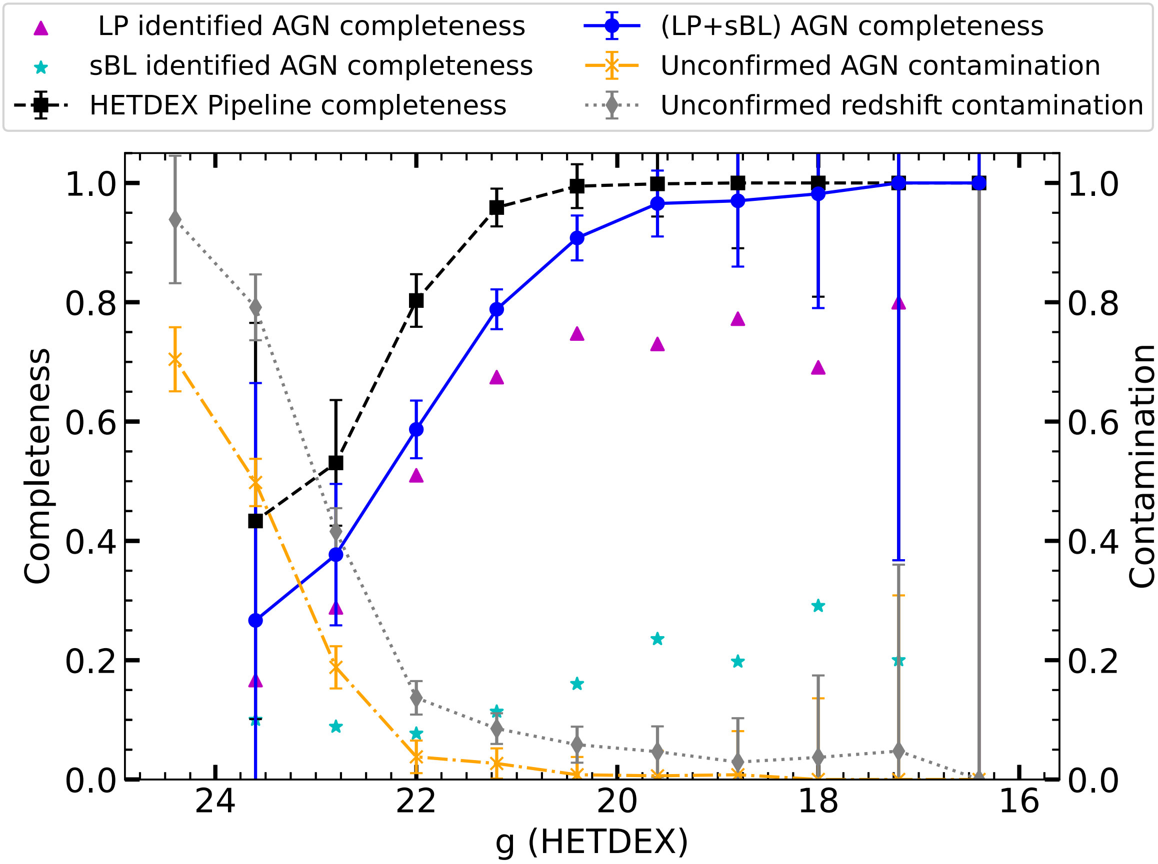


**Figure 4.** The completeness and contamination of the HETDEX HDR4 AGN sample in each HETDEX $g$-band magnitude bin. Completeness is estimated from the crossmatched sample between the SDSS DR16Q sample and all HETDEX HDR4 fibers. The black squares connected with a dashed line show the completeness of the parent sample of this work (Sections 3.1 and 3.2). The blue circles connected by a solid line show the completeness of the AGN identified either by the LP method (Section 3.3.1) or by the sBL method (Section 3.3.2). The magenta triangles and the cyan stars show the completeness of the LP method (Section 3.3.1) and the sBL method (Section 3.3.2), respectively. The gray diamonds connected by a dotted line present the contamination rate from the unconfirmed redshifts (`zflag = 0`, column 6 in Table 2). The orange crosses connected by a dashed–dotted line show the contamination rate from the unconfirmed AGN (`agn_flag = 0`, column 7 in Table 2).

emission lines at longer wavelengths in the future to determine whether they are non-AGN LAEs or low-luminosity Seyferts.

## 7. The Catalog

The HETDEX HDR4 AGN catalog is available in a .tar.gz package.[14] The FITS file (D. C. Wells et al. 1981) is constructed similarly to the HDR2 release in Liu22 with six extensions. Table 2 provides details on the main catalog in Extension 1. A simple Jupyter notebook is also included that describes how to use this FITS file in Python. The FITS extensions are as follows.

*Extension 1.* Table extension (15,940 rows $\times$ 56 columns): Basic information, one row for one unique AGN, arranged in a descending order of redshifts. Column information is given in the header and summarized in Appendix A. We explain each column in more detail as follows.

*Column 1.* Sequential ID number assigned for each AGN organized in a descending redshift order from `agnid` $= 0$ to 15,939.

*Columns 2 and 3.* The R.A. and decl. of the emission-line flux-weighted FoF center of all member detections of a given AGN (J2000). If more than one line is present within the

[14] http://web.corral.tacc.utexas.edu/hetdex/HETDEX/catalogs/agn_catalog_v2.0/

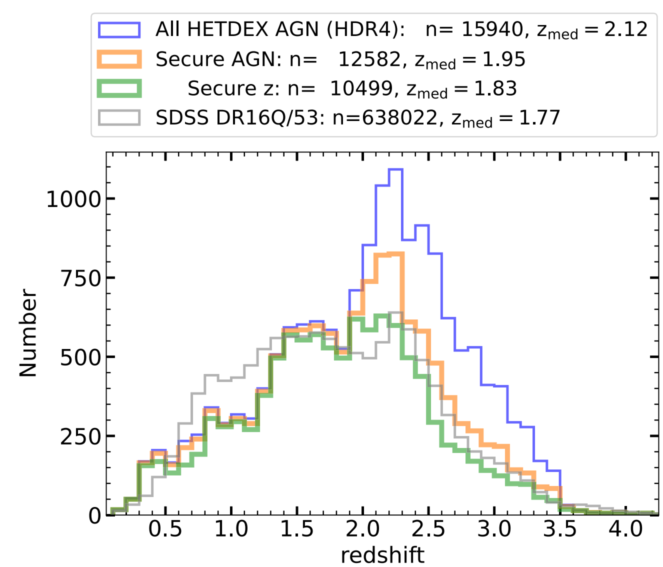


**Figure 5.** The redshift distributions of the HETDEX HDR4 AGN catalog. The blue histogram is for the full 15,940 AGN set. The orange histogram excludes the sIBL AGN candidates (`agn_flag = 0`, column 7 in Table 2). The green histogram further excludes the AGN with estimated redshifts only (`zflag = 0`, column 6 in Table 2). The gray histogram is the scaled redshift distributions of the SDSS DR16Q sample to generally match the number of HETDEX AGN at $z\sim1.5$.

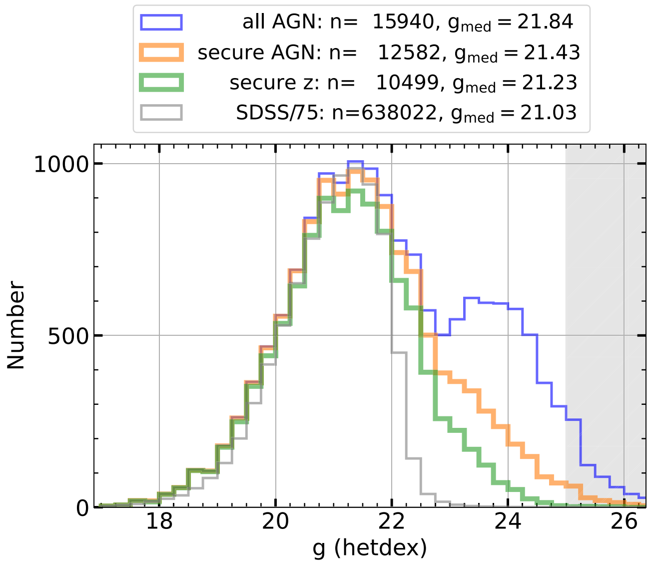


**Figure 6.** Similar format as Figure 5 but for the HETDEX $g$-band magnitudes. While $g$-band magnitudes are plotted as computed, the $5\sigma$ depth of HETDEX is roughly 25, marked by the shaded area.

HETDEX wavelength range, only the strongest line is used as the weight. The FoF grouping method to remove duplicates from the 63,827 AGN detection sample is detailed in Section 4.

*Columns 4 and 5.* The best-fit redshift and its error of the 1D spectrum of the "best" detection, i.e., `detectid_best`. Member detections can have slightly different redshifts due to gas kinematics.

*Column 6.* Redshift flag. Redshifts confirmed either by the LP method or matched SDSS quasars have `zflag=1`, while estimated redshifts from sBLs are flagged with `zflag=0`.

*Column 7.* AGN flag. The sIBL AGN candidates have `agn_flag=0`; the remaining confirmed AGN have `agn_flag=1`.

*Column 8.* Source flag giving the information on a given AGN identified by which specific selection method is described in Section 3.3. `sflag=` 1, 2, and 3 indicate the AGN is identified by the sBL method (Section 3.3.2), LP method (Section 3.3.1), and the crossmatch with SDSS DR16Q method (Section 3.3.3). The few `sflag=0` AGN are identified by old versions of the AGN selection code.

*Column 9.* Field name in the HETDEX survey as indicated by Table 1.

*Column 10.* The number of shots with fiber coverage at (R.A., decl.). A given coordinate can be observed by more than one shot. For example, some of the COSMOS fields were observed more than 10 times.

*Column 11.* Number of member detections related with a given AGN from the 63,827 AGN detection sample visually confirmed by Section 3.4.

*Column 12.* The member detection of a given AGN closest to its emission-line flux-weighted FoF center as indicated by the coordinates recorded in columns 2 and 3.

*Columns 13 and 14.* The R.A. and decl. of `detectid_best`. The position can be slightly off from the FoF center as the spatial detection search resolution is $0\overset{''}{.}15$ (Section 3.1).

*Column 15.* Offset between the coordinate of `detectid_best` (`ra_best` and `dec_best` in columns 13 and 14) and the FoF center (`ra` and `dec` in columns 2 and 3). Distributions of the offset (`roff`) are presented in the left panel of Figure 3.

*Columns 16 and 17.* The HETDEX $g$-band magnitude and its 68% confidence interval calculated by multiplying the 1D spectrum of `detectid_best` with the SDSS $g$ filter's throughput curve using `ELiXer`.

*Column 18.* The identification of the fiber closest to `detectid_best`. It is a standard string with the format of `yyyymmddsss_multi_bbb_ccc_ddd_aa_fff`. The first eight letters give the date `detectid_best` is observed. The next three digits are the shot number on the date catches `detectid_best`. `bbb`, `ccc`, and `ddd` are all three-digit numbers giving the spectrograph number, the IFU number, and the IFU slot number of `detectid_best`. There are four amplifiers for each IFU: LL, LU, RL, and RU. `aa` is a two-letter code among the four indicating the amplifier for `detectid_best`. Each amplifier has 112 fibers. `fff` is the three-digit number of the specific fiber on the amplifier contributes the most flux to the 1D PSF-weighted spectrum of `detectid_best`. The full string of `fiberid` can identify the raw data of a given AGN.

*Column 19.* The ID of the shot for `detectid_best`. This is an 11-digit number consistent with the first 11 digits of `fiberid`.

*Column 20.* The specific amplifier for `detectid_best`. This is a 20-letter string consistent with the 13–32 letters of `fiberid`.

*Column 21.* The throughput of `shotid`. This parameter is described in K. Gebhardt et al. (2021). Basically, it is defined as the nominal throughput treating each exposure as if it were 360 s long through a 50 m$^2$ clear aperture. Throughput varies at different wavelengths. The typical throughput at 4540 Å of our AGN sample is 0.15. In this paper, we require the throughput to be greater than 0.08 for an observation to be included in the catalog.

*Columns 22 and 23.* Image quality and its error. Image quality is calculated as the mean FWHM of the PSFs of bright stars in the full frame (typically ~30 stars). Details can be found in K. Gebhardt et al. (2021) Section 6.15. The typical image quality is `FWHM_virus` ~`1.8″`.

*Column 24.* The aperture correction of `detectid_best`. The typical value of our AGN catalog is 0.92. A value of 1 means no correction.

*Columns 25 and 26.* The best-fit observed line wavelength and its error from `ELiXer`'s single Gaussian fit. For line detections (Section 3.1.2), `ELiXer` fits to the wavelength identified by the HETDEX line detection algorithm. For continuum detections (Section 3.1.1), `ELiXer` finds the most possible line to fit.

*Columns 27–46.* The best-fit continuum subtracted rest-frame line fluxes, SNRs, FWHMs, and EWs of Ly$\alpha$ $\lambda$1215, C IV $\lambda$1549, C III] $\lambda$1909, Mg II] $\lambda$2799, and O VI $\lambda$1034. Parameters of lines falling outside the wavelength range are set to a default value of −99.0.

*Columns 47–52.* The best single Gaussian fit parameters from the HETDEX's pipeline (K. Gebhardt et al. 2021) in the observed frame. For detections in the continuum catalog, these parameters are set to 0.

*Column 53.* The $\chi^2$ of fit to the 2D spectrum. A high value (`chi2fib_pipe > 4`) usually indicates issues such as hot pixels, bleeding cosmic rays, etc.

*Columns 54–56.* Match results with the SDSS DR16Q sample. The match radius is 3″. AGN with no matches within

3″ are set with default values of `rsep_dr16q=999.0″`, `z_dr16q=-999.0`, and `sdssid_dr16q`="?."

*Extension 2*. Image extension (15,940 rows × 1036 columns): 1D spectra for the $\text{detectid}_{\text{best}}$ of each AGN (no extinction correction applied). Rows are arranged in the same order as Extension 1. The wavelength array starts from 3470 Å and contains 1036 elements with a step size of 2.0 Å. The flux densities are in units of $10^{-17}\,\text{erg}\,\text{cm}^{-2}\,\text{s}^{-1}\,\text{Å}^{-1}$. The wavelength solution and the flux density units are recorded in the extension header.

*Extension 3*. Image extension (15,940 rows ×1036 columns): The error array for Extension 2, with the same wavelength solution and flux units.

*Extension 4*. Table extension (17,552 rows × 4 columns): Repeat information for each AGN. Some AGN were observed multiple times. In this table, each unique observation is listed with their `shotid`. Column 1 is the `agnid`, column 2 gives the number of repeat observations of `agnid`, and column 3 gives the `shotid` of the observation. An AGN observed multiple times will have multiple entries in this table with the same `agnid` and `nshots`, but different `shotid`. There are 14,306 AGN observed only once with no repeats, 1125 AGN observed twice, and 240 AGN observed more than twice. Note that 15,671 AGN can be successfully extracted with spectra at their FoF grouping center (R.A., decl.) in Extension 1. 269 AGN are lack of their extracted spectra because their flux-weighted FoF center falling on bad amplifiers (low throughput or other issues).

*Extension 5*. Image extension (17,552 rows × 1036 columns): 1D PSF-weighted extracted spectra at the FoF center, repeat observations included, no extinction correction applied. Rows are arranged in the same order as Extension 4. The wavelength solution and flux units are the same as Extension 2, and can be found in the header.

*Extension 6*. Image extension (17,552 rows × 1036 columns): Error array for Extension 5.

## 8. Summary

The HETDEX HDR4 AGN catalog has 15,940 AGN. A total of 10,499 have redshifts either confirmed by LPs or matched SDSS quasars, while the rest are sBL AGN candidates provided with redshift estimates. The catalog contains observations from January 2017 through 2023 August. The effective survey area is 62.9 deg$^2$, giving a raw AGN density of 253.4 deg$^{-2}$. The redshifts range from 0.1 to 4.6, and peak at $z \sim 2.1$. The $g$-band magnitudes of the faint AGN can reach the HETDEX continuum detection limit at $g \sim 25$ mag. Approximately 39% of the 10,499 redshift-confirmed AGN sample has extended emission-line regions with $>5\sigma$ signals detected at distances significantly larger than the image quality $>$ (`FWHM_virus` + 3 `FWHM_virus_err`)/2.

## Acknowledgments

C.X.L. acknowledges the support from the "Science & Technology Champion Project" (202005AB160002) and the "Top Team Project" (202305AT350002), both funded by the "Yunnan Revitalization" (202401AT070489).

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.

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, http://www.tacc.utexas.edu) 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.

## Appendix A
## Catalog Format and Column Information

The HETDEX HDR4 AGN catalog is made available online in FITS format, and is described in Table 2. Other extensions to the FITS file are described in Section 7.

**Table 2**
Columns of Extension 1 in the FITS File

| Column | Name | dtype | Unit | Description |
|---|---|---|---|---|
| 1 | agnid | int32 | ⋯ | Sequential number for each unique AGN |
| 2 | ra | float32 | degrees | R.A. of the AGN (center of flux-weighted FoF, J2000) |
| 3 | dec | float32 | degrees | Decl. of the AGN (center of flux-weighted FoF, J2000) |
| 4 | $z$ | float32 | ⋯ | Redshift |
| 5 | z_err | float32 | ⋯ | Redshift error |
| 6 | zflag | int32 | ⋯ | zflag=0/1: 1 confirmed $z$; 0 estimated $z$ |
| 7 | agn_flag | int32 | ⋯ | agn_flag=0/1: 1 confirmed AGN; 0 AGN candidate |
| 8 | sflag | int32 | ⋯ | Method flag: sBL=1,2em=2,sdss=3,else=0 |
| 9 | field | bytes10 | ⋯ | Field name in the HETDEX survey |
| 10 | nshot | int32 | ⋯ | Number of the observed shots of (R.A., decl.) |
| 11 | nmem | int32 | ⋯ | Number of member detectids |
| 12 | detectid_best | int64 | ⋯ | Detectid closest to the FoF center |
| 13 | ra_best | float32 | degrees | R.A. of detectid_best |
| 14 | dec_best | float32 | degrees | Decl. of detectid_best |

**Table 2**
(Continued)

| Column | Name | dtype | Unit | Description |
|---|---|---|---|---|
| 15 | roff | float32 | arcseconds | Offset between ra_best,dec_best with FoF center |
| 16 | mag_g_wide | float32 | mag | HETDEX $g$ measured at (ra_best, dec_best) |
| 17 | mag_g_wide_err | float32 | mag | Error of mag_g_wide |
| 18 | fiberid | bytes38 | ⋯ | ID of the fiber that is closest to detectid_best |
| 19 | shotid | int64 | ⋯ | ID of the shot for detectid_best |
| 20 | multiframe | bytes20 | ⋯ | IFU for detectid_best |
| 21 | throughput | float32 | ⋯ | Throughput of shotid |
| 22 | fwhm_virus | float32 | arcseconds | Image quality (FWHM) of shotid |
| 23 | fwhm_virus_err | float32 | arcseconds | Error of fwhm_virus |
| 24 | apcor | float32 | ⋯ | Aperture correction of detectid_best |
| 25 | wobs_elixer | float32 | angstrom | Observed line wavelength fit by ELIXER |
| 26 | wobs_err_elixer | float32 | angstrom | Error of wobs_elixer |
| 27 | flux_LyA | float32 | 1e-17 erg $s^{-1}$ $cm^{-2}$ | Rest-frame flux of the LyA emission |
| 28 | snr_LyA | float32 | ⋯ | SNR of LyA |
| 29 | fwhm_LyA | float32 | km $s^{-1}$ | Rest-frame FWHM of the LyA emission |
| 30 | ew_LyA | float32 | angstrom | Rest-frame EW of the LyA emission |
| 31 | flux_CIV | float32 | 1e-17 erg $s^{-1}$ $cm^{-2}$ | Rest-frame flux of the C IV emission |
| 32 | snr_CIV | float32 | ⋯ | SNR of C IV |
| 33 | fwhm_CIV | float32 | km $s^{-1}$ | Rest-frame FWHM of the C IV emission |
| 34 | ew_CIV | float32 | angstrom | Rest-frame EW of the C IV emission |
| 35 | flux_CIII | float32 | 1e-17 erg $s^{-1}$ $cm^{-2}$ | Rest-frame flux of the C III emission |
| 36 | snr_CIII | float32 | ⋯ | SNR of C III |
| 37 | fwhm_CIII | float32 | km $s^{-1}$ | Rest-frame FWHM of the C III emission |
| 38 | ew_CIII | float32 | angstrom | Rest-frame EW of the C III emission |
| 39 | flux_MgII | float32 | 1e-17 erg $s^{-1}$ $cm^{-2}$ | Rest-frame flux of the Mg II emission |
| 40 | snr_MgII | float32 | ⋯ | SNR of Mg II |
| 41 | fwhm_MgII | float32 | km $s^{-1}$ | Rest-frame FWHM of the Mg II emission |
| 42 | ew_MgII | float32 | angstrom | Rest-frame EW of the Mg II emission |
| 43 | flux_OVI | float32 | 1e-17 erg $s^{-1}$ $cm^{-2}$ | Rest-frame flux of the O VI emission |
| 44 | snr_OVI | float32 | ⋯ | SNR of O VI |
| 45 | fwhm_OVI | float32 | km $s^{-1}$ | Rest-frame FWHM of the O VI emission |
| 46 | ew_OVI | float32 | angstrom | Rest-frame EW of the O VI emission |
| 47 | wave_pipe | float32 | angstrom | Observed line wavelength from the HETDEX pipeline |
| 48 | flux_pipe | float32 | 1e-17 erg $s^{-1}$ $cm^{-2}$ | Observed line flux from the HETDEX pipeline |
| 49 | continuum_pipe | float32 | 1e-17 erg $s^{-1}$ $cm^{-2}$ $Å^{-1}$ | Observed continuum from the HETDEX pipeline |
| 50 | linewidth_pipe | float32 | angstrom | Observed sigma from the pipeline's single Gaussian fit |
| 51 | sn_pipe | float32 | ⋯ | SNR from the pipeline's single Gaussian fit |
| 52 | chi2_pipe | float32 | ⋯ | chi2 from the pipeline's single Gaussian fit |
| 53 | chi2fib_pipe | float32 | ⋯ | chi2 of fit to 2d spec for each fiber |
| 54 | rsep_dr16q | float64 | arcseconds | Separation between the matched SDSS DR16Q AGN |
| 55 | z_dr16q | float32 | ⋯ | Redshift of the matched SDSS DR16Q AGN |
| 56 | sdssid_dr16q | bytes18 | ⋯ | SDSS name of the matched SDSS DR16Q AGN |

## Appendix B
## Example HETDEX AGN Spectra

Figure 7 shows example HETDEX AGN spectra covering the redshift range of the catalog.

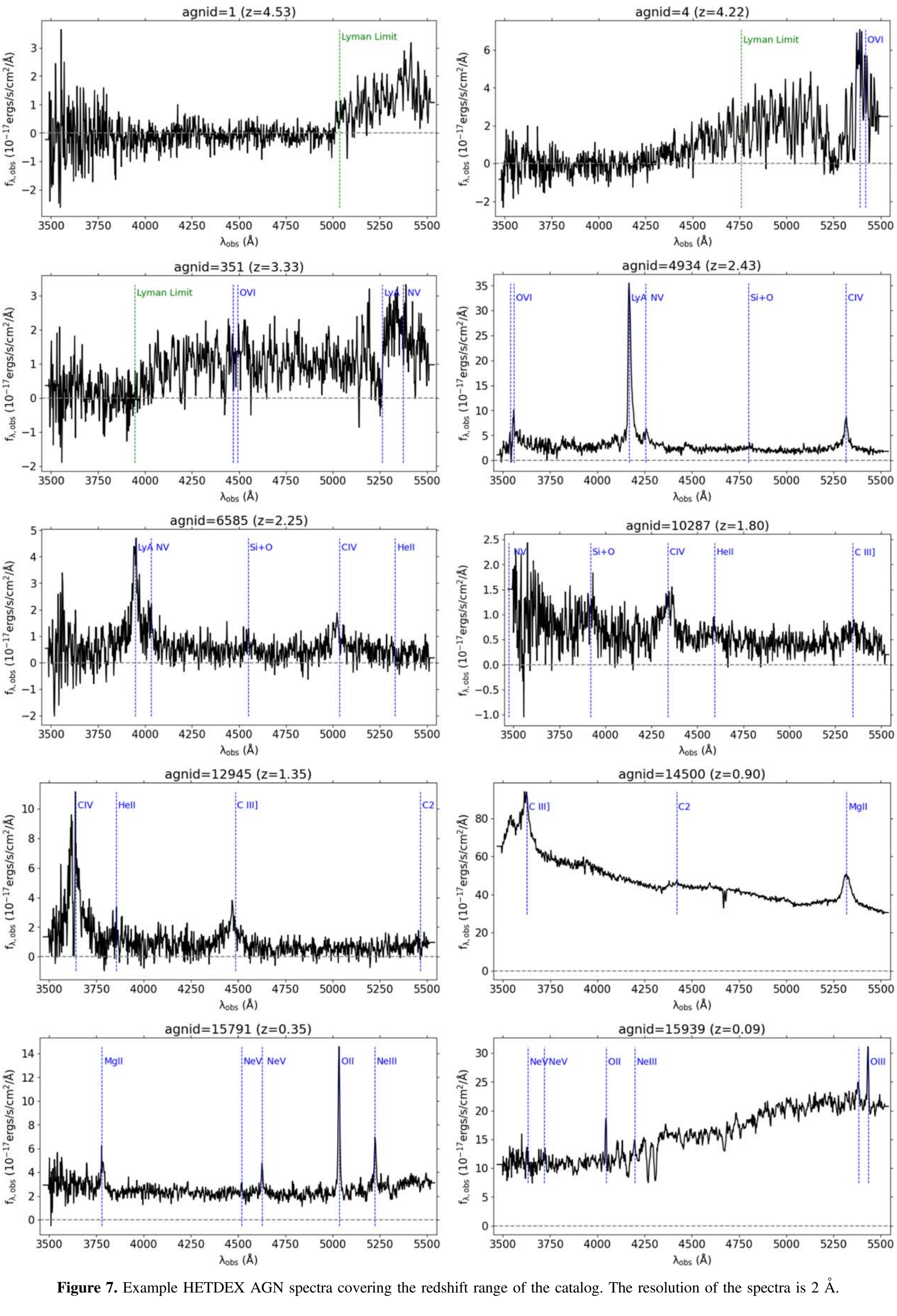


**Figure 7.** Example HETDEX AGN spectra covering the redshift range of the catalog. The resolution of the spectra is 2 Å.

## ORCID iDs

Chenxu Liu https://orcid.org/0000-0001-5561-2010
Karl Gebhardt https://orcid.org/0000-0002-8433-8185
Erin Mentuch Cooper https://orcid.org/0000-0002-2307-0146
Dustin Davis https://orcid.org/0000-0002-8925-9769
Donald P. Schneider https://orcid.org/0000-0001-7240-7449
Matt J. Jarvis https://orcid.org/0000-0001-7039-9078
Daniel J. Farrow https://orcid.org/0000-0003-2575-0652
Steven L. Finkelstein https://orcid.org/0000-0001-8519-1130
Óscar A. Chávez Ortiz https://orcid.org/0000-0003-2332-5505

## References

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
Chávez Ortiz, Ó. A., Finkelstein, S. L., Davis, D., et al. 2023, ApJ, 952, 110
Davis, D., Gebhardt, K., Cooper, E. M., et al. 2023, ApJ, 946, 86
Flewelling, H. A., Magnier, E. A., Chambers, K. C., et al. 2020, ApJS, 251, 7
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Gavignaud, I., Bongiorno, A., Paltani, S., et al. 2006, A&A, 457, 79
Gebhardt, K., Mentuch Cooper, E., Ciardullo, R., et al. 2021, ApJ, 923, 217
Hill, G. J., Lee, H., MacQueen, P. J., et al. 2021, AJ, 162, 298
Jarvis, M. J., van Breukelen, C., & Wilman, R. J. 2005, MNRAS, 358, L11
Liu, C., Gebhardt, K., Cooper, E. M., et al. 2022a, ApJS, 261, 24
Liu, C., Gebhardt, K., Cooper, E. M., et al. 2022b, ApJ, 935, 132
Liu, C., Gebhardt, K., Kollatschny, W., et al. 2022c, ApJ, 940, 40
Lyke, B. W., Higley, A. N., McLane, J. N., et al. 2020, ApJS, 250, 8
Mentuch Cooper, E., Gebhardt, K., Davis, D., et al. 2023, ApJ, 943, 177
Pâris, I., Petitjean, P., Aubourg, É., et al. 2018, A&A, 613, A51
Ramsey, L. W., Adams, M. T., Barnes, T. G., et al. 1998, Proc. SPIE, 3352, 34
Roeser, S., Demleitner, M., & Schilbach, E. 2010, AJ, 139, 2440
Wells, D. C., Greisen, E. W., & Harten, R. H. 1981, A&AS, 44, 363
York, D. G., Adelman, J., Anderson, J. E. J., et al. 2000, AJ, 120, 1579