The Astrophysical Journal Supplement Series, 192:5 (34pp), 2011 January doi:10.1088/0067-0049/192/1/5

THE HETDEX PILOT SURVEY. I. SURVEY DESIGN, PERFORMANCE, AND CATALOG OF EMISSION-LINE GALAXIES*

Joshua J. Adams[1], Guillermo A. Blanc[1], Gary J. Hill[2,3], Karl Gebhardt[1,3], Niv Drory[4], Lei Hao[1,10], Ralf Bender[4,5], Joyce Byun[1,11], Robin Ciardullo[6], Mark E. Cornell[2], Steven L. Finkelstein[7], Alex Fry[1,12], Eric Gawiser[8], Caryl Gronwall[6], Ulrich Hopp[4,5], Donghui Jeong[1,3,13], Andreas Kelz[9], Ralf Kelzenberg[4], Eiichiro Komatsu[1,3], Phillip J. MacQueen[2], Jeremy Murphy[1], P. Samuel Odoms[2], Martin Roth[9], Donald P. Schneider[6], Joseph R. Tufts[1,2,14], and Christopher P. Wilkinson[2]

[1] Department of Astronomy, University of Texas at Austin, 1 University Station C1400, Austin, TX 78712, USA
[2] McDonald Observatory, University of Texas at Austin, 1 University Station C1402, Austin, TX 78712, USA
[3] Texas Cosmology Center, University of Texas at Austin, 1 University Station C1400, Austin, TX 78712, USA
[4] Max-Planck Institüt fur extraterrestrische Physik, Giessenbachstraße, 85741 Garching bei München, Germany
[5] Universittäs-Sternwarte der Ludwig-Maximilians-Universität, Scheinerstraße 1, 81679 München, Germany
[6] Department of Astronomy and Astrophysics, The Pennsylvania State University, 525 Davey Lab, University Park, PA 16802, USA
[7] George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, Department of Physics and Astronomy, Texas A&M University, 4242 TAMU, College Station, TX 77843, USA
[8] Department of Physics and Astronomy, Rutgers, The State University of New Jersey, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA
[9] Astrophysikalisches Institut Potsdam, An der Sternwarte 16, D-14482 Potsdam, Germany
Received 2010 October 12; accepted 2010 November 10; published 2010 December 20

ABSTRACT

We present a catalog of emission-line galaxies selected solely by their emission-line fluxes using a wide-field integral field spectrograph. This work is partially motivated as a pilot survey for the upcoming Hobby–Eberly Telescope Dark Energy Experiment. We describe the observations, reductions, detections, redshift classifications, line fluxes, and counterpart information for 397 emission-line galaxies detected over 169 □′ with a 3500–5800 Å bandpass under 5 Å full-width-half-maximum (FWHM) spectral resolution. The survey's best sensitivity for unresolved objects under photometric conditions is between 4 and 20 × 10^{-17} erg s^{-1} cm^{-2} depending on the wavelength, and Lyα luminosities between 3 × 10^{42} and 6 × 10^{42} erg s^{-1} are detectable. This survey method complements narrowband and color-selection techniques in the search of high-redshift galaxies with its different selection properties and large volume probed. The four survey fields within the COSMOS, GOODS-N, MUNICS, and XMM-LSS areas are rich with existing, complementary data. We find 105 galaxies via their high-redshift Lyα emission at $1.9 < z < 3.8$, and the majority of the remainder objects are low-redshift [O II]3727 emitters at $z < 0.56$. The classification between low- and high-redshift objects depends on rest-frame equivalent width (EW), as well as other indicators, where available. Based on matches to X-ray catalogs, the active galactic nuclei fraction among the Lyα emitters is 6%. We also analyze the survey's completeness and contamination properties through simulations. We find five high-z, highly significant, resolved objects with FWHM sizes >44 □″ which appear to be extended Lyα nebulae. We also find three high-z objects with rest-frame Lyα EW above the level believed to be achievable with normal star formation, $EW_0 > 240$ Å. Future papers will investigate the physical properties of this sample.

Key words: cosmology: observations – galaxies: evolution – galaxies: formation – galaxies: high-redshift

Online-only material: color figures, machine-readable tables

1. INTRODUCTION

The Hobby–Eberly Telescope (HET) Dark Energy Experiment (HETDEX; Hill et al. 2004, 2008a) will survey 60 □° spread throughout 420 □° to discover 0.8 million new Lyα emitting galaxies (LAEs) over $1.9 < z < 3.5$ and use them to map the expansion history of the universe. A further ∼1 million low-z galaxies will have their redshifts determined, primarily in the [O II]3727 transition, over $0 < z < 0.47$. The primary HETDEX science goal is to measure the dark energy equation of state at high redshift by using the three-dimensional power spectrum of LAE positions and redshifts (Jeong & Komatsu 2006, 2009; Koehler et al. 2007; Shoji et al. 2009). An important secondary goal of HETDEX is to investigate the physical properties of star-forming galaxies, through Lyα and [O II] emission, using vastly greater statistics and volumes than currently available. The survey will use an array of 150 integral field spectrographs (IFSs) on the upgraded 10 m HET (Ramsey et al. 1998; Savage et al. 2010) called the Visible Integral field Replicable Unit Spectrograph (VIRUS; Hill et al. 2010).

The HETDEX Pilot Survey (HPS) is the pathfinder to the full HETDEX survey. This pilot survey provides a direct test of equipment, data reduction, target properties, observing procedures, and ancillary data requirements to HETDEX by using one IFS, named the VIRUS prototype (VIRUS-P; Hill et al. 2008b), on the 2.7 m Harlan J. Smith telescope at the McDonald

* This paper includes data taken at The McDonald Observatory of The University of Texas at Austin.
[10] Current address: Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China
[11] Current address: Department of Astronomy, Cornell University, 610 Space Sciences Building, Ithaca, NY 14853, USA
[12] Current address: Department of Astronomy, University of Washington, Box 351580, U.W., Seattle, WA 98195, USA
[13] Current address: California Institute of Technology, Mail Code 350-17, Cahill Center 326, Pasadena, CA, 91125, USA
[14] Current address: Las Cumbres Observatory Global Telescope Network, Inc., 6740 Cortona Drive Suite 102, Santa Barbara, CA 93117, USA

Observatory over 111 nights. To do this, the pilot survey uses the novel technique of blindly targeted, wide-field contiguous spectroscopy to find emission-line objects over a broad redshift range. While large numbers of narrowband-selected LAEs have been assembled by previous surveys (e.g., Hu & McMahon 1996; Cowie & Hu 1998; Rhoads et al. 2000; Steidel et al. 2000; Ouchi et al. 2003; Hu et al. 2004; Hayashino et al. 2004; Santos et al. 2004; Palunas et al. 2004; Venemans et al. 2005; Gawiser et al. 2006; Gronwall et al. 2007; Nilsson et al. 2007; Ouchi et al. 2008; Nilsson et al. 2009; Guaita et al. 2010; Tilvi et al. 2010), these surveys are heterogeneous in nature, with different depths and equivalent width (EW) limits. The HPS is designed to produce a homogeneous sample of LAEs over an extremely large volume, 1.03×10^6 Mpc^3 h_{70}^{-3}, that is nearly an order of magnitude larger than the largest existing blind spectroscopic survey, 2.5×10^5 Mpc^3 h_{70}^{-3} (Cassata et al. 2010), and vastly larger than other blind surveys (Pirzkal et al. 2004; van Breukelen et al. 2005; Xu et al. 2007; Sawicki et al. 2008; Martin et al. 2008). This allows us to evaluate potential redshift evolution of LAE properties and to make comparisons to color-selected high-redshift galaxy populations (e.g., Steidel et al. 1996, 1999; Daddi et al. 2004; Kornei et al. 2010). The HPS also enables us to find a large sample of lower redshift galaxies selected through, primarily, their [O II]3727, Hβ, and [O III] emission and study their properties over a lower redshift ranges (up to $z = 0.56$, 0.19, 0.17, and 0.16 for [O II], Hβ, [O III]4959, and [O III]5007, respectively).

The paper is organized as follows. In Section 2.1, we describe the instrumental capabilities of VIRUS-P, the type and quality of data taken, the necessary calibrations, and the imaging compiled to aid source classification. We detail the data reduction steps, with special care given toward tracking systematic errors in Section 3. In Section 4.1, we describe the methods used to recover objects to the survey's statistical limits and analyze the effect of noise contamination and the emission-line flux measurements. In Section 5, we present our classification methods, relying primarily on imaging counterpart likelihoods and EW measurements. The contamination of the high-redshift LAE sample by active galactic nuclei (AGNs) is presented as well as example classifications. The final emission-line catalog and its summary properties are given in Section 6. Finally, in Section 7, we review the analysis and describe its place in future projects.

In this work, we adopt a standard ΛCDM cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$. All magnitudes are quoted in the AB system (Oke & Gunn 1983). All wavelengths are corrected to vacuum conditions in the heliocentric frame with an assumed wavelength-independent index of refraction for air at the observatory's altitude of $n = 1.00022$.

2. OBSERVATIONS

2.1. Instrumental Configuration

The Visible Integral-field Replicable Unit Spectrograph Prototype (VIRUS-P) was designed for this pilot survey and is described in Hill et al. (2008b), and references therein. The instrument is a fiber-based IFS fed at $f/3.65$ on the McDonald Observatory's 2.7 m Harlan J. Smith telescope. A small focal reducer sits just prior to the Integral Field Unit (IFU) input in the lightpath of the telescope's $f/8.8$ focus. Originally, VIRUS-P operation used a focal reducer labeled FR1, but all data taken after 2008 September used a second focal reducer labeled FR2, which has significantly improved efficiency below 4000 Å compared to FR1 (see Section 2.4). Auto-guiding and sky transparency measurements were performed with an off-the-shelf Apogee Alta camera installed into a field position ∼9′ north of the IFS field of view (FOV). The guider has a square 20.25□′ FOV and uses a $B+V$ filter with a mean wavelength of 5000 Å at a platescale of 0.″53 pixel^{-1}.

Two different IFUs have been used over the course of this pilot survey. Fiber bundle IFU-1, used prior to 2008 March, spans 1.′70 × 1.′77 with 244 functional and 3 broken 200 μm core diameter (4.″235 on-sky) fibers. IFU-2 spans 1.′61 × 1.′65 with 246 functional and 0 broken fibers of the same core size. There is no significant difference in throughput between the bundles. Both IFUs are of the densepak type (Barden et al. 1998) with a filling factor near 1/3, requiring at least three dithered positions to fully sample the FOV. This survey utilizes a six position dithering pattern as illustrated in Figure 1. The nearly × 2 oversampling delivered by this dithering pattern provides improved spatial registration between detected spectral objects and imaging-based continuum counterparts. The wavelength range on VIRUS-P is adjustable from 3400 to 6800 Å, and a set of volume phase holographic gratings delivering various spectral resolutions are available. For this survey the instrument was set to cover 3500–5800 Å at resolutions that range from 4.5 to 5.5 Å full-width-half-maximum (FWHM) over the whole data set through an 831 lines mm^{-1} grating that delivers a dispersion of 1.1 Å pixel^{-1} in the unbinned charge-coupled device (CCD) mode. The spectral resolution over that range weakly and gradually varies with wavelength and between different fibers due to CCD surface shape deviations from planarity, camera design limits, and the residual camera alignment errors. The data are recorded on a 2k × 2k CCD with 15 μm pixels in a custom built, LN_2 cooled, vacuum-sealed camera (Tufts et al. 2008) with electronics that deliver between 3.6–4.2 e^- read noise, making the sky background the dominant source of noise at all wavelengths in our 20 minute exposures. The data have been taken with 2 × 1 binning along the dispersion direction to minimize read noise and still maintain a Nyquist sampling of the instrumental line profile.

Several instrumental properties determine the survey's calibration needs. The instrument's scattered light properties have been discussed in Adams et al. (2008). A weak in-focus ghost of atmospheric OH lines redder than the targeted wavelength range was found to exist at discrete wavelengths. These lines are easily distinguished by their deviations from calibrated wavelength solutions and fiber trace positions. The strength of the scattered light varied over time as alignments changed and baffling was implemented, but the ghost's strength was at maximum 3× the resolution element noise, and more characteristically below the noise in any one fiber. The scattered light affected one resolution element per fiber. Extra masking installed around the grating solved this issue for all data taken after 2008 September. All emission-line sources discussed in this paper from observations prior to the installation of the grating mask have been visually inspected to not lie in the affected regions.

The lab testing and characterization of the VIRUS-P fibers, with particular attention to transmission and focal ratio degradation, have been investigated in Murphy et al. (2008). A high stability in each fiber's throughput over a night, at minimum, is crucial toward the survey's goals. IFU mounting practices have been established from these tests to yield fiber stability sufficient for our purposes. To facilitate mounting on the HET

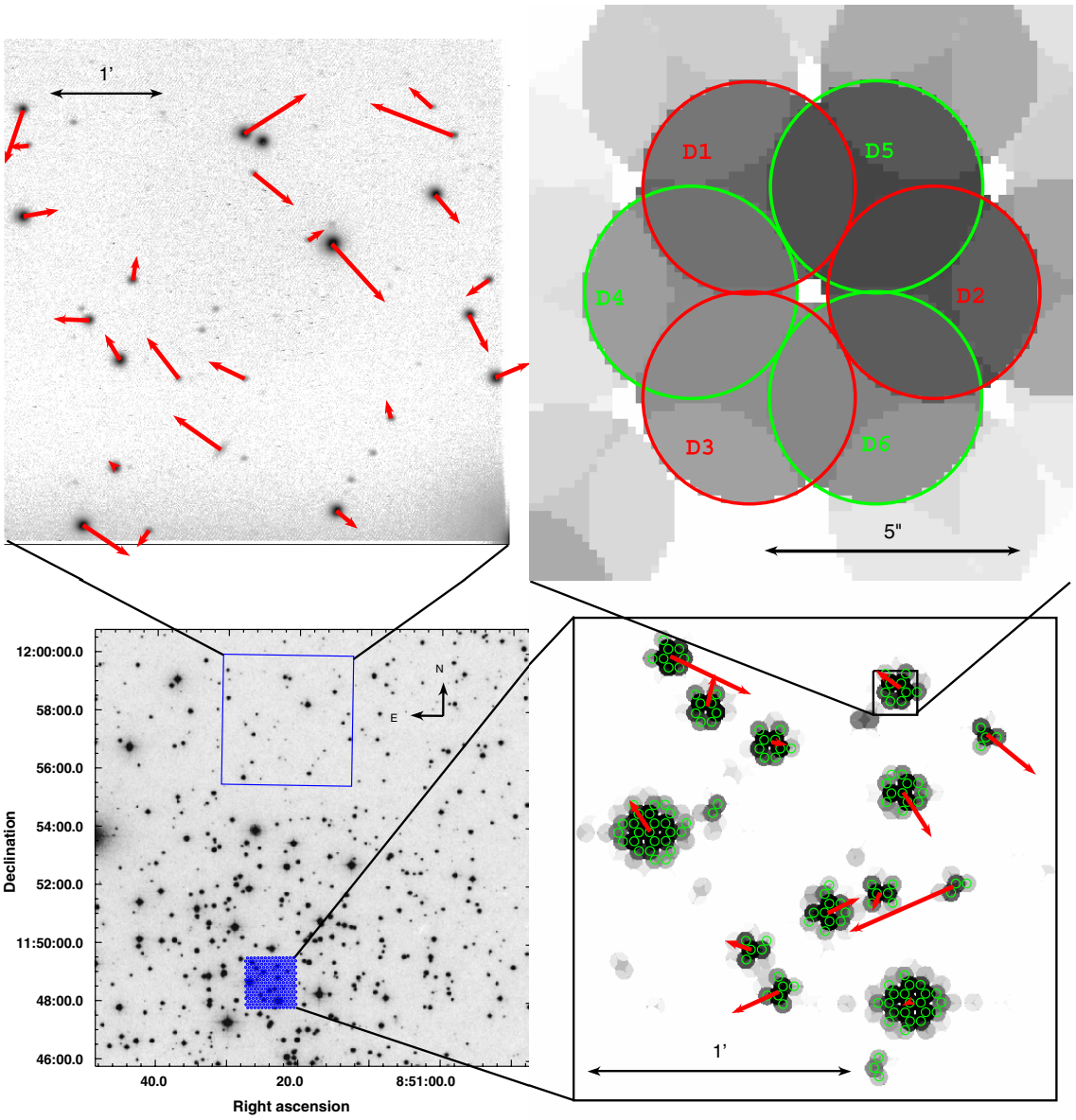

Figure 1. Layout of a VIRUS-P observation and the quality of a guider-based astrometric solution. Bottom left: the footprints of the spectroscopic science FOV and the northernly offset guider FOV overlaid on a Digital Sky Survey (DSS) image of the open cluster M67. This type of field is used to calibrate the astrometry of the guider and the fibers as discussed in Section 2.3. Top left: an expanded view of the VIRUS-P guider field with residuals from the astrometric model. The residuals are shown as red vectors scaled by 60×. The rms is 0.″42. Bottom right: an expanded view of the science FOV. The continuum map is generated from the IFS data summed over 4100 Å $< \lambda <$ 5700 Å. Fibers that have significant flux and border other significant fibers are highlighted with green circles and bunched as point source detections for the astrometric fit. The residuals are shown as red vectors scaled by 60×. The rms is 0.″21. The residuals in the IFU are less than the residuals in the guider as both fields have a similar number of degrees of freedom, but the guider has more data points. Top right: the expanded view of one fiber moving through the six dither positions. The pattern, marked with D1–D6, gives very nearly an oversampling of two. The small offsets necessary to complete the dither patterns are controlled by sending offsets to the guider.

(A color version of this figure is available in the online journal.)

as well as the Smith telescope, the IFU was made longer than otherwise necessary. Since the IFU demonstrated inferior performance when coiled, the fibers were left uncoiled for most of this pilot survey. When the IFU bundle is properly uncoiled, it is measured on-telescope to be stable over nightly operating conditions to 1% root-mean-squared (rms) for the most affected fibers and 0.3% rms for the median fiber. We will explore the effect of this potential systematic on the data in Section 3.6. There, we will show that the VIRUS-P fiber stability is not an important issue for emission-line detections, but can dominate the uncertainty in continuum estimates.

The mechanical design of VIRUS-P has been presented in Smith et al. (2008). The instrument's mechanical structures are all made from aluminum to achieve a uniform coefficient of thermal expansion between components and to maintain the optical alignment. The gimbal mount connecting VIRUS-P to the telescope allows VIRUS-P to swing into a horizontal position for any pointing of the equatorially mounted telescope. This ensures that the trace patterns of fibers on the CCD remain constant to high precision over a night. Although a <0.05 pixel trace shift per night is desired, this could not always be accomplished. A trace could shift by up to 0.3 pixels with temperature under some operating conditions. Consequently, data reduction steps were developed to identify and compensate for this subtle systematic; these are described in Section 3. There is not an atmospheric differential corrector installed on

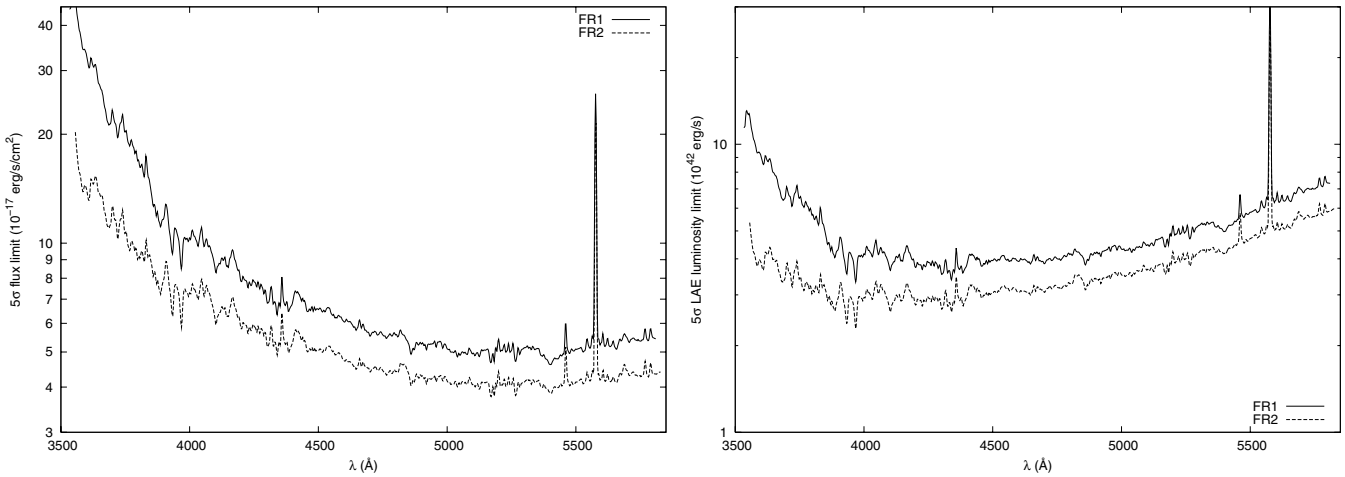

Figure 2. Left: the 5σ detection limit under photometric conditions for an emission-line object perfectly centered in a fiber in three 20 minute exposures. Different source positions can improve or decrease this limit by ∼15% which is captured in our completeness calculation. In both figures, curves are given for the two focal reducers, FR1 and FR2. Right: the 5σ luminosity limit under photometric conditions for objects detected in the Lyα line.

Table 1
Summary of VIRUS-P Observations

Date	Number of Pointings	Fields	Median A_V Used	Range of A_V Used	Number of Emission-line Detections
2007 Oct 4–9	3	MUNICS S2	0.29	0.18–0.44	21
2007 Nov 5–10	5	MUNICS S2;COSMOS	0.33	0.00–1.48	26
2007 Dec 4–9	6	MUNICS S2;COSMOS	0.24	0.23–0.42	36
2008 Jan 3–10	8	MUNICS S2;COSMOS	0.25	0.00–2.84	71
2008 Feb 1–12	6	MUNICS S2;COSMOS;GOODS-N	0.28	0.00–1.63	27
2008 Apr 1–7	3	COSMOS;GOODS-N	0.30	0.00–0.97	30
2008 Apr 28–May 3	3	GOODS-N	0.39	0.00–2.77	43
2008 Jun 3–9	2	GOODS-N	0.31	0.00–1.69	0
2008 Sep 24–29	2	XMM-LSS	0.36	0.19–0.51	28
2008 Nov 24–30	2	COSMOS	0.13	0.07–0.46	8
2008 Dec 22–27	4	COSMOS;XMM-LSS	0.23	0.13–0.47	40
2009 Jan 21–27	3	COSMOS;XMM-LSS	0.23	0.13–0.67	46
2009 Feb 19–23	4	COSMOS;GOODS-N	0.19	0.00–0.53	27
2009 Mar 20–25	4	COSMOS;GOODS-N	0.34	0.00–0.92	29
2009 Apr 20–25	2	GOODS-N	0.20	0.07–0.48	28
2009 May 20–25	1	GOODS-N	0.21	0.00–1.55	5
2010 Feb 9–11	1	COSMOS	0.30	0.10–0.50	0

the telescope. We discuss the atmospheric effects on emission-line source astrometry in the Appendix and the absolute flux calibration of the data in Section 2.4. All observations were taken with airmasses below two.

2.2. Data Collection

We obtained regular fall/winter/spring dark time observations from 2007 September to 2010 February on the McDonald 2.7 m Harlan J. Smith telescope. These observing runs are summarized in Table 1. In total, out of our allocation of 113 nights, 61 were useful for this project. We constructed datacube mosaics in four science fields: the Cosmological Evolution Survey (COSMOS; Scoville et al. 2007), the Hubble Deep Field North (Williams et al. 1996) and the surrounding Great Observatories Origins Deep Survey North (GOODS-N; Dickinson et al. 2003), the Munich Near-IR Cluster Survey (MUNICS; Drory et al. 2001), and the XMM Large-Scale Structure field (XMM-LSS; Pierre et al. 2004). We completed 27, 13, 16, and 4 field pointings, respectively in these fields, by taking three 20 minute exposures at each of the six dither positions. Our effective observation area, accounting for mosaic overlap, is 169.23 □′ over the wavelengths ∼3500–5800 Å with a spectral resolution of ∼5 Å. This corresponds to survey volumes of 1.03×10^6 Mpc^3 h_{70}^{-3} for LAEs and 4.24×10^4 Mpc^3 h_{70}^{-3} for [O II] sources. As described in Section 2.4 and shown in Figure 2, we give the survey's flux and luminosity limits as a function of wavelength under photometric conditions for the case of a spectrally unresolved, point-source emission-line object well centered on a fiber.

In addition to the science data, the following calibration data were obtained one or twice each night. Spectrophotometric standard stars from Massey et al. (1988) were observed. Flats near zenith of the dawn and dusk sky were taken. Calibration with dome lamps was explored but abandoned when none were found with sufficient blue-to-red flux balance. Sets of bias frames were taken and used to construct a master bias for each run. HgCd arc lamps were used to illuminate a dome screen for wavelength calibration. Custom line lists for the HgCd lamps were made by observing the lamps with the 2.7 m Tull Coudé Spectrograph (Tull et al. 1995) at $R = 60$k. The Coudé wavelength calibration was made from ThAr lines. For most of the observing runs, guider frames were saved at

intervals of 2–10 s, depending on the guider star brightness and transparency. The collection of guider frames was prevented 13% of the time due to human error and guider equipment failure. For those observations, the flux calibration was done assuming the median of the observed atmospheric transmission (Section 2.4) from the data set's remaining observations.

2.3. Astrometry

The position of a faint source is not well determined by the IFS data alone since most pointings lack sufficiently bright stars to establish an astrometric solution for the frame. Instead, the positions of stars in the offset guider camera were used to determine the fiber positions; this required precise calibration of the relative astrometry between the fiber array and the offset guider. The relative fiber-to-fiber positions of both IFUs were measured in the laboratory and verified to be very regular due to the precise machining. Illumination and direct imaging in the lab showed that IFU-2 has exceptional uniformity in its fiber matrix, and no deviations from the designed pitch of 340 μm could be measured to an accuracy of 1 μm. IFU-1 is somewhat less uniformity in its fiber matrix than IFU-2. We have mapped the centroid of each fiber to within 0.3 μm, or 0.″007, at the nominal plate scale.

The transformation from guider field position to science field position was calibrated by on-sky measurements. Whenever the guide camera was replaced, we obtained data under a six dither pattern on open clusters at low airmass. In total, seven astrometric solutions were derived, each yielding the plate scales, offsets, and rotations of two image planes under a standard tangent projection (Greisen & Calabretta 1993). We found adequate fits with constant plate scales determined for each IFU axis yielding 12 degrees of freedom in a nonlinear transformation from guider and IFS pixel positions to celestial coordinates. We first determined guider positions by using SExtractor (Bertin & Arnouts 1996) to measure the positions of stars and match to coordinates from the United States Naval Observatory's (USNO) Nomad catalog (Zacharias et al. 2005). Similarly, the continuum intensities of USNO stars in the fibers were measured by summing flux over the wavelength range 4100 Å $< \lambda <$ 5700 Å; this region was chosen to mimic the guider wavelength response and minimize atmospheric refraction differences. Fibers containing signal significantly above the noise were matched with significant detections in adjacent fibers. Centroids were calculated for each source and again matched to the Nomad catalog. A simplex method (Press et al. 1992) was then used to find the least squares minimum robustly in the presence of the many local minima. We show in Figure 1 the fit quality in a representative solution. The range of systematic uncertainty in our seven eras of astrometric solutions was 0.″17–0.″51 with a median of 0.″31.

We further measured the stability of the astrometry over many months from flux standard stars. We anticipated any drift to be negligible due to the design of plastic pins which located the IFU head against the telescope mounting surface. However, we found substantial month-to-month systematic variations of order 1.″8 rms. The only clear dependence was a declination term with temperature, which we attribute to a thermal expansion of the guider camera mount. However, this expansion cannot explain the bulk of the astrometric scatter. Since we find much smaller astrometric scatter in any 1 month, the monthly removal and remounting of the IFU input head from the telescope between observing runs is the plausible source of drift. So, we have chosen to estimate an empirical month-by-month offset in the astrometric zero point which lowers the median monthly rms to 0.″6 and ranges from 0.″0 to 1.″0.

Coarse positional sampling by the large fibers and low signal-to-noise (S/N) limitations forms the final component of the astrometric error budget. In order to quantify this uncertainty, we have simulated the positional recovery for a range of emission-line sources. We describe those simulations in Section 4.3. The result is a fit to the random astrometric uncertainty with a functional form of $\sigma_{r,\mathrm{random}} = 0.''348 + 2.''04/(\mathrm{S/N})$.

We can assess the completeness of our error budget by measuring the observed positional offsets of emission-line objects found with high confidence counterparts. As explained in Section 5, a comparison of our fiber detections with broadband imaging shows that 55% of our emission-line detections have an isolated counterpart detected with $\geqslant$90% confidence. Through a comparison, we find a mean offset of $\Delta\alpha = -0.''53 \pm 0.''05$ and $\Delta\delta = 0.''39 \pm 0.''05$ between the fiber-based emission-line source positions and the broadband photometric centers. The source of this offset is not certain, but we apply it to all our reported emission-line positions. After correcting for this offset, the counterpart associations were iterated to produce our final emission-line positions. In Figure 3, we present the distribution of the data offsets to test the error budget. This error budget serves as an important input in the method (Section 5) for assigning broadband counterparts in crowded fields to the emission-line sources.

2.4. Flux Calibration and Transparency

The majority of the observations were not taken under photometric conditions, hence a proper flux calibration requires a real-time measurement of the atmospheric transparency. Unlike some modern wide-field imagers, the VIRUS-P FOV is not large enough to contain photometrically calibrated stars in the majority of its arbitrary pointings. However, the offset guider with a larger FOV has a size sufficient for this continuous calibration purpose. We recorded all guide camera exposures sampled at 2–10 s that were contemporary with the IFS science exposures. The guider exposure times varied depending on the guide star brightness. Basic bias-subtraction and flat-fielding reductions were implemented on the guider frames. We performed aperture photometry on all stars detected. When available, we used Sloan Digital Sky Survey (SDSS) measurements (Adelman-McCarthy et al. 2008) for our calibrations; otherwise we used the USNO-B1.0 survey (Monet et al. 2003). The SDSS photometric precision is quoted at below 1% for guide stars used, typically $V < 19$. The USNO-B1.0 photometric precision is typically much worse, ~0.25 mag, and this directly leads to an important uncertainty in line fluxes for objects in the MUNICS and XMM-LSS fields. Accordingly, we have added in quadrature a 15% error, assuming the median of three guide stars per field, to the flux and EW measurements for the MUNICS and XMM-LSS sources. We treat these errors as random, since multiple and independent sets of stars were used in different mosaic pointings and multiple spectrophotometric standards were observed. A color term was fit from the guider data considering its non-standard, wide-bandpass filter, a new zero point was calculated each month to correct for periodic equipment changes and mirror cleanings, and non-photometric extinctions were found for each frame after removing a standard airmass term of 0.186 mag AM^{-1}. Typically, we had 2–5 stars per field that were bright enough for this purpose. The resultant distribution of zero-point offsets due to transparency, Δzp, is given in Figure 4. By measuring the scatter in the zero-point

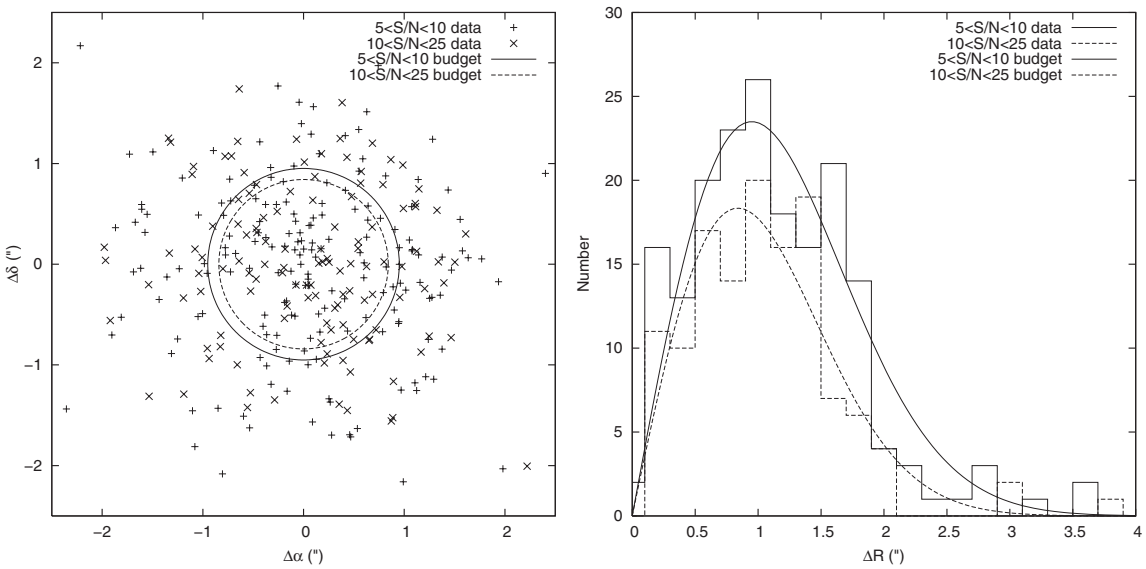

Figure 3. Left: the positional offsets from catalog emission-line detection positions and broadband counterpart image positions divided into two S/N bins. The residual correction discussed in Section 2.3 has been applied. The 1σ astrometric error budget for each bin is also drawn with radii 0.″95 for $5 < S/N < 10$ and 0.″84 for $10 < S/N < 25$. Right: a histogram of the same data shown with a Rayleigh distribution. The same dispersions are used to demonstrate the appropriate characterization of the astrometric error as a two-dimensional Gaussian function.

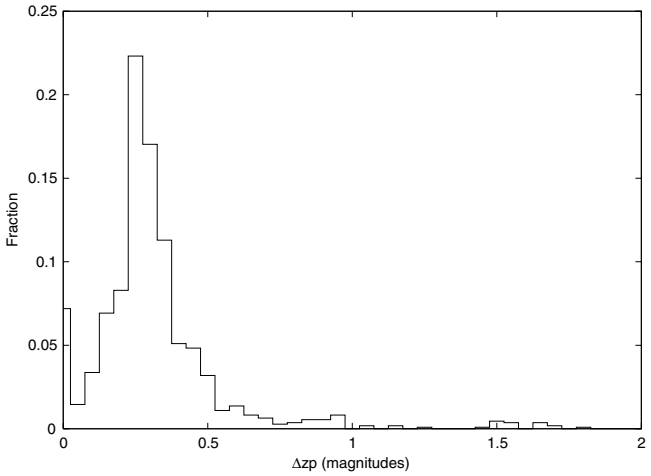

Figure 4. Distribution of zero-point offsets due to non-photometric transparency as measured with the guider camera. This distribution represents the best 60th percentile of the observing allocation with the remaining 40% being too poor to guide or requiring dome closure.

offset from all the stars available in each frame, we find a mean uncertainty of 6% in the guider-based photometric correction.

The flux calibration of IFS data was done in a manner similar to that for long-slit spectroscopy, but with some additional steps to compensate for fiber sampling patterns. We used the spectrophotometric stars and calibrations of Massey et al. (1988) observed under a six dither pattern. Airmass extinction coefficients for photometric conditions with a curve specifically modeled for McDonald Observatory are applied. This extinction curve is similar to the Kitt Peak curve supplied with IRAF. The bright standards allowed us to determine both the source position relative to the fiber grid and the seeing point-spread function (PSF), which in turn yields the exact fiber sampling. In contrast, fainter emission-line sources require statistical sampling corrections that are discussed in Section 4.4. In order to determine the percentage of incident flux captured over the six dither positions, we employed the following analysis. We began by considering the spectra for all fibers positioned within a large radial aperture (operationally, 8″) from the stellar centroid. and adopting a seeing model with a 2D circular, Gaussian PSF. The broadband flux of each fiber was measured by summing over a large wavelength range (operationally, $4000\,\text{Å} < \lambda < 5500\,\text{Å}$). The PSF and Gaussian normalization were determined through a nonlinear least squares minimization by assuming the spatial response of each fiber was top hat. The sampling correction was then formed from the ratio of the Gaussian normalization to the sum of the broadband flux measurements. Then, the spectral count rates of the relevant fibers were resampled to a common wavelength scale, co-added, and normalized using the sampling correction. By using such a broad, circular aperture, we ensured that the effects of atmospheric differential refraction (ADR) on the co-added spectrum were negligible. The final spectral flux calibration curve was then formed from the ratio of the published, absolute flux density to the sampling corrected data count rate. Spectrophotometric standards were taken under a range of conditions, so their comparison required a further correction for transparency as estimated from the guider measurements. Once done, we find an rms between all flux calibration curves of 9.3% and 8.5% for FR1 and FR2. We find no trend with wavelength in this scatter and so validate the assumed gray zero-point correction for all guider transparencies at these levels of uncertainty. The final catalog will list the random line flux errors, but the whole sample may be considered to also be subject to the ∼10% flux calibration systematic uncertainty just discussed. We do not fold the systematic into the tabulated values as relative comparisons within the sample should not be subject to it.

Several statistics from this flux calibration analysis summarize the survey's performance. First, the range of atmospheric transparencies for recorded data is shown in Figure 4. These statistics are biased against periods of weather too poor to attempt observation and represent only the best 60% by time. The median nonphotometric transparency penalty to this survey in the observable periods is 0.28 mag. The total system throughput is shown in Figure 5 as the fraction of light recorded after passing through one photometric airmass (zenith), the telescope, the

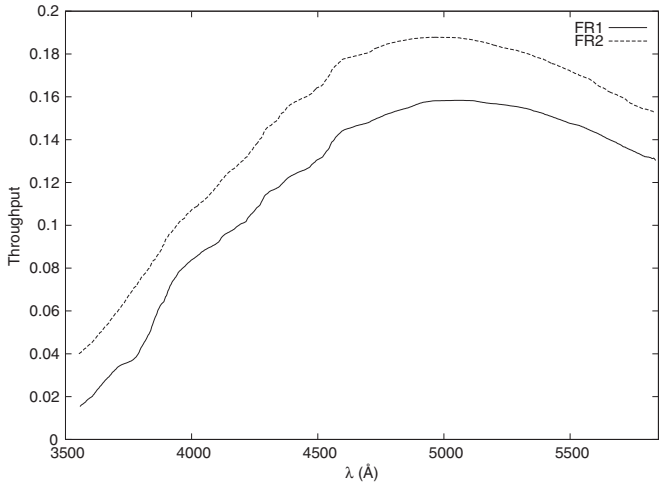

Figure 5. Total system throughput of VIRUS-P, the 2.7 m telescope, and the atmosphere at an airmass of one. Curves are given for the two eras of focal reducers, FR1 and FR2.

focal reducer, and the VIRUS-P instrument. The curves for the two focal reducers show a dramatic difference: FR2 performs better than FR1 at all wavelengths, but particularly in the blue where FR1 has only half the throughput of FR2.

The combination of read noise, system throughput, and sky brightness determines the detection limit for an unresolved emission-line source. Figure 2 shows the 5σ limit in a detection element (defined as $\pm 2\times$ the instrumental dispersion or $\pm 1.9\times$ binned pixels), which is nominally the survey's photometric limit with some modulation for sources sampled under different fiber positions. The luminosity limit for LAEs is also shown in Figure 2. The exact limits will be further explored in Section 4.1 and compensated for with the completeness limit derived in Section 4.3. Finally, in Figure 6 we give the sensitivity maps at 4500 Å for spectrally unresolved point sources, taking into account mosaic overlap, bright objects, dead fibers in IFU-1, guider measured extinctions, and the range in airmass over the data set. Small gaps in the map are due to the slightly different sizes of IFU-1 and IFU-2, and the failure to complete the desired

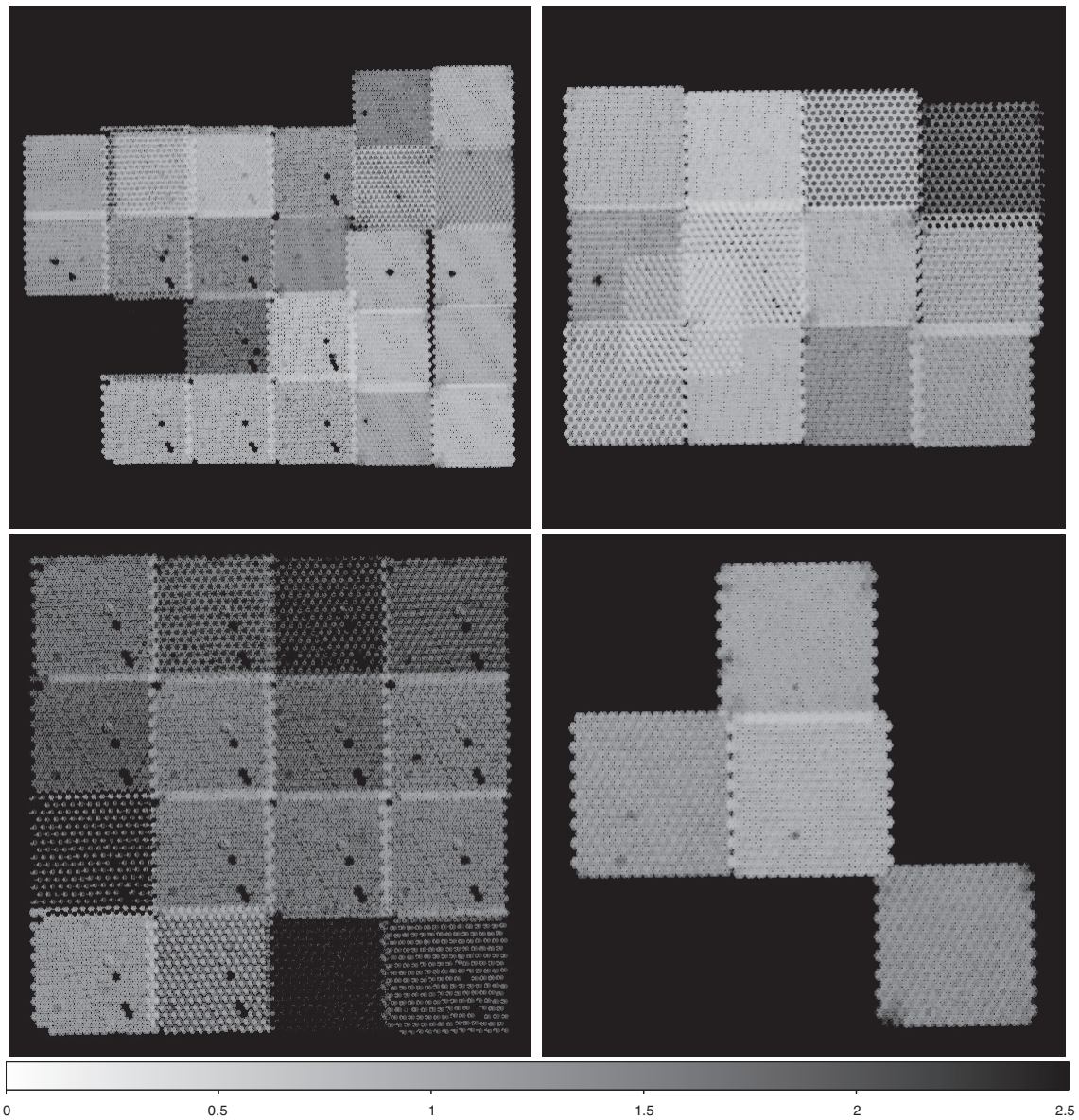

Figure 6. Sensitivity maps (1σ per detection element) at 4500 Å in 10^{-17} erg s^{-1} cm^{-2}. The three broken fibers in fiber bundle IFU-1 are evident. Top left: COSMOS, top right: GOODS-N, bottom left: MUNICS, bottom right: XMM-LSS.

Table 2
Ancillary Broadband Imaging Properties

Field	Central α (J2000)	Central δ (J2000)	$E(B-V)$	Filters	FWHM[a]	Stack[b]	Band$_K$[c]	Depth[d]	VIRUS-P Area ($\square'$)
COSMOS	10:00:30	+02:15:04	0.018	u^a,B_J,V_J,r',i',z'	1.33	$B_Jr'i'$	V_J	26.5	71.6
GOODS-N	12:36:51	+62:12:51	0.012	U_J,B_J,V_J,R_J,I_J,z'	1.26	B_JR_JI	V_J	26.6	35.5
MUNICS-S2	03:06:41	+00:01:15	0.083	B_J,g',i',z'	0.99	$B_Jg'i'$	g'	25.8	49.9
XMM-LSS	02:21:20	-04:30:00	0.027	u^a,g',r',i',z'	0.97	$g'r'i'$	g'	25.8	12.3

Notes.
[a] The worst seeing FWHM in $''$ to which all bands are matched.
[b] Filters combined to form the detection image.
[c] The band chosen for Kron aperture measurement.
[d] The 5σ limit in AB magnitudes for a point source in a $2''$ aperture for the band with the Kron aperture measurement.

six dither pattern in one COSMOS pointing by only completing a three dither pattern. Finally, five fields were chosen to overlap with previous fields for cases where transparency in the first pass yielded poor depth.

2.5. Ancillary Imaging

This survey discovers and spectroscopically measures LAEs in one pass, as opposed to narrowband surveys that often require spectroscopic confirmation on a subsample. The depth and bandpass restrictions of VIRUS-P, however, still make discrimination between LAEs and low-z contaminants challenging. For both LAEs and [O II] emitters at many redshifts, we expect to have only one strong emission line in the VIRUS-P bandpass. [O III]λ5007, [O III]λ4959, and Hβ will be lost at $z > 0.158$, $z > 0.170$, and $z > 0.193$, respectively, and the survey's spectral resolution does not resolve the [O II] doublet. Furthermore, the variation observed in local galaxies for strong line ratios (Kennicutt 1992) never guarantees that two statistically significant lines will be detected. By necessity, we resort to an EW cut, as used extensively in LAE narrowband surveys, to classify single emission-line detections. We discuss the EW cut further in Section 5. However, the VIRUS-P spectra are not sufficiently sensitive for continuum detections for the majority of the emission-line detections. To reach the necessary limits, we must supplement the spectra with deep imaging.

These data set's fields are located in regions of the sky with existing deep images and catalogs (Drory et al. 2001; Fernández-Soto et al. 1999; Capak et al. 2004, 2007; Ilbert et al. 2009). The XMM-LSS field does not have a published catalog but is covered by the Canada–France–Hawaii Telescope Legacy Survey[15] (CFHTLS) wide-field W1. The deep MUNICS images, which were not part of the original publications, consist of B_J, g', i', and z' data taken with the Large Area Imager for Calar Alto (LAICA) on the Calar Alto Observatory 3.5 m, with zero points made by matching stellar photometry to the published catalog. Instead of using the literature catalogs, we have chosen to produce our own SExtractor catalogs on the images and error maps; this ensured a consistent analysis for the fields and pushed the S/N to a lower threshold for a more complete emission-line association. We list select properties of the relevant broadband data in Table 2. The table also gives the Galactic extinction values (Schlegel et al. 1998) we applied to the continuum and emission-line fluxes under the extinction curve fit of O'Donnell (1994).

Care was taken in the photometry to ensure our photometric colors were robust. Two measures of seeing FWHM are relevant: the one for the particular band where a Kron (Kron 1980) aperture is measured (FWHM$_{\rm Kron}$) and another larger value to which the other photometric bands will be matched (FWHM$_{\rm match}$). For each field, we formed a detection image by stacking the deeper available bands without matching each band's seeing (see Table 2). The detection parameters of SExtractor were then set to find a minimum of three neighboring pixels detected with 1σ significance over sky without filtering. Since we will only be using sources with 3σ significance in their photometry, the exact detection weights and filters have little importance. Also, the return of spurious continuum sources from the low significance thresholds is acceptable for our application. A chosen band with good depth for each field, labeled here as i, was compared to the detection image using SExtractor dual image mode, in order to measure flux densities in a blending corrected Kron aperture, $\hat{f}_{\nu,i,\rm Kron}$. The Kron ellipse dimensions a and b were also measured. Blending correction was crudely accomplished with the SExtractor `AUTO` flux measurements and the flag `MASK_TYPE` set to `CORRECT`. Under this setting, SExtractor sums the flux from the opposite side of the Kron aperture whenever it encounters pixels covered by multiple Kron apertures. In the remaining bands, labeled here as j, each frame was matched in seeing to FWHM$_{\rm match}$ and run in dual detection mode to measure the flux density in a circular aperture of diameter $1.4 \times$ FWHM$_{\rm match}$, $\hat{f}_{\nu,j,\rm circ}$. The term $f_{\rm corr} = (1 - e^{-0.5ab/\sigma^2_{\rm Kron}})$ then forms a correction factor for the fraction of flux lost to the Kron aperture from a point source under seeing with dispersion $\sigma_{\rm Kron}$. The final aperture-corrected flux density in each band j was then estimated from Equation (1). Standard error propagation was applied

$$\hat{f}_{\nu,j} = \frac{\hat{f}_{\nu,i,\rm Kron} \times \hat{f}_{\nu,j,\rm circ}}{\hat{f}_{\nu,i,\rm circ} \times f_{\rm corr}}. \quad (1)$$

This resultant source catalog was used only in cross-correlation with our VIRUS-P emission-line catalog to identify object counterparts. The method of assigning counterparts is described in Section 5. The emission-line fluxes are subtracted off from the broadband measurements according to the filter transmission curves as supplied by Brammer et al. (2008) once counterparts are assigned.

[15] Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada–France–Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the CFHT Legacy Survey, a collaborative project of NRC and CNRS.

3. DATA REDUCTION

The science goals of this survey required the development of a custom reduction pipeline. Several IFS reduction pipelines already exist (e.g., Valdes 1992; Zanichelli et al. 2005; Turner et al. 2006; Sánchez 2006; Sandin et al. 2010) and are well suited to many applications. In particular, we first tried using a predecessor of p3d (Sandin et al. 2010; Becker 2001). The crucial limitation of the p3d package and all other IFS pipelines at the time is that they resample the spectrum of each fiber onto a common wavelength scale at some step in the processing. This step correlates errors and complicates the detection statistics. In fact, we found by running simulated, source-less VIRUS-P data through p3d that many more resolution elements were flagged to have 5σ significance than was possible from the input Poisson statistics. The use of p3d would have either produced too high a contamination fraction or required higher S/N cuts and survey flux limits. This consideration led us to develop a set of scripts and FORTRAN routines collectively called Vaccine. Many of the pipeline steps are standard to all spectroscopic reductions. However, the primary Vaccine requirement to avoid data resampling is done in a manner similar to the Kelson (2003) pipeline developed for long-slit spectroscopy and affects the flat fielding and sky subtraction steps.

3.1. Preliminaries

The first operation done to each VIRUS-P frame is to measure a single bias value from the overscan regions, subtract it from the frame's data section, and trim the overscan. A master bias then is created from all the overscan-subtracted biases taken during an observing run (typically 100–200 frames). Overall, the noise statistics in bias frames were remarkably stable and indistinguishable over weeks. Next, we cleaned the images with a bad pixel mask made by exposing the camera to scattered white light and finding the pixels with relative quantum efficiency outside 10% of the CCD's median. The VIRUS-P CCD has very clean cosmetics: besides the two rows nearest the readout register, this bad pixel mask only contained 13 total pixels in three patches. Data combination for all co-additions of frames is accomplished using the biweight estimator (Beers et al. 1990); this algorithm was chosen for its robust performance regarding outliers such as cosmic rays. The master bias and individual overscans are subtracted from all calibration, science, and flux standard frames. Calibration frames, consisting of arc frames and twilight flats, are taken at the beginning and end of each observing night. The dawn arcs and flats were preferentially used over those frames taken in dusk, as they were a better match to the temperature of the night-time conditions.

As is common to both IFS and slitlet multi-object spectroscopy, the traces of all fibers are not strictly parallel to the CCD pixels or to each other. The fiber profiles, taken from a flat-field calibration, must be traced to define an extraction aperture of each fiber. Moreover, the dispersion axis is not necessarily parallel to each fiber's trace. However, with the camera alignment in VIRUS-P, we found the maximum deviation of this misalignment is 0.2 resolution elements, so we ignored this distinction and defined the dispersion axis along the fiber trace to be perpendicular to the cross-dispersion direction. This assumption effectively broadens, slightly, the resolution in some fibers. The tracing is then made by fitting Gaussian functions to cuts along the cross-dispersion axis at a series of wavelengths for each fiber. The Gaussian centers are fit by a fourth-order polynomial across the CCD. This fit was tested against repeated flats and shown to be precise to <0.1 pixels across the CCD. Trace information is displayed for the user, who can iterate the fit tolerances if required. All further operations are done in the traced coordinates with cross-dispersion apertures of 5 pixels. Vaccine propagates errors for all operations starting with the read noise and keeps track of the Poisson noise from sources and the background sky.

3.2. Wavelength Calibration

An automated peak finding algorithm is run on the arc lamp frames, and line identifications are made from a user entered initial wavelength solution. Typically, seven unblended HgCd lines are found with their central pixel locations determined by a Gaussian fit to the line profile. The pixel-to-wavelength mapping is then fit with a fourth-order polynomial in the dispersion direction. The first-order term of that polynomial is found to vary smoothly for all fibers as a function of the cross-dispersion direction. Hence, for increased accuracy, this first-order term is refit as a function of the cross-dispersion distance from the camera optical axis using a fourth-order polynomial. Finally, the wavelength polynomial as a function of dispersion direction pixel is refit, this time with the constrained first order term. The residuals of this procedure are typically one hundredth the size of a resolution element and the solutions are stable to a tenth of a resolution element over several weeks.

The heliocentric correction is found for each frame by using a FORTRAN implementation (written by G. Torres[16]) of the IRAF task bcvcorr in the rvsao package (Kurtz & Mink 1998). The small, <1 km s^{-1} differences in heliocentric velocities for exposures at the same dither position but taken over different nights are ignored and only the mean heliocentric correction between them is applied. All reported wavelengths are in the heliocentric frame. A correction to vacuum conditions is made assuming an index of refraction for air of $n = 1.0002$ for all observed wavelengths.

3.3. Flat Fielding

Typically 15 twilight flats were taken each night and combined using the biweight estimator. To ensure high S/N in the twilight flats, each frame was exposed to near but below the CCD's 1% nonlinearity specification which occurs at 50% of full well. Four signals are present in the twilight flats, (1) the solar spectrum, (2) the relative throughputs between fibers, (3) the fiber profile in the cross-dispersion direction, and (4) the relative pixel-to-pixel responses. To remove the first of these we employed a bspline fit (Dierckx 1993) constrained by input from large subsets of fibers. Such a fit is robust against outlier datapoints (i.e., our cosmic rays or faint sources that fill a subset of the data) and fits curvature that a linear interpolation would miss. The advantage of the bspline fit is best leveraged when a spectrum is highly supersampled, and the camera's optical distortions naturally deliver this quality in different fibers, predominantly as a smooth function of cross-dispersion direction. However, the slight (10%) spectral resolution variation across the CCD disfavors a single fit for all the fibers' data. As a compromise, we consider each fiber with its 20 nearest fibers in CCD coordinates. Within these sets the spectral resolution variations at any wavelength are less than 2%. We do not make more complicated corrections for the spectral resolution variation beyond this. The bspline fit for each fiber, serving as a model of

[16] http://tdc-www.harvard.edu/iraf/rvsao/bcvcorr/bcv.f

the solar spectrum, is then divided into the original flat-field data, resulting in a precision between different sets of frames to <1% rms.

3.4. Background Subtraction

The majority of VIRUS-P fibers and resolution elements in this blind survey record blank sky. This enables the noise if our sky model to be driven down by stacking measurements over many fibers, so long as the noise is statistical. By using the 50 nearest fibers in the cross-dispersion direction, the statistical noise in the sky can be reduced to only 14% of a single fiber's noise. In this way, the uncertainty in the post-sky-subtracted data can be made very close to that of the pre-sky-subtracted data (as long as the flat-fielding systematics are understood). Our sky background models were formed identically to the flat-field models.

We note, however, that this semi-local sky estimation method is only robust for sources that fill a small fraction of the combination window, which on-sky is approximately $\Delta\alpha = 100''$ by $\Delta\delta = 20''$. No bright, broadband sources have such sizes in the survey fields. Moreover, in order to further avoid oversubtracting bright sources, we constructed an object mask prior to the bspline fit. Any fibers that yield $>2\sigma$ significance in the continuum, as estimated by combining the data and errors across all VIRUS-P wavelengths, were placed in the object mask.

3.5. Data Combination

The count rates in the three frames taken at each dither position were first corrected by the airmass-based photometric extinctions and the guider-based transparency measurements and then combined. The three frames and the 5 pixel cross-dispersion aperture delivered 15 input values to the biweight estimator at each wavelength. The VIRUS-P flux standard frames are passed through Vaccine exactly as the primary science data. Finally, the science spectra (and errors) are scaled by the flux calibration (Section 2.4) to form a set of calibrated, one-dimensional spectra at each fiber and dither position.

3.6. Systematic Errors

We identify three potential sources of systematic error in VIRUS-P data, one unimportant, and two that require monitoring. First, we discuss why crosstalk between fibers is not important in VIRUS-P data. Next, we identify the effects of throughput variations and the accuracy of flat field cross-dispersion profiles on the error budget as the most prominent systematics. Finally, we describe an empirical, frame-specific estimate of the systematics that must be added to the random errors.

IFS crosstalk occurs when the profile of a fiber in the cross-dispersion direction significantly overlaps that of other fibers projected nearby on the CCD. We make no crosstalk correction in Vaccine for two reasons. First, the fibers are measured to have cross-dispersion profiles of 4 pixel FWHM size. This is a factor of two smaller than the center-to-center fiber spacing on the CCD and results in larger relative fiber spacing than in many IFS instruments. As a result of our 5 pixel extraction aperture, sources of equal strength in neighboring fibers imply only a <0.5% contamination. Second, the blind field selection of this survey leaves most fibers seeing only uniform sky background and leaves little risk from cross-talk contamination. A fiber aligned on a source will usually be isolated and trade an equal flux from the background sky with its crosstalk neighbors. The flux calibration (Section 2.4) steps use the same cross-dispersion aperture, and therefore correct for the source flux lost by crosstalk.

The stability in the throughput of fibers can cause significant systematic errors in some measurements. As discussed earlier, our fiber throughput is very stable, with 1% rms variation at worst and 0.3% median variation over a night. However, our background sky is 25–40× stronger than the statistical noise limits in each resolution element. As a result, the systematics can overwhelm the statistical errors in spectral apertures of six resolution elements or more during the worst stability conditions. Continuum estimates using large wavelength ranges may thus be severely affected in our survey, and we make no claims on such properties. However, the situation for emission lines is far better. First, the systematics in a detection element (approximately two resolution elements) are at worst 56% of the statistical error and at median are 13% before background subtraction. Second, most of the throughput variation is captured in the background subtraction step. As described in Section 4.1, before we detect emission lines we subtract off a locally estimated continuum value using roughly 90 independent spectral pixels. Since fiber throughput variations manifest uniformly across wavelength, the spurious signal is a small multiple of the sky spectrum and relatively featureless over our bandpass (exempt for the bright [O I] 5577 Å sky line which we mask prior to all detections). The systematic error in a post-background-subtraction detection element therefore drops to 5.9% of the statistical error in the extreme case and 1.4% of the statistical error in the median case. We include this systematic uncertainty in both the detection and flux calibration error budgets via the empirical correction described below.

The final known source of systematic error is occasional variability in the cross-dispersion profile that occurs with time and temperature for different fibers. These profile changes can appear as both a trace position shift and a width change, and while small, are important. Between twilight flats spaced 8 hr apart and through maximum dome temperature changes of 10°C, we have measured trace shifts of up to 0.3 pixels and profile FWHM changes of 0.3 Å. Our goal was to limit this systematic to 10% for any pixel in the flat. The FWHM variation already meets this criterion, but the maximum trace shift is too large by a factor of six. Moreover, although the trace shift also appears to be coherent between adjacent fibers on the CCD, it sometimes goes in opposite directions at the opposite ends of the fiber bundle, as if the traces are subject to a "breathing mode." We have developed a heuristic solution that mitigates this problem. The core idea is to measure the offset over subsets of fibers, alter the flat fields to maintain the fiber-to-fiber and pixel-to-pixel patterns but resample the fiber profile to produce a shifted flat tailored to each exposure.

For each pre-sky-subtracted data frame, the fiber centroids at each wavelength along the cross-dispersion direction are calculated with respect to the corresponding flat. These trace shift estimates are then median smoothed with their 12 nearest fibers on the CCD. Rather than presume a cross-dispersion profile shape, which displays non-Gaussian features, we use sinc interpolation to resample the profile. Linear interpolation fails to recover the strong curvature in this profile. In each fiber and each wavelength, the flat field is resampled at the fiber-specific estimated offset relative to the polynomial trace peak. However, additional smoothing is still required to leave pixel-to-pixel features unaltered. To do this, we run a boxcar smoother

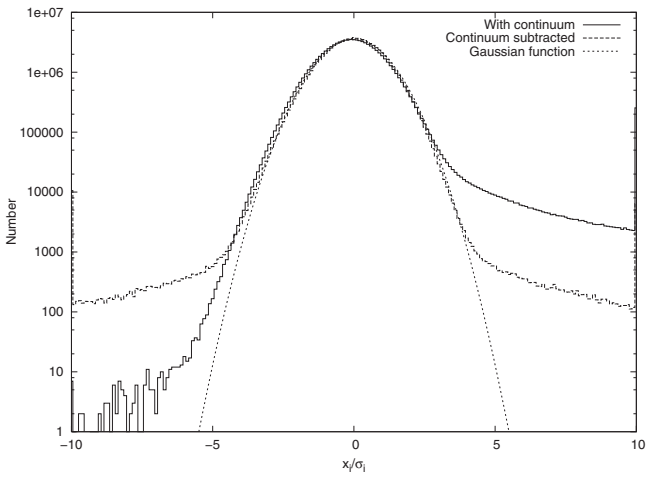

Figure 7. Histogram of the ratio between the reduced data and error for all 87.8 M independent spectral elements in this survey. Most elements only see the sky background with residuals consistent to a normally distributed noise model. A small but obvious fraction of the elements also see bright, positive signal from continuum sources. The normalized Gaussian function shows the expected distribution in the absence of signal and systematics. The influence of the fiber throughput systematic (Section 3.6) likely broadens the distribution. This is most evident on the negative side which becomes nearer the normal distribution after background subtraction. The distribution after continuum subtraction appears much more symmetric and with a better matching width. The data wings at high and low ends represent the fiber positions with strong continuum. The boxcar-based continuum fitting is a rather crude tool that does not characterize all the continuum signal, but the emission-line catalog is uncompromised by its use.

of 81 pixels along the dispersion direction for both the original flat field and the sinc resampled flat field. The biweight of each forms a pure profile model in the original and resampled frames, and the pixel-to-pixel variations are isolated in a separate image. The total, shifted flat is then formed by multiplying the pure, shifted profile model by the isolated pixel-to-pixel estimate. A final scaling is then applied to maintain the fiber-to-fiber throughputs and total flat normalization, as sinc interpolation does not automatically conserve flux. The use of these shifted flats rather than the original flats results in lower systematic errors and meets the goal of $<10\%$ flat-field profile error.

To capture any remaining systematics we have made a second, independent estimate of the error using the rms of the 15 measurements that go into the final data combination (Section 3.5). This error estimate is itself noisy, but the ratio between this empirical error and our formal error over all pixels is useful as a diagnostic. We find the median of this ratio per frame is 0%–20% above the random noise alone, and the median over all data is 5%. Therefore, we increase the errors by this amount prior to the detection steps. Figure 7 shows the distribution of all 87.9 million independent datapoints divided by the error estimates of this data set. Versions prior to and after continuum subtraction are shown. If the data set were entirely without signal, if all the systematics were understood, and if all the noise were uncorrelated, the distribution should match the given Gaussian function with a dispersion of unity. Clearly the distribution is asymmetric, distorted on the positive end by signal and the negative end presumably by the previously discussed fiber throughput variations. However, the continuum subtracted data with the fiber throughput variation removed are much more symmetric and show a distribution that is a much better match to the Gaussian width. Emission-line objects are detected in the continuum-subtracted data, and the noise model is validated.

4. EMISSION-LINE SOURCE SELECTION

The controlled selection of emission-line objects is the next step in producing this survey's catalog. The primary task of the detection process is to optimally use the source signal that has been distributed into, potentially, several fibers. The challenge is to push to a high completeness level at low S/N under a contamination constraint. The approach we adopt is to define emission-line detection seed apertures at a low S/N significance, test the combination of the seed apertures and all nearby fibers on sky, and allow the aperture to grow if the significance of the encompassed signal increases. The growth process is iterated. To understand the completeness and contamination rates of this method, we also present simulations with mock data. In similar data sets such as blind long-slit spectroscopy (Gilbank et al. 2010) and grism spectroscopy (Meurer et al. 2007), detection algorithms based on data convolution have been used. We have tested this approach on our data set, but found it inferior in completeness to our adopted technique (see Section 4.3).

4.1. Detection Method

Several terms require definition before we describe the detection method. A fiber position carries a set of neighboring fibers, defined as all other fibers offset by $\leqslant 3''$ in their center-to-center coordinates. The detection aperture starts with one fiber and, by iteration, is allowed to grow by accepting neighboring fibers. A detection aperture may be composed of multiple fibers and has its own set of neighbors, defined as the union of all neighbors to the current member fibers. The S/N of a potential emission line is calculated in a specific spectral window around the fit central wavelength. We define this detection window as spanning $\pm 2\sigma_{\rm res}$ where $\sigma_{\rm res}$ is the dispersion of the VIRUS-P resolution element (2.2 Å). Within this window, data are summed and errors added in quadrature. Pixels that straddle the window are included by their fractional overlap.

We begin with the fully calibrated spectra, errors, and fiber sky coordinates. First, a local continuum for each fiber is estimated and removed through a 200 Å wide biweight boxcar. Second, seed apertures are defined as all pixels that have 1σ positive significance under a 6 Å wide boxcar smoothing. Seeds are merged when found in the same fiber and at contiguous wavelengths. Third, a Gaussian model is fit to each seed with variable width, wavelength, and intensity using a data window of 30 Å. We anticipate emission-line widths for LAEs to lie below the VIRUS-P spectral resolution, but the detection method is designed to be general to all line widths. We experimented with basing the detection aperture on the Gaussian function's fit width instead of the instrument's resolution, but simulations showed that the broad fits produced an unacceptable level of contamination. Fourth, fits with the seed apertures and each of the neighboring fibers are made. When making fits using multiple fibers, each fiber's emission-line intensity is allowed to vary, but constrained to a common wavelength and width. Fifth, if the inclusion of any prospective neighboring fiber increases the total S/N over a particular threshold, the fiber with the greatest increase is added to the detection aperture. Operationally, we use a threshold of $\Delta {\rm S/N} = 0.3$. Sixth, these steps are iterated until the apertures no longer grow or the aperture size reaches six fibers. The cut at six fibers is chosen because in the dither pattern, a point source can be equidistant from at most six fibers. Seventh, a final significance cut is made on the potential detections. If the detections had only been made using single, independent apertures, simple counting statistics

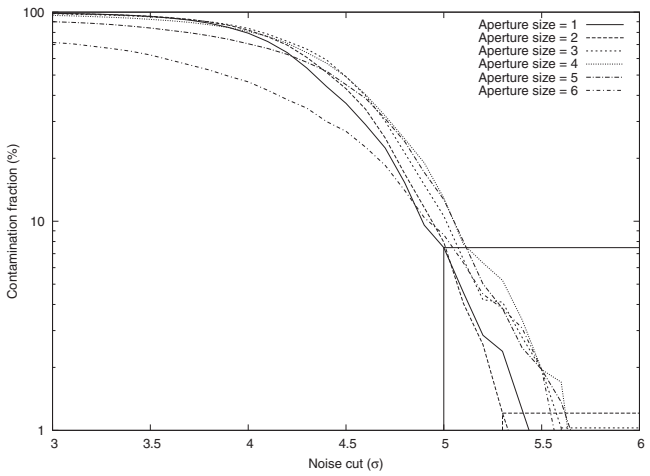

Figure 8. Fractional contamination to the LAE sample by S/N cut. The detection method used is described in detail in Section 4.1. The LAE predictions are the same as in Section 2.4. Curves are given for differently sized collections of neighboring fibers. The growth of apertures was allowed whenever the S/N was increased by >0.3 with the inclusion of another fiber. At a high, constant S/N, the greatest contamination is produced by large apertures. The optimum-sized aperture for a point source under all dither-source geometries is two at median. To optimize our selection, we make a staggered series of S/N cuts based on the number of fibers used, N, as S/N $> 5.0 + 0.3 \times (N - 1)$. The horizontal and vertical lines show the evaluation points of this cut to the simulation curves. This procedure predicts 10% ± 1.6% noise contamination.

could be used to meet the <10% contamination goal. For example, when applied to the luminosity function of Gronwall et al. (2007), our S/N $\geqslant$ 5 cut and no galactic extinction implies that we should see 2.4 LAEs per VIRUS-P pointing under photometric conditions. Similarly, a VIRUS-P pointing (over six dithers) has 756k independent resolution elements, so an S/N $\geqslant$ 5 cut would deliver 8% contamination. Unfortunately, the more complicated detection algorithm used here is not so straightforward to assess. While the growth steps will recover some sources that would otherwise be missed, they can also bundle noise from neighboring fibers. We therefore have made simulations of mock noise frames in order to optimize our selection thresholds.

4.2. False Source Tests

To test for false sources, we began by simulating full, two-dimensional spectral data for 25 VIRUS-P fields using the observed median sky brightness. The mock data were made with noise realizations from the actual sky background and CCD read noise but were otherwise without sources. The fields were then analyzed for emission-line sources exactly as in Section 4.1 for all detections that reached S/N $\geqslant$ 3. The number of spurious sources was then compared to the expected number of true LAEs (Gronwall et al. 2007) as a function of S/N cut, aperture, size, and survey depth. Evidence indicates that the LAE luminosity function does not evolve strongly at $z = 3$ and higher redshifts (Ouchi et al. 2008), but there is less certainty about the rate of evolution over the lower redshifts that we also probe (Nilsson et al. 2009; Cassata et al. 2010). The results of this analysis are shown in Figure 8. Interestingly, at higher S/N the larger apertures begin to contribute the most contamination. Under the typical survey observing conditions and the majority (∼80%) of source-fiber geometries, the optimal number of fibers to include in a simultaneous detection is two. Point source emission objects, which we anticipate most LAEs to be (Bond et al. 2010), rarely (<5%) benefit from fiber apertures of four or more. Conversely, the S/N for extended low-z objects is often improved by including more fibers, so we should not avoid large apertures altogether. Finally, it is clear that a common cut of S/N $\geqslant$ 5 would deliver an unacceptably high rate of (60%) contamination. The situation can be improved by varying the S/N limit as a function aperture size. The choice we adopt is for an aperture of N fibers to have an S/N cut of S/N $\geqslant 5 + 0.3 \times (N - 1)$. Under the assumption of a non-evolving LAE luminosity function, we predict a 10% ± 1.6% contamination of spurious sources to the LAE sample. We project there are 17 ± 3 spurious sources in the data catalog. A sample essentially free of contamination can be produced by using this catalog with an S/N > 6 cut, which by the limited number statistics of these simulations may contain 0^{+5}_{-0} spurious sources.

In addition, we have also performed an empirical test for spurious sources by analyzing the inverse of the survey data frames. All sources with a detected continuum were masked (so that we would not find the inverse of absorption features as spurious sources), and our detection algorithm was re-run. This analysis found seven spurious sources in 28 fields; a rate that is significantly lower than that estimated from the simulations. This suggests that our estimate of the systematic error is conservative and the true contamination fraction likely lies somewhere between 4% and 10%.

4.3. Completeness Tests

Not every source at the flux limit of Figure 2 will be recovered by the detection scheme. Beyond the usual statistical fluctuations introduced by noise, different source positions and seeing variations will cause the signal to be distributed over different numbers of fibers and cause varying fractions of light to be lost to the gaps between fibers. While this partial image sampling is an undesirable feature, IFS mitigates these uncertainties compared to serendipitous long-slit observations (Rauch et al. 2008; Lemaux et al. 2009; Cassata et al. 2010), where the slit losses can range (nearly uniformly) from 0% to 100%.

We have simulated our completeness limit using 25 mock fields of full, two-dimensional data with noise generated from the mean McDonald sky spectrum and the CCD read noise. Each simulated image contained 3000 emission-line sources randomly chosen in position and wavelength, but constrained to avoid object blending and spaced by the seeing from the IFU edges. We used the same detection routines as for the real data. For all these simulations, the seeing was held constant at the survey's 1.″5 FWHM median. These mock sources were modeled as spectrally unresolved point sources with fluxes randomly drawn from an unevolving Gronwall et al. (2007) LAE luminosity function over the luminosity range $41.5 < \log L(\mathrm{erg\ s^{-1}}) < 44.5$ where the lower bound was chosen to yield S/N = 0.5 over most of the wavelength range. Figure 9 compares our simulated emission-line fluxes to the fluxes that were measured. As the S/N decreases, the error in our measurements increases. Moreover, at the faintest limits, there is a slight systematic trend, with the measured fluxes being overestimated. This is the well-known Eddington (1913, 1940) correction which, if ignored, can lead to an underestimate of a luminosity function's slope. The least-squares fit shown in the figure will be used to statistically correct all our LAE fluxes prior to luminosity function computation. The completeness results are shown in Figure 10. We reach 50% and 95% corrected completeness at 5.6σ and 8.3σ, respectively. Compared to a

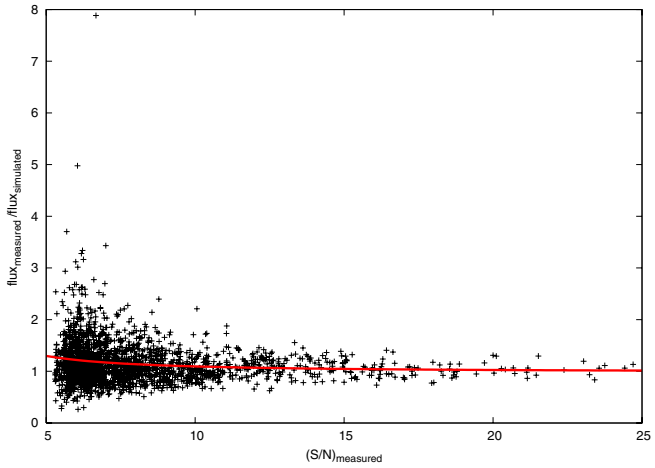

Figure 9. Ratio of measured and input fluxes from point-source simulation and a least-squares fit (Section 4.3). The curve-of-growth method is used here to measure emission-line flux. The upturn at low S/N is expected from the Eddington bias. This ratio can be applied to the LAE flux measurements of Table 3 for some applications. The average correction is fit as $\mathrm{flux}_{\mathrm{measured}}/\mathrm{flux}_{\mathrm{true}} = 0.98 + 0.74/(\mathrm{S/N})_{\mathrm{measured}} + 4.27/(\mathrm{S/N})^2_{\mathrm{measured}}$.

(A color version of this figure is available in the online journal.)

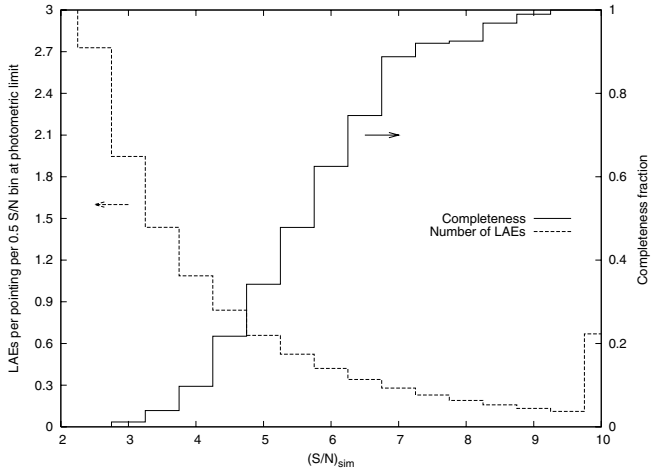

Figure 10. Survey's catalog completeness function for multiples of the flux limit. When combined, this function, the survey's photometric wavelength-based flux limit in Figure 2, and the actual mosaic pattern with photometric calibration in Figure 6 define the limits necessary for volume and luminosity function calculation (Blanc et al. 2010). The number of predicted LAEs per S/N bin assuming photometric conditions and a non-evolving Gronwall et al. (2007) LAE luminosity function is also plotted for reference. The long tail of the completeness curve down past the luminosity function break shows that the incomplete bins are crucial toward maximizing the catalog size.

step function completeness limit at S/N > 5 at the photometric limit of this survey which we consider the ideal goal, the number of detected LAEs is degraded by 13%. The long, low S/N tail helps mitigate the loss of objects to the non-ideal completeness.

Our source simulations also allow us to quantify the statistical astrometric error as a function of S/N. This is an important ingredient to our algorithm for associating VIRUS-P emission-line objects with sources found in broadband imaging (see Section 5). If we adopt a Rayleigh distribution for the form of the radial errors, i.e., $\sigma = a + b/(\mathrm{S/N})$, then a maximum likelihood fit for the coefficients yields $a = 0\farcs348$ and $b = 2\farcs04$. Figure 11 shows this relation with the individual measurements overplotted.

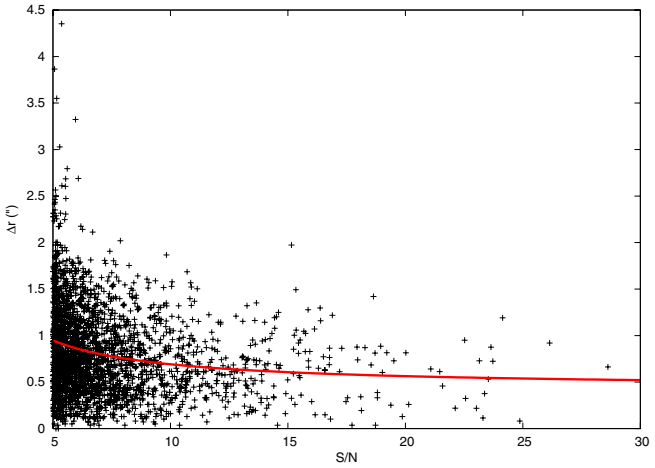

Figure 11. Distribution of simulated source positional errors under VIRUS-P sampling patterns with a fit to the dispersion as a function of measured S/N. The maximum likelihood fit to the peak of a Rayleigh distribution gives $\sigma = 0\farcs348 + 2\farcs04/(\mathrm{S/N})$.

(A color version of this figure is available in the online journal.)

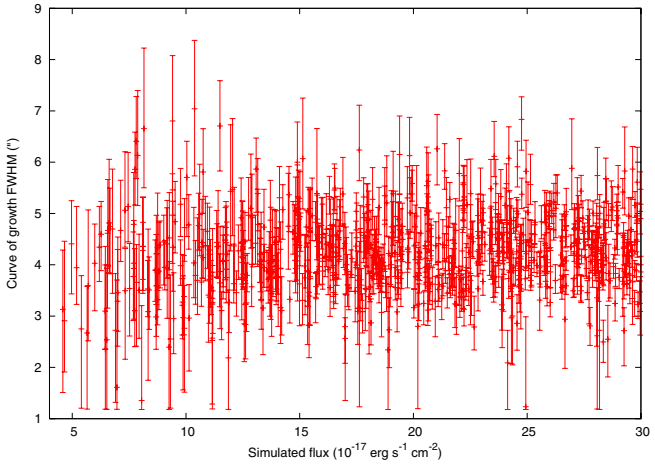

Figure 12. Distribution of measured source sizes in simulated data. The simulation input source sizes were drawn from the survey seeing distribution. The large fiber sizes set a large resolution threshold. Based on the 99.7% confidence interval, we only claim significantly resolved measurements for curve-of-growth FWHM sizes $>6\farcs81$.

(A color version of this figure is available in the online journal.)

The large VIRUS-P fibers lead to poor spatial resolution. Nevertheless, we have also simulated one mock field of 3000 point sources at and above the survey's flux limit and seeing distribution in an effort to quantify the minimum resolvable source size. To do this, we modeled the seeing FWHM distribution as a Gaussian function centered on $1\farcs5$ with a dispersion of $1''$ but truncated below $1\farcs2$. With the oversampled pattern of dithers, we expect the Nyquist limit to be near the diameter size of a fiber. The same curve-of-growth (CoG) photometry routines as described in Section 4.4 were used to measure the sizes of simulated point sources. Figure 12 shows the distribution of emission-line flux and measured size. The distribution is mostly flat with either flux or source S/N. Based on the simulation, we label a threshold of $7\farcs5$ as the resolution limit of our survey. This can be compared to the usual definition for Lyα blobs, i.e., emission over an isophotal area of $>16\square''$ at a certain surface brightness threshold. The Lyα blob surveys of Matsuda et al. (2004) and Yang et al. (2010) used thresholds of

2.2×10^{-18} erg s^{-1} cm^{-2} arcsec^{-2} and 5×10^{-18} erg s^{-1} cm^{-2} arcsec^{-2}, respectively. Our HPS should detect many Lyα blobs based on this flux limit, but will only be able to resolve the very largest objects. The full HETDEX survey will have $\sim$3$\times$ better spatial resolution.

4.4. Line Flux Measurement

A source's detection aperture described in Section 4.1 does not contain the total source flux. The imposed S/N cut omits some fraction of the flux in the detection aperture; this fraction is a function of source strength and orientation to the fiber dither pattern. In order to determine an unbiased emission-line flux in the presence of these complications, we describe here a CoG procedure used to measure a source's total line flux after detection. While other total flux estimators are possible, we advocate this method as generally robust against the range of sizes and morphologies encountered in the survey and the rather large astrometric errors and seeing variations inherent in this data set. The algorithm is similar to CoG (Stetson 1990) fits previously developed for CCD imaging photometry, but is new to spectrophotometry.

We begin a flux measurement by considering the positions, central wavelengths, and line widths ($\sigma_{\rm det}$) obtained from the emission-line detection algorithm described in Section 4.1. A circular aperture is formed around the centroid emission-line position of variable radius. Fibers overlapping this aperture are given fractional weights determined by their enclosed areas. Specifically, we form 15 apertures linearly spaced between radii 2.″2 and 9.″0. In each aperture, the enclosed fibers have their continuum-subtracted data summed and errors summed in quadrature for wavelengths within $\pm 2\sigma_{\rm res}$ of the detection wavelength. A spectral correction factor is defined as the flux fraction of a Gaussian line profile that falls within the fixed, spectral window defined by Equation (2)

$$f_{\rm spec,corr} = {\rm erf}\left(\sqrt{2}\sigma_{\rm res} \Big/ \sqrt{\sigma_{\rm res}^2 + \sigma_{\rm det}^2}\right). \quad (2)$$

Note that the fluxes returned by directly summing all fibers in a circular aperture of radius r, $\hat{f}(r)_{\rm raw}$, may oversample or undersample the source flux depending on the data completeness and overlap regions of mosaic. For example, the ideal six dither pattern produces an oversampling of very near two. Let the number of fibers at a particular position lying within one fiber radius, $r_{\rm fib}$, be $N(\Delta r < r_{\rm fib}, r, \theta)$ in polar coordinates. Equation (3) gives the raw flux measured for arbitrary sampling of a source with total flux $f_{\rm total}$ and normalized profile $P(r, \theta)$; $f(r)_{\rm samp} \equiv \int_0^r f_{\rm total} P(r, \theta) r dr d\theta$ is an estimate of the cumulative flux corrected for sampling. This approximation is correct when $N(\Delta r < r_{\rm fib}, r, \theta)$ does not systematically depend on r, which is nominally true for the randomly positioned observations presented here. The approximation is necessary to cleanly estimate an unbiased flux without knowing the exact profile.

$$\hat{f}(r)_{\rm raw} = \int_0^r f_{\rm total} N(\Delta r < r_{\rm fib}, r, \theta) \times P(r, \theta) r dr d\theta$$
$$\approx \hat{f}(r)_{\rm samp} \times \frac{\int_0^r N(\Delta r < r_{\rm fib}, r, \theta) r dr d\theta}{\pi r^2}. \quad (3)$$

We fit, by nonlinear least squares minimization, a cumulative two-dimensional Gaussian function, $A_{\rm CoG} \times (1 - e^{-0.5 r^2/\sigma_{\rm CoG}^2})$, to the highly correlated distribution $\hat{f}(r)_{\rm samp}$, where we enforce the limits $1'' < \sigma_{\rm CoG} < 10''$. In addition, we create Monte Carlo realizations by varying each fiber's intensity from the best-fit model. The CoG data points are highly correlated, so we took care to estimate the errors from the uncorrelated data of each fiber. The final, total flux estimate is given by Equation (4), with errors similarly propagated from the raw data and the uncertainty in $\sigma_{\rm CoG}$. Figure 13 gives CoG examples for an [O II] emitter and an LAE

$$\hat{f}_{\rm total} = A_{\rm CoG} / f_{\rm spec,corr}. \quad (4)$$

We tested the reliability of the CoG flux measurement, particularly for correlated errors with the source size, by using the simulated data discussed in Section 4.3. We first measured the flux from the fibers chosen as the detection aperture (Section 4.1) and compared this to the simulated flux. The mean and dispersion of the measured-to-simulated ratio are 0.93 and 0.31; unsurprisingly, the fluxes are systematically underestimated. Next, the set of all fibers within 6″ of the detected position was used as the flux aperture. This reduced the scatter found by the fixed aperture method, but a systematic error still remained with a mean of 0.94 and dispersion of 0.20. Finally, the CoG flux measurement was considered. Under this procedure, the bulk systematic flux measurement error vanished, giving a mean of 1.00 while still maintaining a low dispersion of 0.23. All three flux estimation methods are shown in Figure 14 against the simulated source size. A systematic offset with input source size can be seen for all cases, but the CoG photometry is preferred as the least biased method investigated.

5. SOURCE CLASSIFICATION

An emission-line galaxy catalog is of limited value without secure redshift identifications. Unfortunately, the uncertainty in identifying single emission lines is a common hindrance to high-redshift galaxy surveys (e.g., Stern et al. 2000). We here describe the two steps necessary to robustly assign redshifts to the emission-line catalog. Tables 3 and 4 present the catalogs. We give the detailed description of these tables in Section 6.2. We further summarize the statistics of commonly found objects and compare the sample to other works where available.

5.1. Spectral Classification

As mentioned in Section 2.5, the presence of multiple, strong emission lines can be used to identify some low-z objects, but the absence of such lines is not sufficient evidence to classify a source as an LAE. We begin all source classifications by cross-correlating the primary emission line at various assumed redshifts to other bright, expected emission lines. We automatically search all the detection spectra for Mg II 2798, [O II]3727, Hγ4341, Hβ4861, [O III]4959, and [O III]5007 assuming the detected line to be, variously, [O II]3727, Hβ4861, [O III]4959, and [O III]5007. At high redshift we test Lyα for the presence of C IV1549. We have manually tried using the other, commonly weaker lines as confirmation of the primary detection, but have found only two cases of interest. For emission-line index 4 of Tables 3 and 4, the C III]1909 line is detected with the also-significantly detected [O II]3727 line of index 5. For emission-line index 85 of the same tables, the broad Mg II 2798 line is brighter than the also-significantly detected [O II]3727 line. We have also misidentified index 400 as an [O II] emitter in the first pass of analysis; it is known to be an [O III]5007 emitter from the literature (Barger et al. 2008), but we find no other detections at other wavelengths.

A demonstration of this cross-correlation process for a multiple-emission-line source is shown in Figure 15. We find

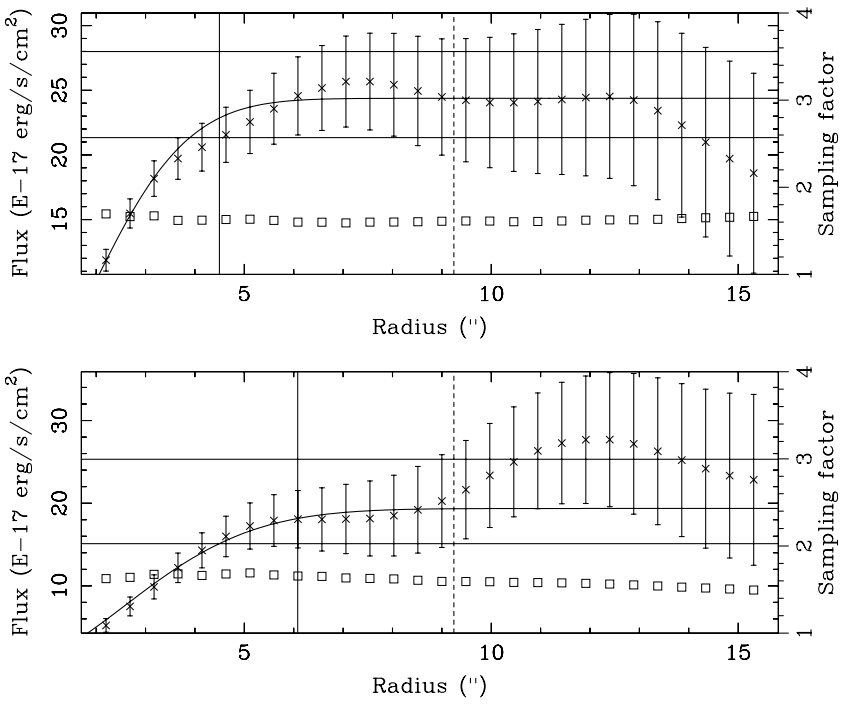

Figure 13. Data, fits, and final evaluation for the curve-of-growth line flux measurements of an LAE (top) and an [O II] emitter (bottom). In both, the open square symbols display the cumulative sampling factor on the right-hand scale. The sampling factor is written in Equation (3) as $N(\Delta r < r_{\rm fib}, r, \theta)$ and is the average number of fibers overlaying the surface enclosed in radius r. The points with errors show the estimated cumulative flux on the left-hand scale. The vertical dotted lines mark where the fit is truncated. This threshold has been selected with consideration toward being significantly larger than the widest objects found and small enough to limit unnecessary noise. The vertical solid line marks the radius to the fit where 90% of the flux is enclosed. The horizontal solid lines show the 1σ confidence of the fit's normalization. The top fit returns a total flux of $24^{+3.6}_{-3.0} \times 10^{-17}$ erg s^{-1} cm^{-2}. The bottom fit returns a total flux of $19^{+6.0}_{-4.3} \times 10^{-17}$ erg s^{-1} cm^{-2}. The errors are correlated on the displayed scale, but the Monte Carlo fit varies the data from each fiber independently to generate proper errors.

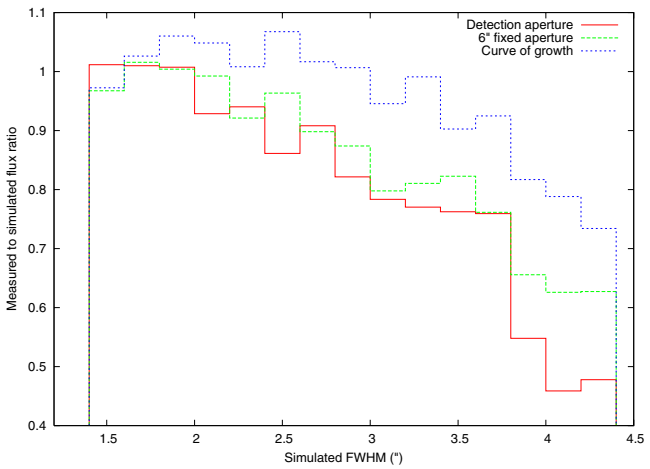

Figure 14. Binned ratios of measured and input fluxes for simulated data under a range of source sizes. The curve-of-growth flux estimator is preferred as the least biased for extended sources compared to either the detection set of fibers or a fixed radius set of fibers as the photometry aperture.

(A color version of this figure is available in the online journal.)

only two cases where the correlation against data below the catalog signal-to-noise cut aids classification as shown in Figure 16. The first is emission-line index 234, which is formally a single emission-line detection. However, we find that an identification of the primary line with [O III]5007 leads to an S/N = 3.2 detection at the wavelength of [O II], an S/N = 5.1 detection at Hβ, and an S/N = 3.6 detection at [O III]4959. The second case is emission-line index 430 which is also a single emission-line detection. We again find that an identification of the primary line with [O III]5007 leads to an S/N = 3.7 detection at the wavelength of [O II] and an S/N = 2.9 detection at Hβ. In practice, the primary utility in the emission line cross-correlation is to discriminate between various low-z possibilities with high S/N detections.

5.2. EW-based Classification

In any LAE survey at sufficient redshift, the most likely contaminants are [O III]5007 and [O II]3727. Many of the former objects can be identified by the presence of [O III]4959 or Hβ. The latter may be identified by either splitting the [O II] doublet, or by using line EW as a discriminant (Cowie & Hu 1998). Since we lack the resolution to split [O II]3727, we follow Gronwall et al. (2007) and require LAEs to have EW$_{\rm rest}$ > 20 Å. A number of different EW estimators are possible with measurements in many filters. We look at two ways to estimate the EW using broadband data and conclude that the cleanest selection of LAEs is obtained when the R-band data are used alone.

The observed wavelengths and EWs are shown in Figure 17. Emission lines without counterparts are shown as limits. We calculate the EW first by using the nearest-available filter that lies redward of the entire sample. For XMM-LSS, GOODS-N, and COSMOS, this is the R band. MUNICS lacks an R-band image, so we used i'. The redward choice is important to avoid attenuation by the intergalactic medium for these data and the Lyman break. Although there may be some diversity in LAE dust content (Finkelstein et al. 2009), it appears that most LAEs at our redshifts of interest have only small amounts of dust and exhibit flat continua (Gawiser et al. 2007; Guaita et al. 2010; Blanc et al. 2010). Of course, low-redshift, star-forming galaxies may also exhibit flat continua or, more likely, some level of a Balmer/4000 Å break, but by extrapolating the continua

Table 3
HETDEX Pilot Survey Emission-line Catalog (Abridged)

HPS Index	α (J2000)	δ (J2000)	λ_{det} (Å)	FWHM (km s^{-1})	S/N$_{det}$	Flux (10^{-17} cgs)	Spatial FWHM ($''$)	Matching Indices
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
001	02:21:11.16	−04:31:25.0	5219.16	229	8.1	$17.4^{3.5}_{-3.9}$	$4.7^{0.8}_{-0.6}$	..
002	02:21:12.21	−04:32:25.3	5448.72	307	5.6	$12.2^{3.6}_{-4.3}$	$4.3^{0.8}_{-1.0}$	..
003	02:21:14.28	−04:31:38.2	4973.93	422	7.5	$19.9^{4.7}_{-3.1}$	$4.4^{0.8}_{-0.5}$	..
004	02:21:14.86	−04:31:56.6	5261.37	1285	6.3	$42.6^{11.2}_{-12.4}$	$5.1^{0.7}_{-0.9}$	5
005	02:21:15.14	−04:31:54.0	4270.67	1841	33.1	$342.1^{16.5}_{-14.3}$	$4.8^{0.2}_{-0.1}$	4
006	02:21:16.26	−04:29:32.8	4591.58	399	14.8	$32.7^{3.5}_{-3.6}$	$4.6^{0.4}_{-0.2}$	..
007	02:21:16.35	−04:31:14.6	5161.72	293	19.5	$49.4^{2.6}_{-4.4}$	$4.7^{0.2}_{-0.2}$	..
008	02:21:17.25	−04:27:55.7	5820.13	118	6.7	$19.1^{5.0}_{-3.2}$	$6.6^{0.9}_{-0.6}$	..
009	02:21:17.25	−04:30:10.4	5464.33	78	12.1	$14.1^{1.5}_{-2.5}$	$3.6^{0.5}_{-0.4}$	..
010	02:21:17.47	−04:27:30.6	4808.33	357	15.9	$38.9^{3.4}_{-3.6}$	$4.6^{0.2}_{-0.3}$	..

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

Table 4
HETDEX Pilot Survey Emission-line Classifications (Abridged)

HPS Index	Counter-part	Counter-part m_R^a	Counter-part P	EW$^a_{R,rest}$ (Å)	EW$_{interp,rest}$ (Å)	Transition	z_{est}	Lyα P	X-ray Counterpart
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
001	J0221112-043126	23.05	0.93	$51.9^{14.7}_{-14.6}$	$62.9^{17.0}_{-17.4}$	[O II]	0.4004	0.07	..
002	J0221122-043225	23.17	0.96	$42.0^{14.9}_{-16.3}$	$59.8^{20.5}_{-23.0}$	[O II]	0.4620	0.04	..
003	J0221143-043138	24.31	0.98	$58.8^{22.2}_{-15.1}$	$109.0^{36.8}_{-26.5}$	Lyα	3.0915	1.00	..
004	J0221150-043156	21.05	0.98	$10.4^{3.4}_{-3.5}$	$7.1^{2.3}_{-2.4}$	C III]1909	1.7561	0.02	J0221151-043156
005	J0221150-043156	21.05	1.00	$54.9^{10.9}_{-9.7}$	$55.1^{9.9}_{-9.2}$	C IV1549	1.7570	0.00	J0221151-043156
006	J0221164-042933	23.82	0.89	$56.7^{12.6}_{-11.5}$	$74.1^{15.8}_{-14.8}$	Lyα	2.7770	1.00	..
007	J0221163-043116	21.38	0.98	$31.2^{6.2}_{-6.0}$	$48.3^{8.8}_{-8.9}$	[O II]	0.3850	0.02	..
008	J0221171-042757	22.82	0.67	$51.2^{16.8}_{-12.1}$	$57.4^{18.3}_{-13.3}$	[O II]	0.5616	0.01	..
009	J0221174-043001	23.21	0.98	$51.0^{11.2}_{-12.4}$	$49.0^{9.9}_{-11.6}$	[O II]	0.4661	0.02	..
010	J0221174-042729	21.43	0.99	$24.0^{5.1}_{-4.6}$	$39.1^{8.2}_{-7.5}$	[O II]	0.2901	0.01	..

Notes.
[a] The Johnson or SDSS R-band filters used are listed in Table 2. The MUNICS field, at $\alpha \approx 3$ hr, instead uses an SDSS i filter.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

from the R band, the low-redshift EWs will be somewhat underestimated while the LAE EWs should remain unaffected. Still, while such a property is beneficial to the classification process, an unbiased EW is also desirable for physical studies. So, we next calculate the EWs in the right panel of Figure 17 by interpolating each emission line with the two nearest, bounding broadband filters. Clearly, the high and low EW populations have more overlap in the interpolated EW measurements. For this reason, we adopt the R-band EW in our classification scheme. Figure 18 shows the emission-line flux against continuum magnitude for each emission line. We have also checked the *GALEX* (Martin et al. 2005) GR4/GR5 database for all objects. None of the LAE classified objects are *GALEX* sources, while most of the low-z classified objects do have counterpart *GALEX* detections.

There are nine objects for which we make exceptions: four low EW objects we identify as LAEs and five high EW sources we believe are low-redshift interlopers.

1. The lowest wavelength exception is observed at $\lambda = 3765.6$ Å with EW$_{obs} = 41^{+21}_{-17}$ Å as index 313. If this were [O II], the galaxy would be extremely nearby (45 Mpc) away and have $M_R = -10.5$. The photometric redshift of Ilbert et al. (2009) suggests the line to be Lyα and excludes all the low-z options with 95% confidence.
2. The next low EW object is in the MUNICS field as index 51 and has $m_{i'} = 23.7$. The detected wavelength is 4981.6 Å with EW$_{obs} = 61^{+38}_{-29}$ Å. The case for this object is not terribly strong, but the dim continuum and lack of a *GALEX* detection suggest this to be an LAE.
3. The third low EW object is in the GOODS-N field as index 447 at wavelength 5017.2 Å with EW$_{obs} = 81^{+31}_{-18}$ Å. It was originally listed as an Lyman Break Galaxy (LBG) in Steidel et al. (2003), but no redshift measurement exists in the literature. The counterpart has $m_R = 24.2$.
4. The final low EW object is in the MUNICS field as index 92 at wavelength 5683.3 Å with EW$_{obs} = 84^{+34}_{-31}$ Å. The counterpart has $m_{i'} = 23.3$, but no *GALEX* detection. Again, this is a borderline classification.

Next, we consider the five high EW objects reclassified as being at low redshifts.

5. The first high EW low-z object is in COSMOS as index 289 at wavelength 5235.9 Å with EW$_{obs} = 96^{+28}_{-26}$ Å.

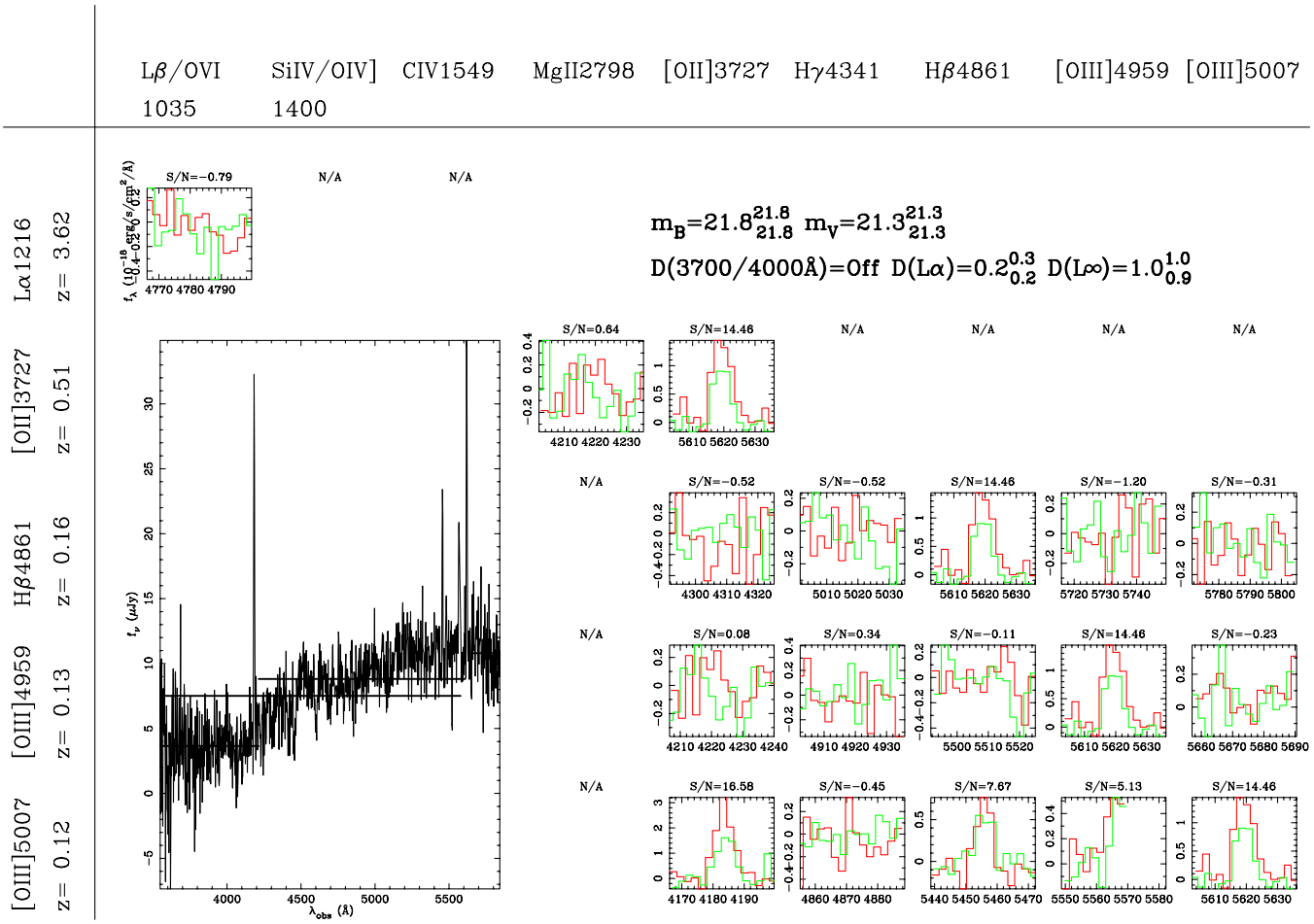

Figure 15. Attempts to find matching lines to an [O III]5007 detection in index 323 of Table 3. The detection is formed from two fibers represented independently with the red and green lines. The leftmost column shows the prospective identifications of our originally detected line at 5619 Å. For each prospective identification, we attempt fits to the emission-line possibilities in the top row. In this case, matches to [O II] and Hβ both give a clear identification. The [O III]4959 is detected but overlapping with the mask around 5577 Å. This technique only rarely gives positive evidence for Lyα classification by ruling out low-z emission-line combinations because our wavelength bandpass is not much larger than common bright optical emission-line spacings. However, matching emission lines often aid in classifying transitions between different low-redshift options as in this case. The full spectrum is shown in the large window to the left. Various continuum regions are evaluated by assuming the primary emission line to be Lyα and [O II] shown by the horizontal lines. The continuum fits are used to look for various breaks as calculated in the upper right. This galaxy is also identified by significant detections of Hβ, [O III]4959, and [O II] in entries 325, 326, and 327 of Table 3.

(A color version of this figure is available in the online journal.)

It does not have a *GALEX* detection, but the counterpart has $m_R = 23.4$. The photometric redshift of Ilbert et al. (2009) suggests the line to be [O II] and excludes all other reasonable options with 95% confidence.

6. This COSMOS object is index 234 with $m_R = 23.4$, wavelength 5466.7 Å and $\mathrm{EW_{obs}} = 149^{+28}_{-23}$ Å. As discussed in Section 5, the source shows low significance emission lines such that the primary detection is likely [O III]5007. Such an identification is possible since unlike [O II], [O III]5007 can have extremely high EWs (Hu et al. 2009). The source also has a *GALEX* detection.
7. This GOODS-N object is index 356 at $m_R = 22.8$, wavelength 5700.5 Å and $\mathrm{EW_{obs}} = 104^{+32}_{-26}$ Å. It has a *GALEX* detection and has a measured redshift in Barger et al. (2008) as being from [O II] emission.
8. This is index 439 from the GOODS-N field with $m_R = 24.2$, wavelength 5762.4 Å, and $\mathrm{EW_{obs}} = 119^{+52}_{-39}$ Å. The object is detected with *GALEX*.
9. This is index 94 from the MUNICS field with $m_{i'} = 21.0$, wavelength 5768.4 Å, and $\mathrm{EW_{obs}} = 107^{+22}_{-20}$ Å. The object is detected with *GALEX*.

We next review the likely levels of contamination in the LAE sample from low-redshift objects based on previous studies. The frequency of EW in bright, rest-frame-optical lines at low redshift has been studied in Hammer et al. (1997), Hogg et al. (1998), Treyer et al. (1998), Sullivan et al. (2000), Gallego et al. (2002), and Teplitz et al. (2003). By combining the observation that ∼2% of [O II] emitters have $\mathrm{EW_{rest}} > 60$ Å (Hogg et al. 1998) with the $0 < z < 0.4$ [O II] luminosity function of Sullivan et al. (2000) and assuming no redshift evolution of either the [O II] or LAE EW distributions (Gronwall et al. 2007), we estimate that our sample may contain 1.6 high EW [O II] interlopers. Similarly, if we use the local luminosity function of Gallego et al. (2002), we predict zero high EW interloping [O II] emitters. As a second comparison, Kakazu et al. (2007) present narrowband imaging and limited spectroscopic follow-up in their search for low metallicity galaxies. They find [O III]5007 at $z = 0.63$ and $z = 0.83$, and [O II] at $z = 1.19$ and $z = 1.45$ at high enough EW values to contaminate our sample. By comparing their high EW [O II] number density to the Schechter (1976) function fits of Ly et al. (2007) at $z = 1.18$ and $z = 1.47$ without extinction corrections, we

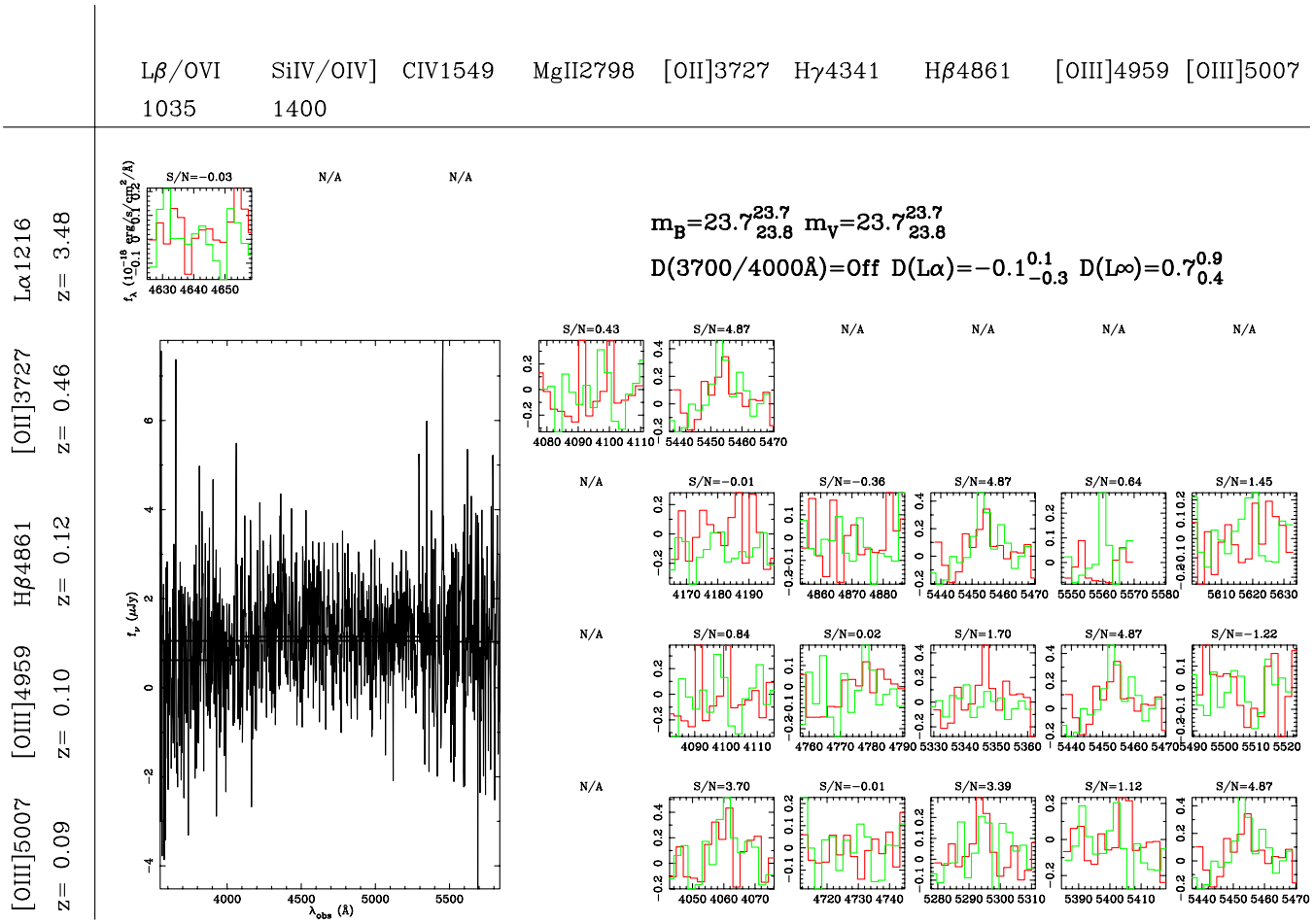

Figure 16. Automated search for multiple emission line in index 430. The format is the same as in Figure 15. This is one of the few cases where the consideration of marginally significant counterpart lines aids the classification. The primary detection is revealed to be [O III]5007, with marginal detections in [O II] and Hβ that did not make the primary emission-line catalog.

(A color version of this figure is available in the online journal.)

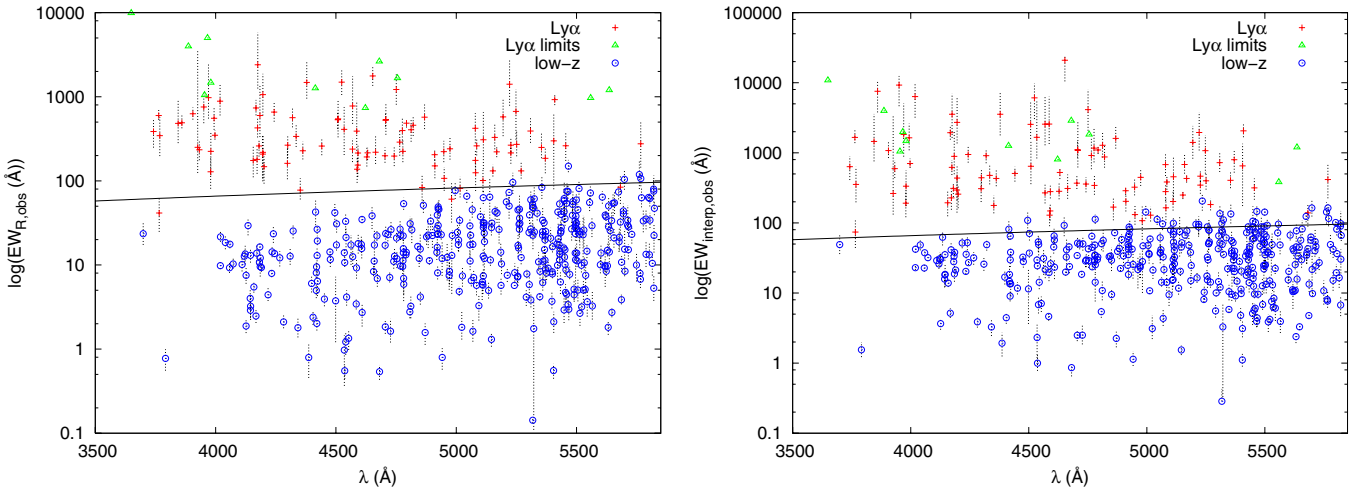

Figure 17. Distribution of observed wavelength and observed EW. The line marks the usual EW cut used in narrowband imaging and adopted here. Exceptions to the EW selection are discussed in Section 5.2. Left: continuum estimated only from the *R*-band photometry (or the *i′* band in MUNICS). Right: continuum estimated from interpolation with the two nearest filters bounding each emission-line's wavelength.

(A color version of this figure is available in the online journal.)

find that the high EW [O II] fraction should only be 3%. In contrast, the same analysis suggests that the high EW [O III]5007 fraction is much higher (33%). However, there is no evidence for such a large fraction of high EW [O III]5007 over our redshifts of interest, and the VIRUS-P bandpass will always enclose [O II] and [O III]4959 when [O III]5007 is observed. Thus, neither high EW [O II] nor [O III]5007 emitters should form important catalog contaminants. The wavelength spacing

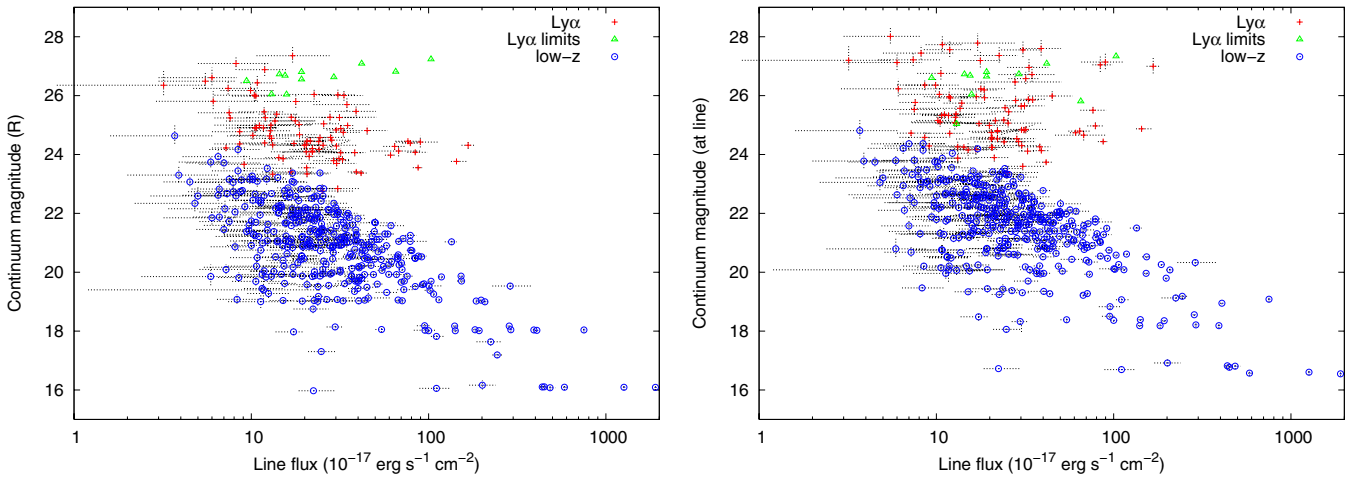

Figure 18. Distribution of emission-line flux and broadband magnitude. Left: continuum estimated only from the *R*-band photometry (or the i' band in MUNICS). Right: continuum estimated from interpolation with the two nearest filters bounding each emission-line's wavelength.
(A color version of this figure is available in the online journal.)

between AGN lines is smaller than for [O II] and the other optical lines, so AGN lines should be identifiable with multiple detections. For example, index 4 is our only C III]1909 detection, identified with a co-detection in index 5 as C IV1549. The best available contamination estimate is that this LAE sample contains 0–2 contaminants from misidentified redshifts if only EW information is used. A complementary question is how our rest-frame EW affects the selection of high-redshift galaxies. Assuming the LAE distribution in Gronwall et al. (2007), the answer is that ∼21%–26% of potential detections are lost by the EW cut.

We briefly state how we propagate errors to the EW estimation. In cases where the flux density measurement is very noisy, the usual first-order error propagation breaks down. Importantly, the error on EW becomes asymmetric in the case of a low S/N continuum even if the original errors on flux and flux density are symmetric. One simple solution is to treat the maximum likelihood distributions in flux and flux density as Gaussian functions, transform the flux into EW and flux density, and define the EW errors using the extrema of the 68% confidence interval. Similarly, in the case of asymmetric errors for the line fluxes, we use the same equation evaluated with each one-sided error to arrive at final EW limits. When we find no upper limit, we list the upper uncertainty as 1000 Å.

5.3. Counterpart Association

The coarse spatial resolution of our VIRUS-P survey often prevents us from associating with certainty a given emission line to a unique broadband counterpart. However, stringent redshift probabilities can often be made by marginalizing over all possible counterparts and their implied rest-frame emission-line EWs. We quantify this association probability by using the astrometric error, discussed in Section 2.3, and the differential number counts for the *R*-band images. Since MUNICS lacks *R*-band data, we use the i' band there. The exact band choice for this step is not critical, so long as the filter samples a fairly flat spectral region for both low-z and high-z objects. We describe the method for *R*-band continuum association as it applies to EW-based redshift discrimination. We use the same formalism for AGN association through X-ray data in Section 5.4.

The problem of assigning counterpart probabilities to detections in multiple bandpasses has been explored by Bayesian methods in Sutherland & Saunders (1992) and is commonly implemented in X-ray surveys (e.g., Luo et al. 2010). We choose not to use the Bayesian technique here since it requires assuming a prior on the continuum counterparts to the emission-line detections. We instead make a simpler, frequentist estimate that still uses the information from multiple candidates. The only assumed inputs are the astrometric error and the number counts of background and foreground objects.

The probability of an emission line being associated with any one image-based counterpart can be constructed as the joint probability of all the remaining imaging detections being unassociated and drawn from established number counts and the preferred counterpart having the observed offset evaluated against the astrometric error budget. For simplicity, we treat all the individual probabilities as independent; this simplification is justified since the range of distances in our redshift range is much larger than cross-correlation scales between galaxies. We begin by identifying all the significant imaging detections within some large area of the detected emission line. We then define: i as the set of all imaging detections in the survey field, Δr_i as the angular offset between the position of the emission line and the centroid of counterpart i, S_i as the flux density of counterpart i (or X-ray flux in a defined bandpass), σ_i as the astrometric error for the emission line under consideration and counterpart i, and $n(S)$ as the differential number count of galaxies in the observed bandpass. We begin by assembling the set of imaging detections with cardinality j as $C_j = (j \in i : \Delta r_j < 10'')$. The exact value of the angular limit is not important so long as it is several times the astrometric error. The chance of a superposition by one or more imaging detections without them being actual counterparts is then $P_{\rm nc,j} = 1 - f(0, \lambda)$ where $f(n, \lambda) = \frac{\lambda^n e^{-\lambda}}{n!}$ is the Poisson probability distribution and λ is the expectation value for the number of galaxies brighter than S_j within Δr_j, so $\lambda = \pi \Delta r_j^2 \int_{S_j}^{\infty} n(S) dS$. Alternatively, the detection j may be the true counterpart. If we model the astrometric error as a two-dimensional Gaussian distribution in the astrometry error and take its cumulative evaluation from infinity, the chance of measuring the true counterpart at Δr_j or further is $P_{c,j} = \exp(\frac{-\Delta r_j^2}{2\sigma^2})$. It may also be that we have not

Table 5
Emission Line/X-ray Counterpart Statistics

Field	Low-z Counterparts[a]	High-z Counterparts	Depth[b]
COSMOS	2/112	4/55	0.73
GOODS-N	27/94	2/25	0.14
XMM-LSS	1/24	0/8	∼27
MUNICS	4/62	0/17	∼20

Notes.
[a] Counterparts/total emission lines.
[b] Assuming a point source, 10^{-15} erg s^{-1} cm^{-2} in 2–10 keV, if available, or 2–8 keV. The depth for XMM-LSS varies over the observed regions, and the sensitivity map for the X-ray coverage in MUNICS is not published.

measured the true imaging counterpart, either due to imaging depth or the emission-line detection being spurious, in which case all imaging detections must be explained as superpositions.

We give in Equation (5) the full joint probabilities assembled from the individual probabilities just described, under the assumption that either one or none of the imaging detections is the true counterpart to the emission-line detection. A similar calculation is done to evaluate the significance of X-ray counterparts in Section 5.4:

$$P = \begin{cases} \dfrac{\prod_{1\leqslant k\leqslant j} P_{\mathrm{nc,k}}}{\prod_{1\leqslant k\leqslant j} P_{\mathrm{nc,k}} + \sum_{1\leqslant\ k\leqslant j}\left(P_{c,k}\times\prod_{1\leqslant m\leqslant j, m\neq k} P_{\mathrm{nc,m}}\right)} & : \text{no counterpart} \\ \dfrac{P_{c,k}\times\prod_{1\leqslant m\leqslant j, m\neq k} P_{\mathrm{nc,m}}}{\prod_{1\leqslant k\leqslant j} P_{\mathrm{nc,k}} + \sum_{1\leqslant\ k\leqslant j}\left(P_{c,k}\times\prod_{1\leqslant m\leqslant j, m\neq k} P_{\mathrm{nc,m}}\right)} & : \text{counterpart } k. \end{cases} \quad (5)$$

In the case of imaging, the astrometric error is dominated by the positional uncertainty of the emission lines, but in the case of X-ray data the positional uncertainty of both the emission-line and X-ray detections are comparable and important. The normalization is simply chosen to make the probabilities sum to unity.

In order to match R-band objects, we performed a least-squares minimization fit to the R-band differential number counts of Furusawa et al. (2008) with a double power-law function to estimate $n(S)$ as given in Equation (6):

$$n(f_\nu)[\text{per }\square^\circ] = \begin{cases} 5142\times(f_\nu/10^{-28})^{-1.996} & : f_\nu > 7.81 \\ & \times 10^{-30}\,\mathrm{erg^{-1}\ s^{-1}\ cm^{-2}\ Hz^{-1}} \\ 10882.6\times(f_\nu/10^{-28})^{-1.702} & : f_\nu \leqslant 7.81 \\ & \times 10^{-30}\,\mathrm{erg^{-1}\ s^{-1}\ cm^{-2}\ Hz^{-1}}. \end{cases} \quad (6)$$

In practice, we consider a threshold distance of $1''$ when calculating the expected number counts of sources based on common seeing conditions to avoid the claim of total counterpart certainty at $\Delta r_j = 0$ regardless of other counterpart options, although the exact threshold makes little difference. Finally, in defining Δr_j, we take the radial offset from the emission-line centroid to the nearest position contained in the imaging's Kron aperture instead of the Kron aperture center. This is motivated by the fact that [O II] emission in nearby galaxies may be from H II regions located at large galactocentric radii.

Object classification starts by identifying all imaging catalog counterparts within $10''$ from the emission-line centroid. The association probability for each possible counterpart is calculated using Equation (5). The emission line is classified under

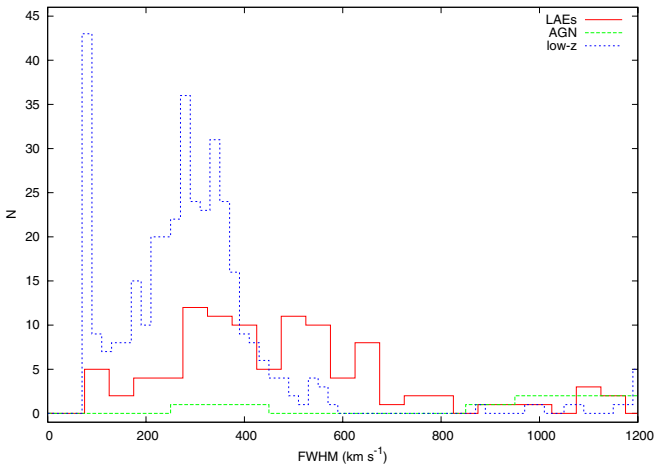

Figure 19. Line width distributions for all survey objects. We have subtracted an instrumental resolution of 300 km s^{-1} in quadrature. There is significant overlap between all populations making width-based classification impossible. No attempts have been made here to deconvolve the blended [O II] doublet. The low-z objects are generally contained to low widths, but the LAE distribution overlaps heavily with both the AGN and low-z distributions.
(A color version of this figure is available in the online journal.)

the EW rule of Section 5.2 to be at either low (most likely as [O II]) or high (most likely as Lyα) redshift. In 74% of the cases, the best counterpart probability exceeds 90%; we refer to these objects as the isolated sample. In another 3% of the cases, our analysis is most consistent with there being no broadband counterpart; we classify these sources as LAEs since all the image depths imply $\mathrm{EW_{Ly\alpha,rest}} > 20$ Å. For the remaining cases, our classification is less certain due to there being multiple likely counterparts; nevertheless, the most probable association is always presented. We illustrate step-by-step two representative classification cases in Sections 5.5.1 and 5.5.2, and a case with less certainty in Section 5.5.3.

We confirm the proper classification of many low-redshift objects by observing multiple emission lines. In total, there are 118 emission-line sources with one or more associated emission lines in combinations of [O II], Hδ, Hγ, Hβ, [O III]4959, and [O III]5007. Of these, all are classified automatically by our EW cut as being at low redshift.

5.4. AGN Contamination

We attempt to identify the Lyα sources that are AGNs through existing X-ray data. All our survey fields have either *Chandra* or *XMM/Newton* coverage, although to quite non-uniform depths. We use the point-source catalogs of Elvis et al. (2009) and Cappelluti et al. (2009) in COSMOS, Alexander et al. (2003) in GOODS-N, Pierre et al. (2004) in XMM-LSS, and Watson et al. (2009) in MUNICS. The data covering MUNICS is described in Severgnini et al. (2005) but not cataloged. The same methodology for determining broadband imaging counterparts is applied to the X-ray data. The cataloged X-ray spatial uncertainty is added in quadrature to the emission-line spatial uncertainty, and the fit of Cappelluti et al. (2007) from 2 to 10 keV is used for the differential number count. Unlike with the imaging counterparts, the association of an X-ray source with a VIRUS-P emission line is nearly binary in nature: there is either a single likely counterpart or no likely counterpart. Table 5 summarizes our results by listing the fraction of X-ray sources in the low-z and high-z objects. We find 6%–8% contamination of LAEs by AGNs over all the fields with the

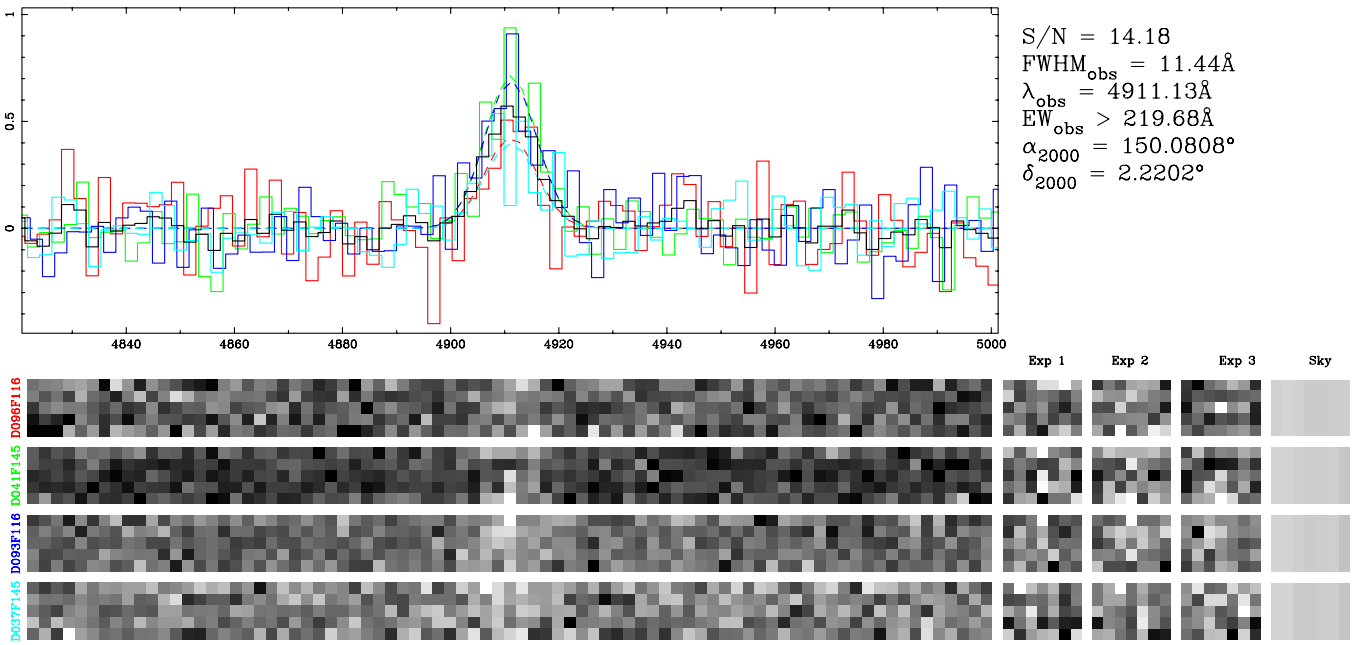

Figure 20. VIRUS-P emission-line detection image for index 229 which lies in the COSMOS field. This object is undoubtedly a real detection with one of the largest S/N ratios we find for any high-z object. In this case, the aperture is grown to include four fibers. The four rows in the figure's bottom half show the spectra from the four detection fibers. The right side, square cut-outs show the spectra from individual 20 minute exposures and the sky model. The three exposures in each fiber are then biweight combined into the two-dimensional, bottom left spectra and the one-dimensional spectra in the upper left line plot. The collapsed, one-dimensional spectra are color-coded by fiber number. The Gaussian fits to each fiber are given by dotted curves. For visual clarity, the spectra are resampled and stacked into the black histogram. Continuum is not detected within a 200 Å boxcar around the line, and the high level of flux permits the Lyα classification from the spectrum alone in this rare case. The tabulated EW instead is based on the flux density of the imaging counterpart. The quoted central wavelength in this figure has not yet had the heliocentric and vacuum corrections applied as is done with the tabulated values.

(A color version of this figure is available in the online journal.)

range depending on what fraction of the line detections we attribute to noise. AGN contamination is likely a strong function of flux limit, but we compare briefly to other, deeper surveys. The sample of Gawiser et al. (2007) and Gronwall et al. (2007) at $z = 3.1$ contains 1% AGN contamination, the sample of Nilsson et al. (2009, 2010) at $z = 2.3$ contains 6%–15% AGN contamination, and the sample of Guaita et al. (2010) at $z = 2.1$ contains 5% AGN contamination. These numbers, all utilizing X-ray detections of AGN, are consistent with the value we find. However, other work with mid-IR and far-IR AGN identification has potentially shown a much higher AGN fraction of 75% at $z = 2.2$ (Bongiovanni et al. 2010). We do not perform any mid-IR or far-IR AGN analysis here. There is no significant variation between AGN fractions of GOODS-N with the deeper X-ray data and the COSMOS field. The small number statistics and shallower X-ray data in MUNICS and XMM-LSS explain the lack of AGN detections in those fields.

There are two other potential indicators of AGN activity: broad Lyα emission-line widths (which may also be seen in Lyα "blobs" where AGN activity is not evident; Francis et al. 2001; Bower et al. 2004; Matsuda et al. 2006; Smith & Jarvis 2007; Saito et al. 2008), and the presence of C IV1549 over a fraction of the redshift range. The distributions of line widths, without any deblending of the [O II] doublet, is given in Figure 19. From the distributions, it is clear that line width information does not aid object classification. We find two cases where broad line objects (FWHM line widths >500 km s^{-1}) have been classified as Lyα without any X-ray detections, but none at >1000 km s^{-1}. However, only one Lyα and C IV1549 source (indices 461 and 462) was not detected in X-ray.

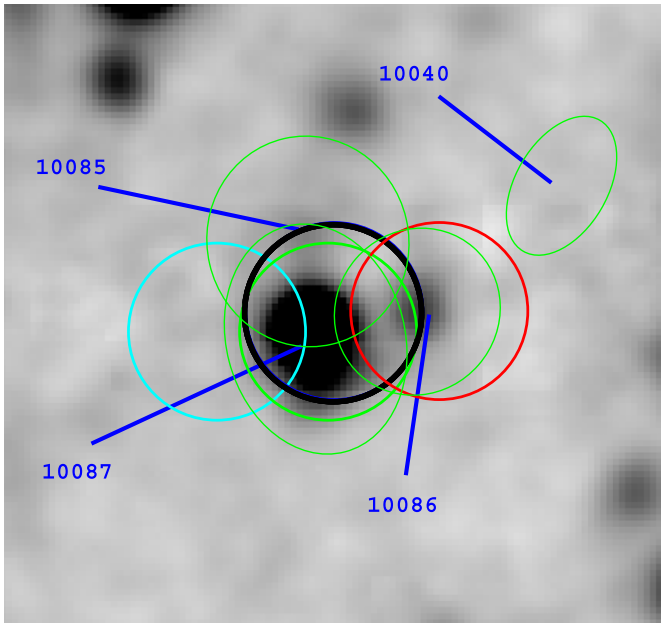

Figure 21. Ground-based, V_j imaging cutout for the Table 3 index 229. The four color circles represent fiber positions and are color coded in accordance with the spectra of Figure 20. The black circle indicates the emission-line centroid. The Kron apertures from the imaging catalog are drawn as green, numbered ellipses. The best centered and brightest source is 10087 with an association likelihood of 87% and $r^+ = 23.4$. The next two most likely counterparts are 10085 at 6% and $r^+ = 25.1$ and 10087 at 5% and $r^+ = 24.9$. Either of these would also make the LAE EW cut, and it is possible that one or more LAEs at similar redshifts are jointly contributing to the emission-line flux, but this is unlikely.

(A color version of this figure is available in the online journal.)

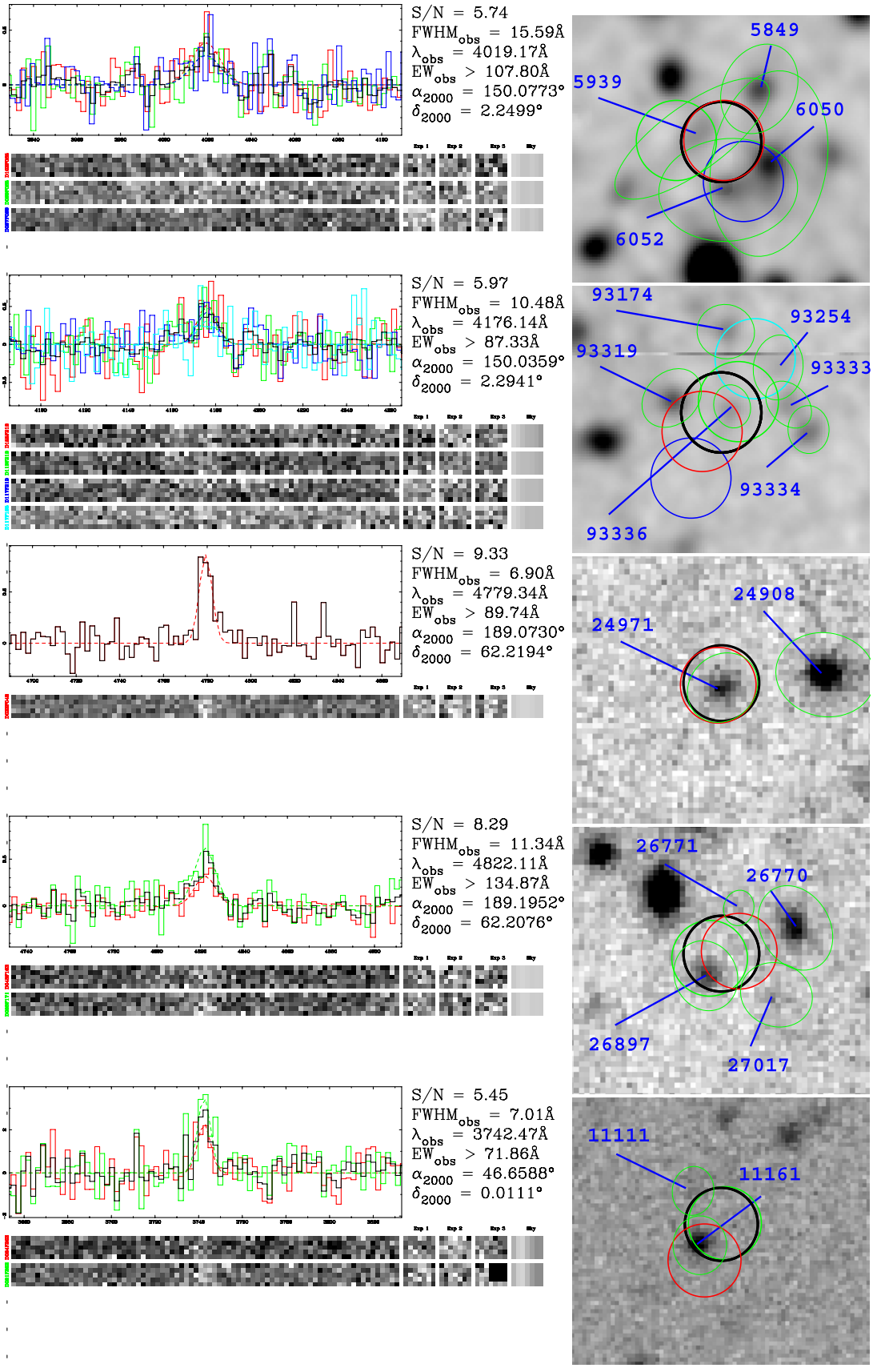

Figure 22. Five example detections of LAEs with the same formats as Figures 20 and 21. The first two lie in the COSMOS field, the next two lie in the GOODS-N field with the first redshift previously measured and the second new, and the final one lies in the MUNICS field. The entries from Table 3 for these five are indices 223, 160, 341, 402, and 62. The best continuum counterpart matches are to 5939, 93336, 24971, 26897, and 11161.

(A color version of this figure is available in the online journal.)

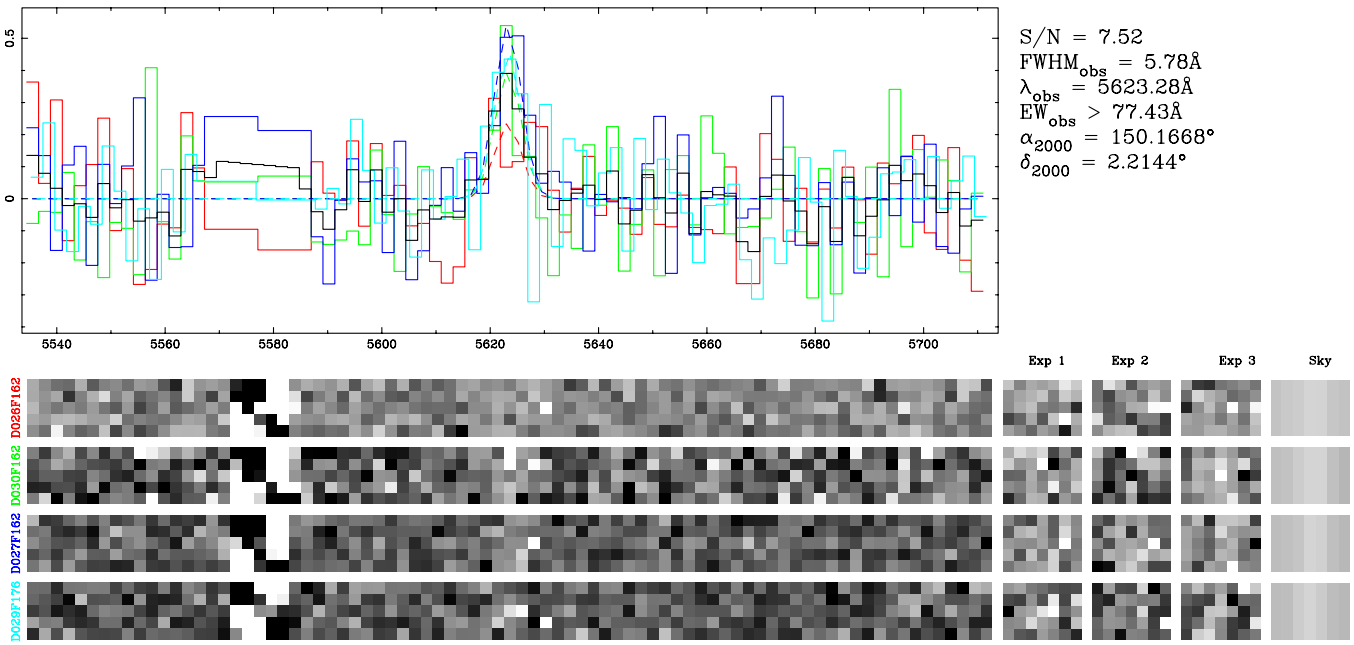

Figure 23. Emission-line source detection for index 308 which lies in the COSMOS field. The format is the same as in Figure 20. The spectrum-based EW does not go deep enough to discriminate between the classifications. This source neither shows alternate emission lines nor has an X-ray counterpart.

(A color version of this figure is available in the online journal.)

5.5. Example Sources

The rules to classify the emission-line objects have been described, but the display of the steps on actual VIRUS-P data is useful to establish confidence and the range of objects encountered. We will walk through the evidence for one emission source of each type and then give a summary display of representative subsamples.

5.5.1. LAEs

The detection image of Figure 20 shows source index 229 in Tables 3 and 4 as a broadened, bright emission line detected in four fiber positions. The high flux ($41.6^{+4.2}_{-5.0} \times 10^{-17}$ erg s^{-1} cm^{-2}) and the lack of a spectral continuum detection are already sufficient to meet the classification cut as an LAE. However, classification from the spectrum alone is only possible for the brightest emission-line sources in this sample. There are no counterpart emission lines at any of the tested transitions (Section 5.1), nor is there any associated X-ray detection. We next move to the deep, COSMOS *R* image in Figure 21. There, we find three plausible broadband counterparts with the brightest counterpart dominating the likelihood. There is no literature redshift for this object, but our Lyα line identification leads to EW$_0 = 51^{+8}_{-8}$ Å, so the object is classified as an LAE. We give compact detection images for five additional LAEs in Figure 22.

5.5.2. Low-z Objects

The detection image in Figure 23 shows source index 308 as a high S/N, only slightly and not significantly broadened, emission-line source detected jointly in four fiber positions. The line flux is $18.4^{+3.5}_{-4.2} \times 10^{-17}$ erg s^{-1} cm^{-2}, and the lack of a spectral continuum is insufficient for our EW-based classification scheme. There are no other emission lines detected in the object nor does the source have a X-ray counterpart. The COSMOS image (Figure 24) shows one bright continuum source barely offset from the emission-line centroid. A second, fainter object at larger separation is also analyzed at a much decreased likelihood, but would carry an LAE classification. Based on the most probable counterpart, we find a rest-frame EW, assuming the line to be Lyα, of EW$_0 = 8^{+2}_{-2}$ Å. This fails the EW cut, so we classify this as a low-redshift object, presumably an [O II] emitter. The actual rest-frame EW is then EW$_0 = 25^{+6}_{-6}$ Å. We note that association with the other possible

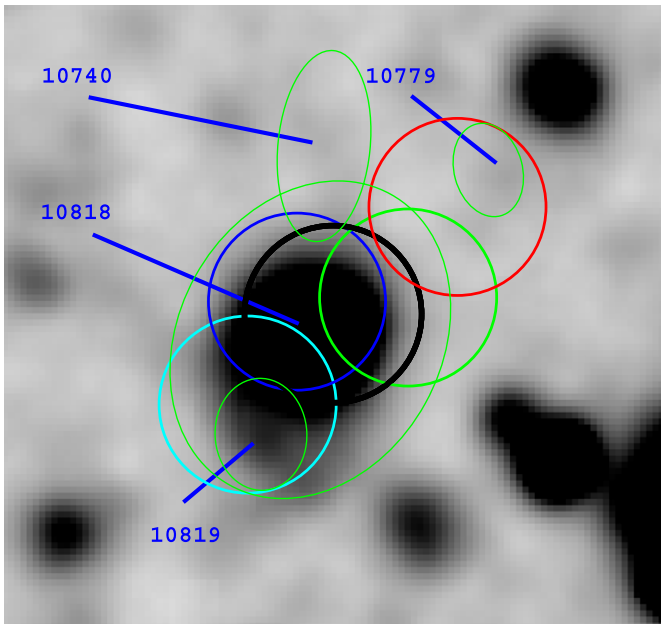

Figure 24. Ground-based, V_j imaging cutout for the Table 3 index 308. The detection spectra are given in Figure 23. The format is the same as in Figure 21. The counterpart 10818 at $r^+ = 22.1$ has a 99.5% likelihood of being associated. The observed EW of $30.9^{+14.9}_{-9.5}$ Å leads to a firm low-*z* classification, presumably for [O II].

(A color version of this figure is available in the online journal.)

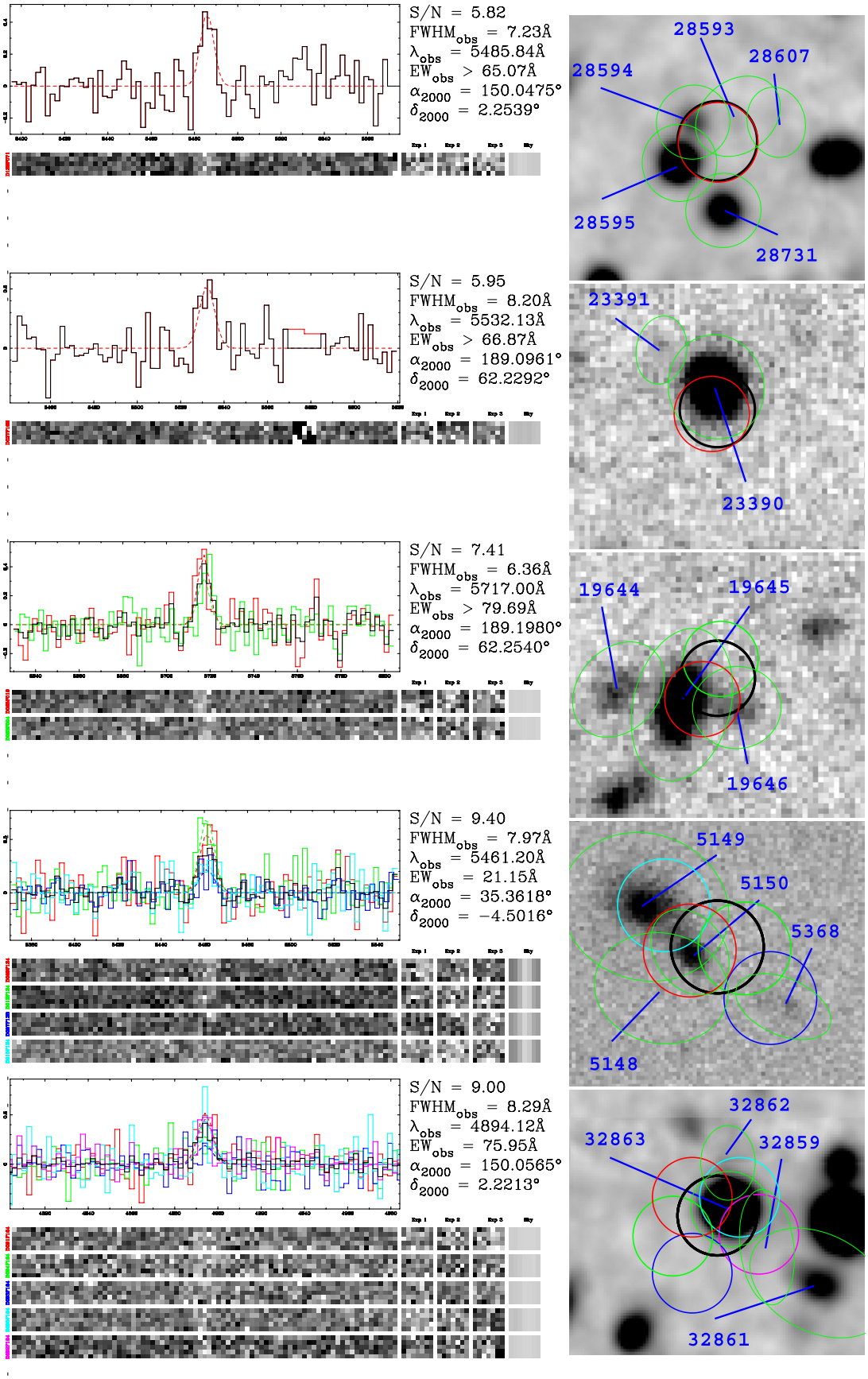

Figure 25. Five example detections of low-z objects with the format of Figure 23. The first and fifth lie in the COSMOS field, the second and third lie in the GOODS-N field both with previously measured redshifts, and the fourth lies in the XMM-LSS field. The entries from Table 3 for these five are indices 178, 351, 406, 33, and 192. The best continuum counterpart matches are to 28595, 23390, 19645, 5150, and 32863.

(A color version of this figure is available in the online journal.)

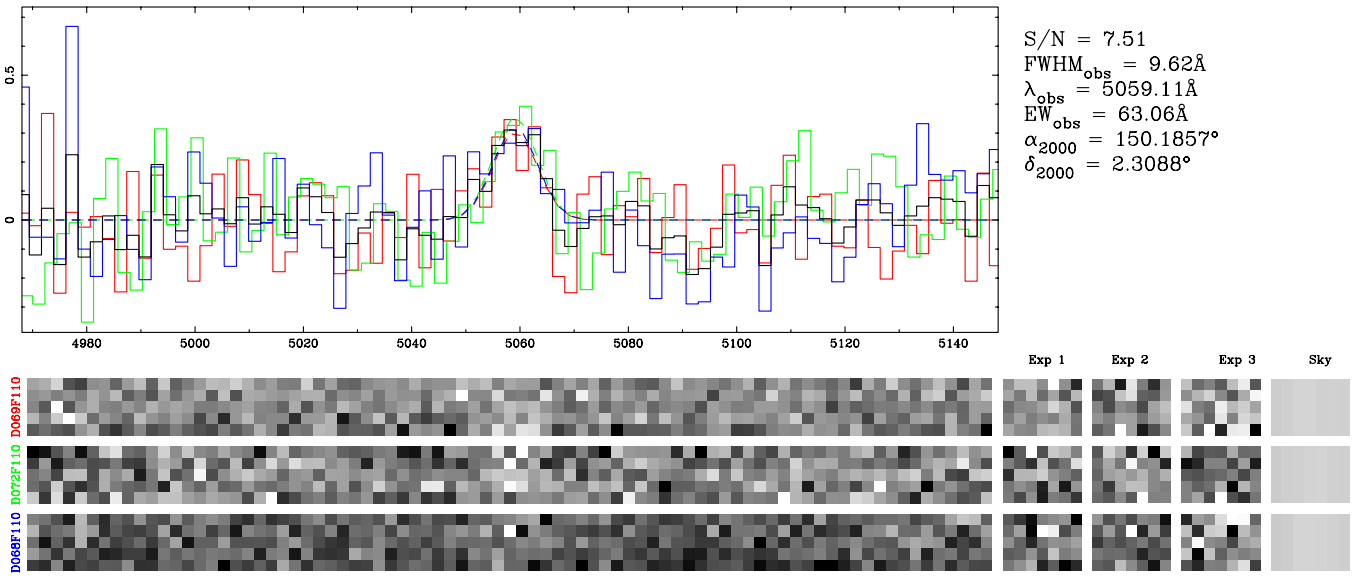

Figure 26. Emission-line source detection for index 322 which lies in the COSMOS field. The spectrum-based EW suggests a low-z classification. The format is the same as in Figure 20. Note that the emission line is broadened compared to the instrumental resolution.

(A color version of this figure is available in the online journal.)

counterpart would lead to the opposite conclusion. However, the likelihood of that association is quite low, $P = 0.1\%$, so we confidently classify the source as an [O II] emitter. We give compact detection images for five additional low-redshift objects in Figure 25.

5.5.3. Objects with Uncertain Redshifts

The detection image in Figure 26 shows source index 322 as a high S/N, broadened line along with a continuum detection in three fiber positions. The emission-line flux is $16.7^{+4.4}_{-3.3} \times 10^{-17}$ erg s^{-1} cm^{-2}. The COSMOS image (Figure 27) reveals three plausible counterparts. The most likely (84%) counterpart implies an easy classification as an [O II] emitter with EW$_{obs}$ = 37^{+8}_{-9} Å. However, there is a non-trivial likelihood (8%) that the counterpart not the bright galaxy, but instead the fainter object. In this case, the source would be classified as an LAE. Despite this uncertainty, we place this object in Tables 4 with an [O II] classification.

6. EMISSION-LINE SOURCE CATALOG

6.1. GOODS-N Comparisons

Most of the detections and redshift classifications in this catalog are new to the literature. For instance in the COSMOS field, the magnitude limit of the spectroscopic cut ($I_{AB} < 22.5$) to zCOSMOS (Lilly et al. 2009) gives little overlap with this sample. Fortunately, the large number of deep spectroscopic observations in the GOODS-N field comprises a better test sample. We have made a detailed comparison of our measurements to those of Barger et al. (2008), which includes most previous GOODS-N measured redshifts. We further include one Lyα match from Lowenthal et al. (1997) and one [O III]5007 source from Wirth et al. (2004). We note that the observations from the literature often have larger spectral coverage and higher resolution than our data, allowing alternate classification methods.

We find 119 unique emission-line sources in GOODS-N. Three of these do not have measured optical broadband counterparts from the ground-based imaging, appear to be blended

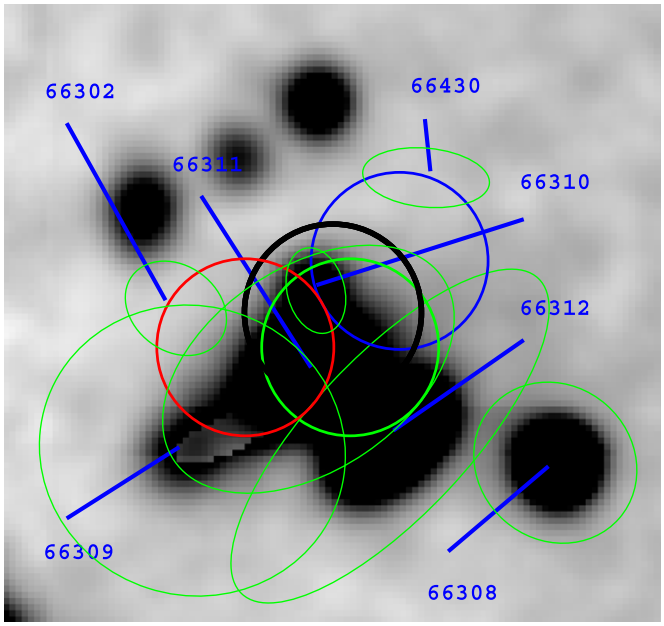

Figure 27. Ground-based, V_j imaging cutout for the Table 3 index 322. The detection spectra are given in Figure 26. The format is the same as in Figure 21. The counterpart 66311 at $r^+ = 21.1$ has a 84% likelihood in association. The counterpart 66312 is already assigned to the VIRUS-P detection of Table 3 index 310 at $\lambda_{obs} = 4948.2$ Å. The counterpart 66310 at $r^+ = 24.9$ would be an LAE based on EW and looks like a reasonable candidate system resolved from 66311 in the *HST* image, but it only holds a 6% chance of association. This source is associated with 66311 as an [O II] emitter in the catalog.

(A color version of this figure is available in the online journal.)

with foreground objects when examined with the *Hubble Space Telescope* (*HST*) images of Giavalisco et al. (2004), and are without published redshifts. We classify these as LAEs. In addition, there are nine other LAEs where we do measure a robust continuum counterpart but that are without published redshifts. We give the 12 new LAEs in GOODS-N in Figures 28 and 29. In addition, we find 92 low-z objects in common to the literature and 13 high-z objects (12 Lyα and 1 [C IV]). Finally, we find

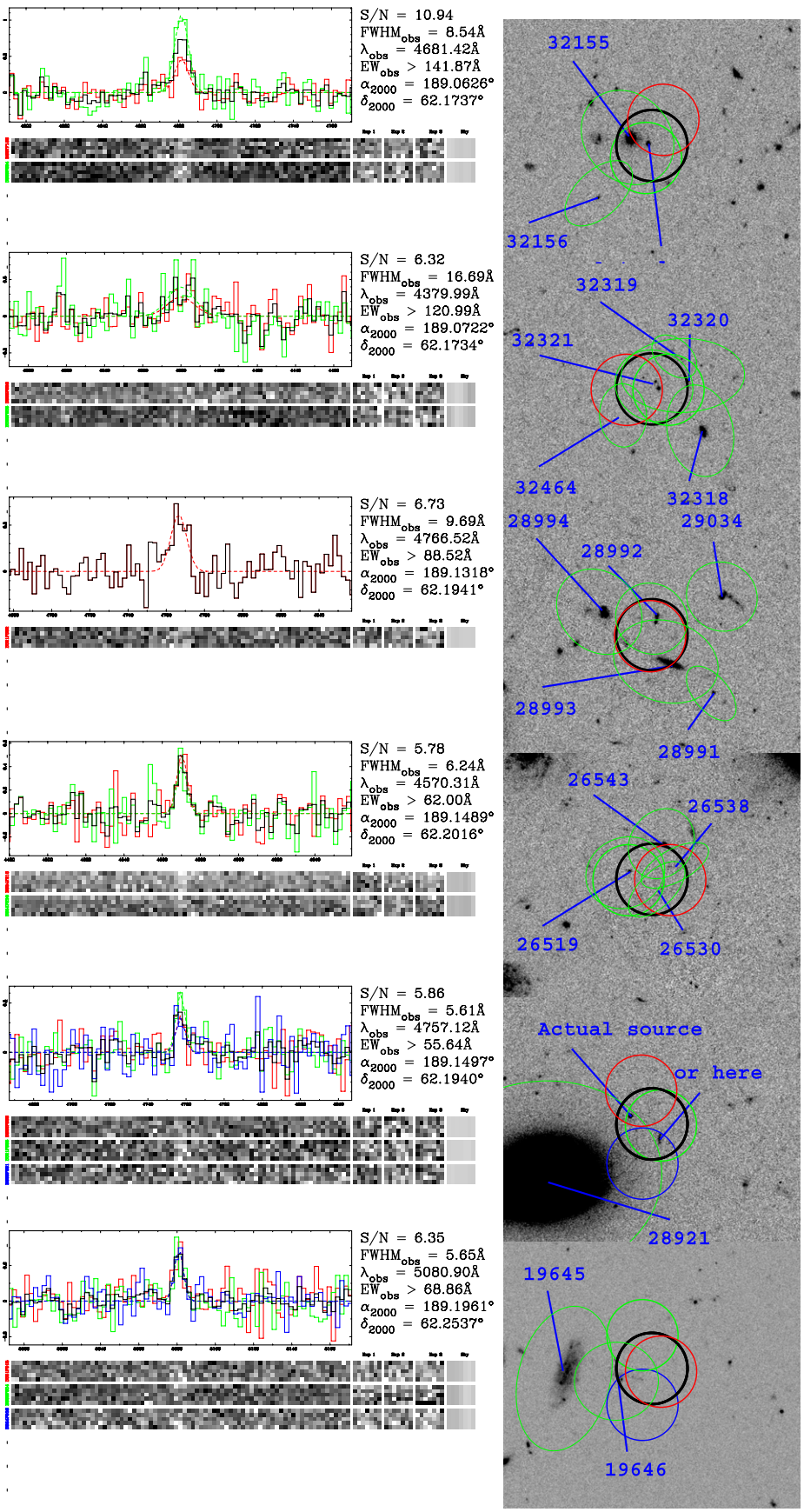

Figure 28. Spectral detection and *HST* ACS *F*606*W* (Giavalisco et al. 2004) cutouts for 6 of the 12 new, high-z redshift measurements in GOODS-N. From top to bottom, the objects are indices 334, 338, 360, 372, 373, and 403 in Table 4. The first and fifth objects do not have identified counterparts in the ground-based images due to blending, although likely counterparts are identified in the *HST* data. The best counterparts for the second, third, fourth, and sixth objects are 32321, 28992, 26519, and 19646.

(A color version of this figure is available in the online journal.)

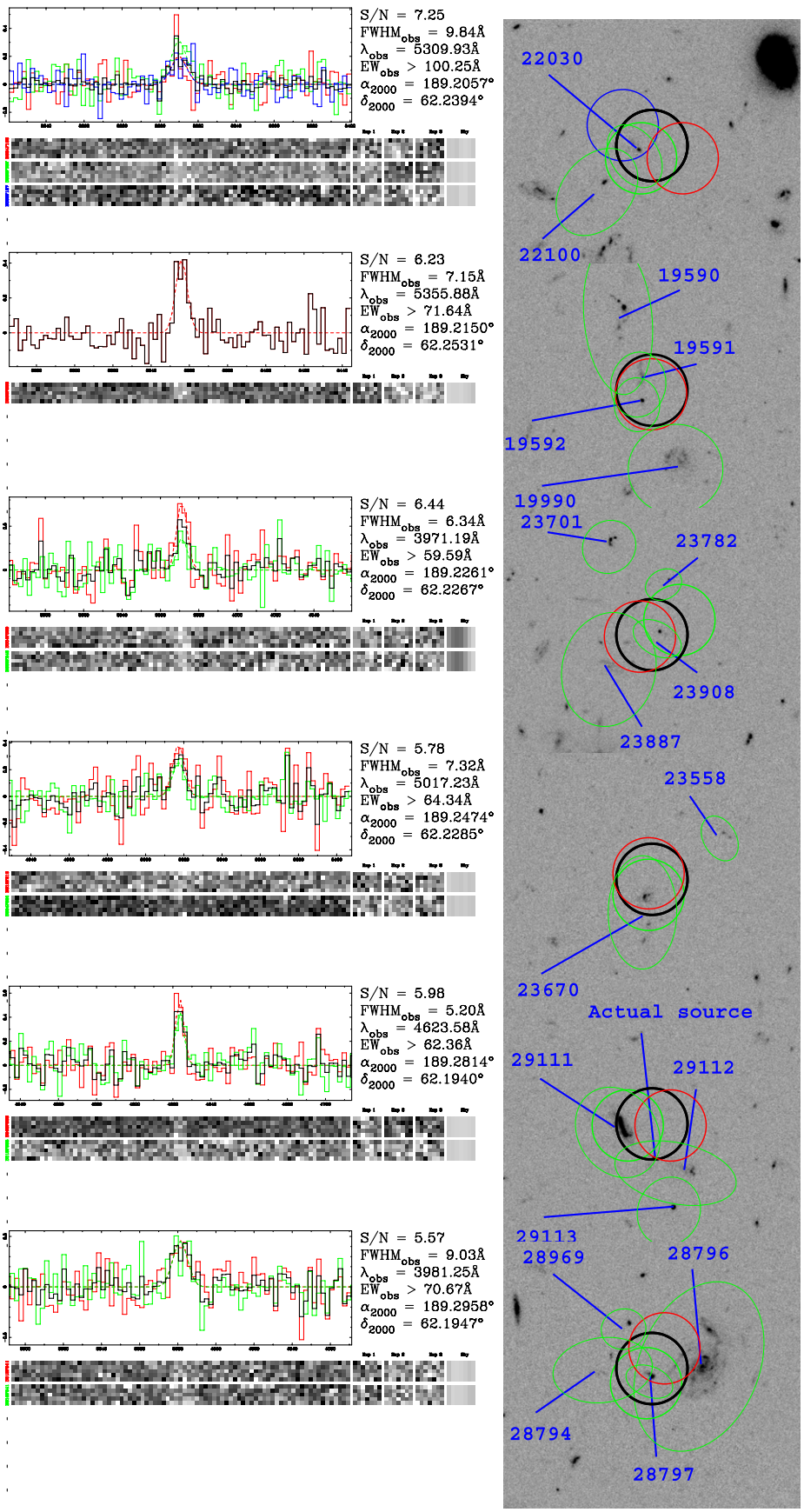

Figure 29. Spectral detection and *HST* ACS *F*606*W* (Giavalisco et al. 2004) cutouts for the final 6 of 12 new, high-z redshift measurements in GOODS-N. From top to bottom, the objects are indices 415, 426, 434, 447, 467, and 474 in Table 4. The object with index 467 does not have an identified counterpart in the ground-based images due to blending, although a likely counterpart is identified in the *HST* data. The best counterparts for the remaining five, in order of listing, are 22030, 19592, 23908, 23670, and 28797.

(A color version of this figure is available in the online journal.)

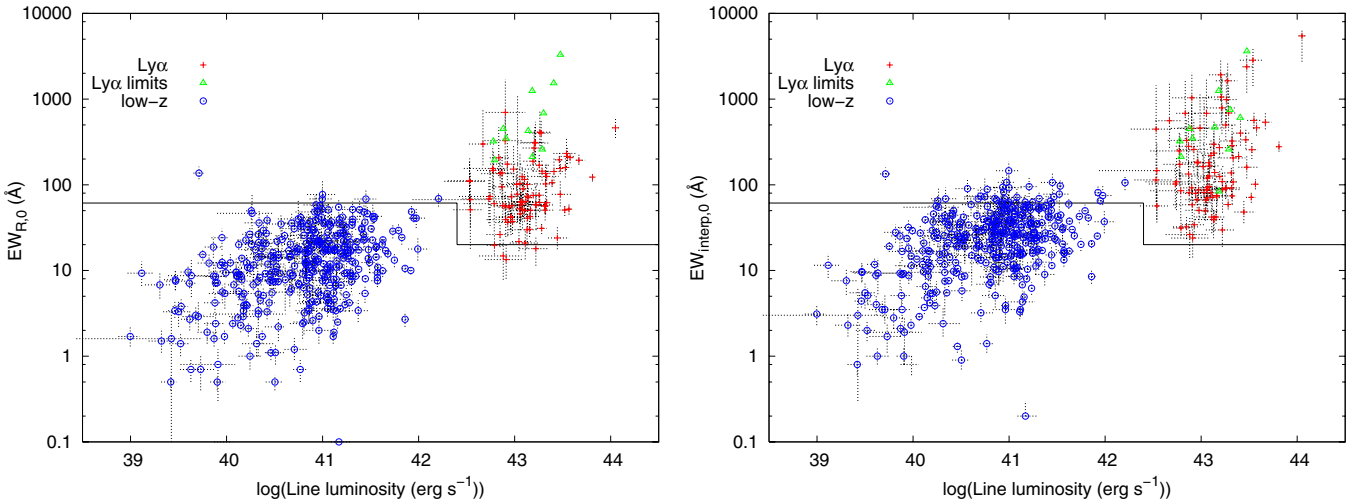

Figure 30. Distribution of rest-frame equivalent width (EW) and line luminosity for both the LAE and low-z samples. The primary classification line based on EW is drawn. The jog in the EW cut line is simply due to the cut being defined as EW$_{rest}$ > 20 Å assuming the line to be Lyα so the equivalent threshold in the [O II] rest frame is 61 Å. The drawn EW cut does not strictly apply to the low-z objects with emission at transitions other than [O II]. A trend between higher EW and line luminosity in the LAEs is somewhat visible but noisy over this survey's dynamic range. The same trend is seen in surveys with lower flux limits and discussed in Cassata et al. (2010). Left: continuum estimated only from the R-band photometry (or the i' band in MUNICS). Right: continuum estimated from interpolation with the two nearest filters bounding each emission-line's wavelength.

(A color version of this figure is available in the online journal.)

only two objects in our catalog were misclassified. Source index 371 was originally called an LAE, but the literature reveals it to be a [C IV]1549 emitter. Source index 400 was originally called an [O II] emitter, but the literature reveals it to be an [O III]5007 emitter. We have rectified Table 4 to reflect these two cases, and we then find an rms in $\frac{\Delta z}{1+z}$ of 0.001 and no offset compared to the literature, which is completely consistent with our 0.5 Å line center uncertainty. A weakness in the literature samples is the lack of emission-line flux calibration, so we cannot use the previous samples to quantitatively test this survey's completeness. We have qualitatively confirmed the completeness by searching for literature objects in our spectra and finding many dozen at $3 < S/N < 5$.

6.2. Catalog Summary

Table 3 contains a segment of the detected emission-line catalog with the full version available electronically. The entry "..." is given where there is not an applicable value. Each emission line is prefixed with the identifier "HPS" to stand for HETDEX Pilot Survey. The column descriptions are (1) the catalog number, (2) the emission-line right ascension in hr:minute:s (J2000), (3) the emission-line declination °:′:″ (J2000), (4) the observed emission wavelength in a vacuum, heliocentric frame (Å) with an estimated 0.5 Å uncertainty based on simulations, (5) the intrinsic FWHM of the emission line (km s^{-1}) after removal of a 5 Å FWHM instrumental resolution and (with an estimated 300 km s^{-1} uncertainty), (6) the S/N of the emission-line flux detected within the aperture set of fibers, (7) the emission-line flux and error in 10^{-17} erg s^{-1} cm^{-2} as measured with the CoG method (Section 4.4), (8) the spatial FWHM of the emission line (″) as measured with the CoG method (Section 4.4), and (9) any additional entries in the table that share a position and redshift with the emission line (i.e., as with detections of other emission lines from the same source).

Table 4 shows a segment of the counterpart and classification information for each emission-line detection with the full version available electronically. The entry "..." is given where there is not an applicable value. The column descriptions

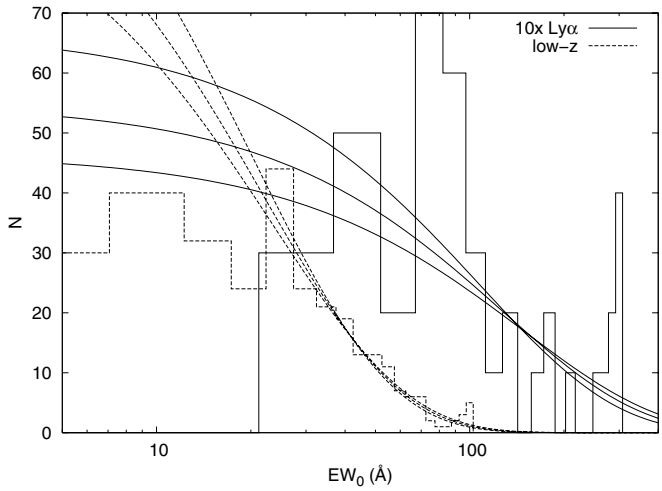

Figure 31. Distribution of rest-frame equivalent width (EW) for values with S/N > 3 for both the LAE and low-z sample by interpolating from the bounding broadband data. The LAE histogram has been scaled by 10× for visual clarity. A maximum likelihood fit was made by taking EW$_{rest}$ > 20 Å where the samples should be complete. An exponential scale length of 128 ± 20 Å fits the LAE distribution and 22 ± 1.6 Å fits the [O II] distribution. The exponential fits and error ranges are also plotted. The largest plotted bins contain all values that lie higher than the histogram range.

are: (1) the catalog number, (2) the best continuum-selected counterpart in the standard J2000 naming convention, (3) the R-band magnitude for this best counterpart (or the i' magnitude for MUNICS), (4) the probability of counterpart association (from Equation (5)), (5) the rest-frame EW and uncertainties for this counterpart and the selected transition based on the R-band photometry, or the i'-band in MUNICS where no R band is available (Å), (6) the rest-frame EW and uncertainties for this counterpart and the selected transition based on an interpolation between the two nearest filters (Å), (7) the transition of the emission line based on the EW cut and/or the presence of multiple emission lines, (8) the estimated redshift, (9) the probability of the emission line being Lyα as calculated by

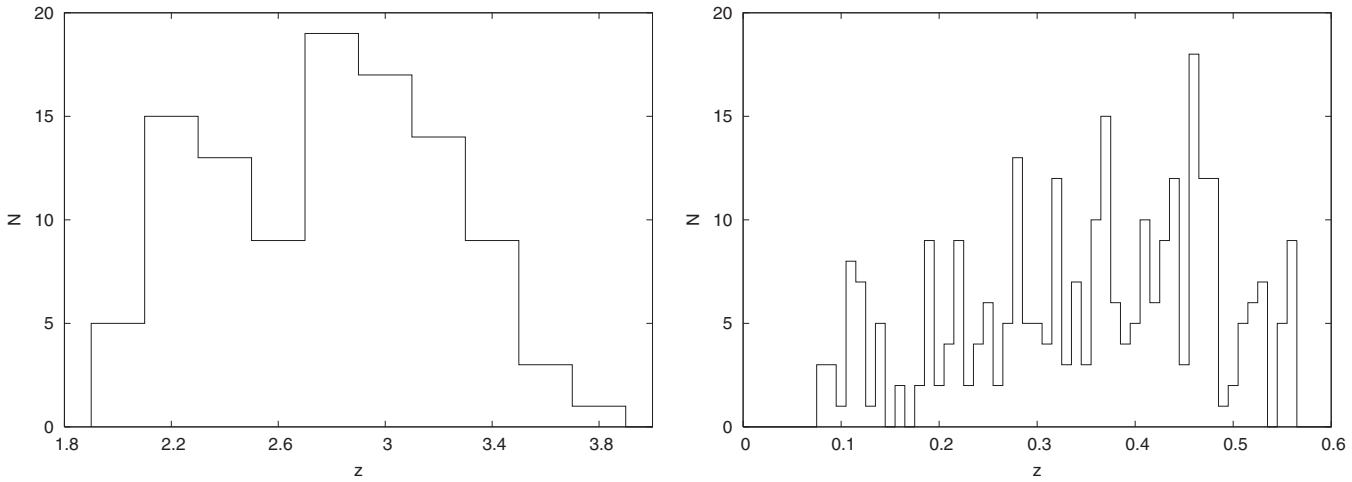

Figure 32. Distribution in redshift for all the survey objects. Left: the LAE sample. Right: the low-z sample. Each significantly overdense low-z bin is primarily contained in one survey field. The overdensities may be early indicators of groups or clusters, but the low number statistics preclude firm classification. There are no clusters from the Koester et al. (2007) catalog in any of this pilot survey's area.

marginalizing over all potential counterparts, and (10) the X-ray counterpart in the standard J2000 naming convention.

6.3. Catalog Properties

Figure 30 compares the distribution of rest-frame EW for LAEs and [O II] emitters to the emission-line luminosity. Histograms of the rest-frame EW distributions of both low- and high-redshift sources are shown in Figure 31. A maximum likelihood fit was made by taking $EW_{rest} > 20$ Å where the samples should be complete. An exponential scale length of 128 ± 20 Å fits the LAE distribution and 22 ± 1.6 Å fits the [O II] distribution. The redshift distribution of all sources is given in Figure 32. No previously identified groups or clusters lie in our fields (e.g., Koester et al. 2007). In Figure 33, we give the color–color diagram for the sample's LAEs. We do not try to transform the filter systems into filter sets which are usually applied to LBG and BX galaxy samples, but we do plot the location of the LAE spectral template from Gawiser et al. (2007), made from Maraston (2005) stellar population synthesis modeling, over the relevant redshifts. We also show the locus of stars from Pickles (1998). Many of the LAEs appear consistent with color space expectations based on continuum-selected samples.

6.3.1. Spatially Extended High-z Sources

Based on the detection threshold of Section 4.3, we find five objects whose Lyα emission appears significantly extended. Figure 34 gives the detection and broadband images for the objects, and Figure 35 shows the CoG analysis which determines their sizes. These five objects are indices 99 and 126 in the MUNICS field and indices 162, 164, and 261 in the COSMOS field. Index 99 is also a high EW object. Indices 162 and 261 are high EW if one use the R-band continua for the EW estimation, and they are both X-ray sources.

6.3.2. High EW LAEs

LAEs with $EW_{rest} > 240$ Å are potential sites of exotic energy sources or unusual metallicity since stellar population modeling has shown that a normal initial mass function cannot produce such high EWs (Charlot & Fall 1993). If we consider our whole catalog and use the EW measurements derived from interpolating with the two nearest filters, we find 11 LAEs without broadband counterparts and a further 21 LAEs with counterparts that have $EW_{rest} > 240$ Å at $>1\sigma$ significance. However, in order to make a conservative estimate, we instead use EW estimates based on the R-band photometry only and restrict the discussion to sources with emission-line detection $S/N > 6.5$ to avoid false detections. This instead leaves only one LAE without a counterpart and two LAEs with counterparts meeting the high EW criterion. We note that a number of the emission lines without broadband counterparts may have their origins obscured by ground-based seeing. For instance, three of the objects with new redshifts in GOODS-N have their best counterparts blended behind foreground object and are shown as part of Figures 28 and 29. For homogeneity between all fields, we only measure continua from the ground-based images in this work. Some of the entries in Table 4 as being without counterparts may be caused by blending and not image depth. In fact, the only emission-line detection with high confidence as being without a counterpart is index 314.

Figure 36 shows the detection and image data for the three significant high EW LAEs. The top figures show the data for index 314, an LAE with $z = 2.6312$ but with no counterpart in the COSMOS image and $EW_{rest} > 348$ Å (1σ). The middle figures show the data for index 126 in MUNICS. Although the counterpart is fairly bright at $m_{i'} = 24.3$, the very bright emission line implies $EW_{rest} > 352$ Å (1σ). We note that this $z = 2.8276$ object is also significantly extended in Lyα with FWHM $= 7\farcs5$. Finally, the bottom figures show the data for index 231 in COSMOS with $EW_{rest} > 282$ Å (1σ) and $z = 2.7215$. This object is marginally extended in Lyα (FWHM $= 6\farcs3$), but also compatible with a point source and poor seeing. We find a 3% high-EW fraction in the LAE sample by our best estimates. However, the fraction could be as high as 31% by our most inclusive criteria.

6.3.3. LAE Number Density Expectation

The spectral and spatial sensitivity limits along with the completeness simulation of Section 4.3 completely define the survey's selection characteristics and are necessary inputs to the luminosity function calculations that will follow in future papers (Blanc et al. 2010). By considering all these effects, namely, the completeness distribution we are able to achieve with the detection routine (Section 4.1) and simulated data (Section 4.3),

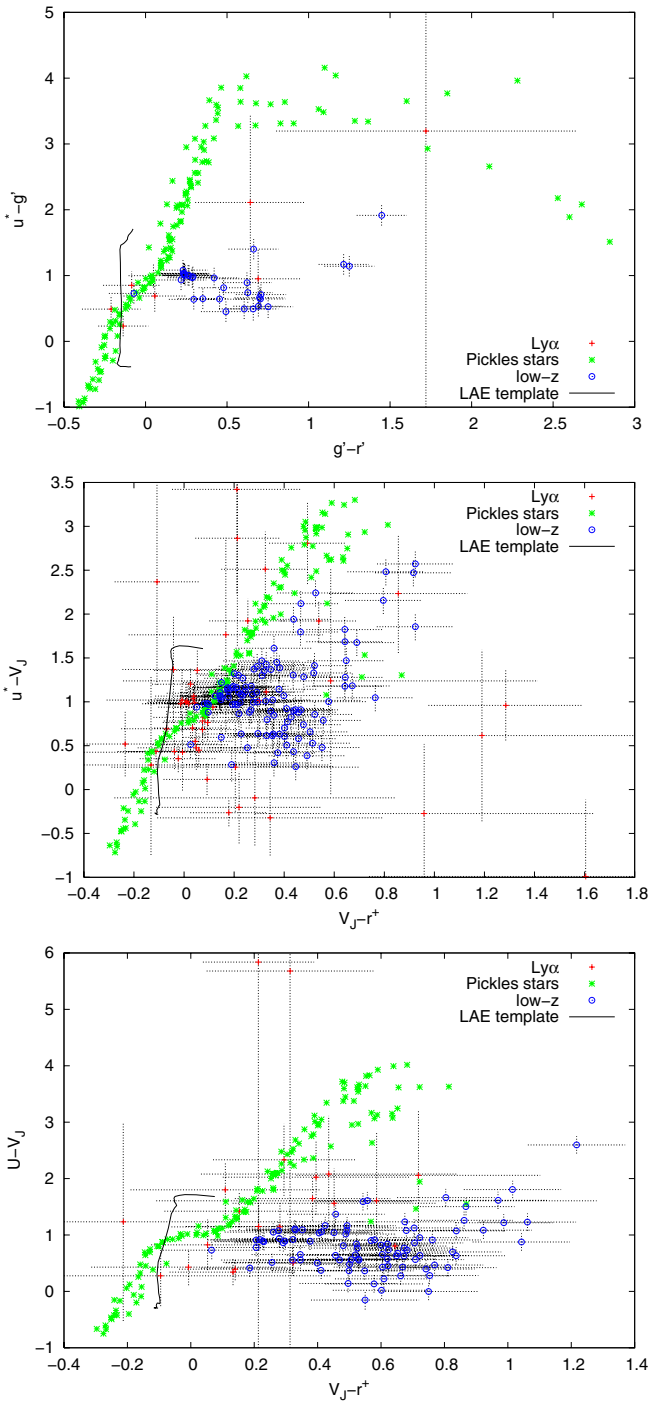

Figure 33. Color–color plots in the photometry bands that commonly define the Lyman Break Galaxy selection. Detections are not shown for the MUNICS field where we lack *U*-band data. The LBG selection rules are sensitive to the exact filter and telescope choice, so we do not transform these filter data into systems with published LBG rules. Instead, we synthesize colors of the Gawiser et al. (2007) LAE template as the solid, black curve for $1.3 < z < 4.5$ and stars (Pickles 1998). Albeit with some exceptions and frequently large color errors, the LAE sample is segregated from the low-z objects and lies where expected. Top: XMM-LSS objects, middle: COSMOS objects, and bottom: GOODS-N objects.

(A color version of this figure is available in the online journal.)

and finally the LAE luminosity function of Gronwall et al. (2007), we predict that this sample should contain 121 LAEs. The dominant uncertainty in this prediction is cosmic variance, which can be approximated by linear theory for a given redshift, volume, and number density from Somerville et al. (2004, Figure 3). The effective Lyα survey volume in Gronwall et al. (2007) is $1.1 \times 10^5\ h_{70}^{-3}$ Mpc^3 which implies a relative cosmic variance of $\sigma_v \sim 35\%$ and $\sigma_v \sim 15\%$ for the volume of this survey. Within these factors, the LAE number statistics from this survey are low but not in serious conflict with earlier determinations.

7. SUMMARY

We present untargetted integral field spectroscopic observations over 169 □′ with the goal of characterizing emission-line galaxies at low ($z < 0.56$) and high ($z \gtrapprox 2$) redshifts. In this first of a series of papers, we describe the design, observations, calibrations, reductions, detections, measurements, and classification methods for the survey. The primary classification method we employ uses EW cuts computed by matching the emission-line objects to continuum counterparts in existing, deep images. We find that effective object classification can be made using $\mathrm{EW_{rest}} > 20$ Å where the continuum is defined using a single band of deep photometry, preferably in the *R* band. We find 397 unique emission-line galaxies: 168 over a 71.56 □′ area in the COSMOS field, 118 over a 35.52 □′ area in the GOODS-N field, 79 over a 49.85 □′ area in the MUNICS field, and 32 over a 12.30 □′ area in the XMM-LSS field. The two transitions most frequently observed are [O II] (285 galaxies) and Lyα (105 galaxies). Based on a non-evolving Gronwall et al. (2007) luminosity function, we should have detected 121 LAEs in this survey; the difference is within the range of cosmic variance. The field with the deepest X-ray data (GOODS-N) shows an AGN fraction in the LAE sample consistent with that of the shallower fields (6%). We compare our data to the extensive GOODS-N targeted spectroscopy to verify our object classification and confirm 92 low-z and 13 high-z galaxies. Moreover, we derive new redshifts for a further 2 low-z and 12 high-z galaxies in the GOODS-N field. Over all fields, 11 high-z objects do not possess optical counterparts despite the imaging depth; these are either very high EW objects, contamination from noise, or objects whose counterparts have been blended by ground-based seeing. However, within the remaining LAEs we find a distribution of $\mathrm{EW_0}$ that can be described by an exponential scale length of 128 ± 20 Å and with only three objects at $\mathrm{EW_0} > 240$ Å at $>1\sigma$ significance. Many of the newly discovered LAEs lie in the color ranges consistent with previous work.

The main contaminant in our LAE sample is simply noise, which should be 10% of the LAE sample based on simulations. A totally pure subsample of 68 LAEs can be defined using this catalog at $\mathrm{S/N} > 6$. We find five sources of Lyα emission that have a high significance as being spatially resolved, at least two of which are AGNs. The pilot survey has validated that IFS searches for LAEs perform as expected. The forthcoming larger FOV of the full HETDEX survey will vastly improve the survey efficiency of this method.

We thank the staff of McDonald Observatory and the engineers and machinists of UT-Austin and the Astrophysikalisches Institut Potsdam for their indispensable work. We thank Povilas Palunas for field planning and instrument commissioning work, Carlos Allende Prieto for Coudé reduction advice, Niel Brandt for advice on the field selection, and Hisanori Furusawa for generously providing his SXDS imaging number counts in electronic form. We thank the Cynthia and George Mitchell Foundation for funding the VIRUS-P instrument. This work required substantial time investment from the McDonald

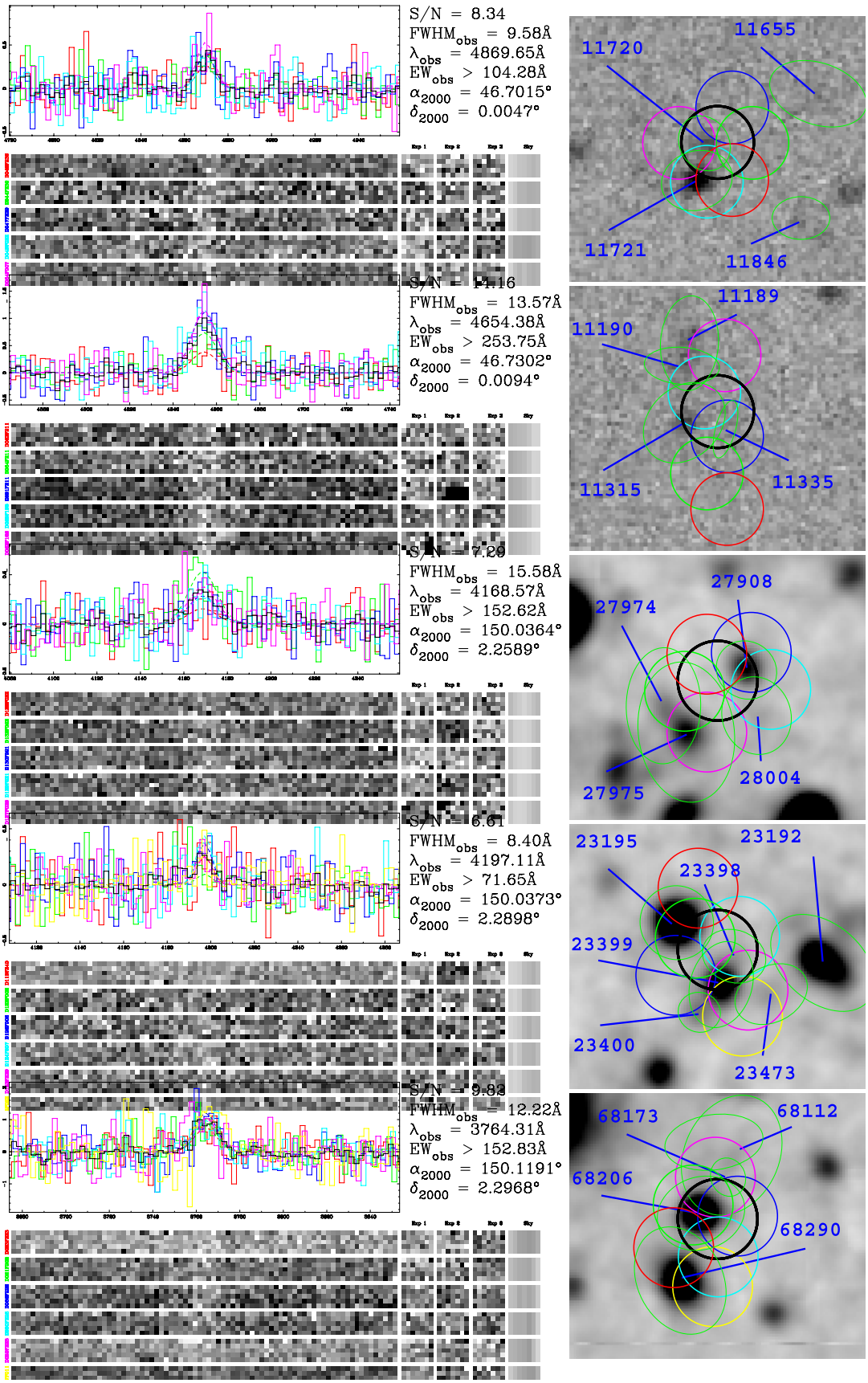

Figure 34. Five objects significantly extended in Lyα. Left: the spectral detections. Right: the *R*-band images (i' for the first two MUNICS objects). First: Index 99 in MUNICS. The best counterpart is 11720. Second: Index 126 in MUNICS. The best counterpart is 11315. This source is also a high EW LAE. Third: Index 162 in COSMOS. The best counterpart is 27975. This source has an X-ray detection. Fourth: Index 164 in COSMOS. The best counterpart is 23399. Fifth: Index 261 in COSMOS. The best counterpart is 68206. This source has an X-ray detection.

(A color version of this figure is available in the online journal.)

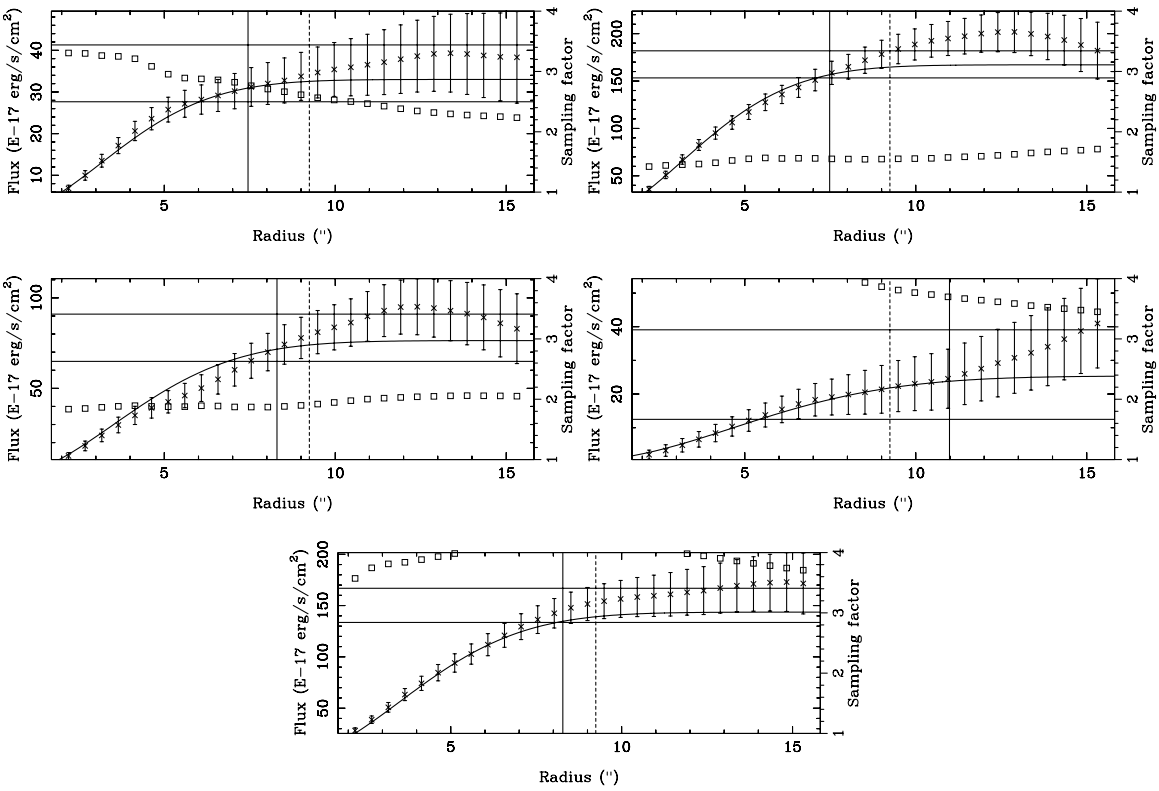

Figure 35. Curve-of-growth plots for the five significant objects extended in Lyα. The format is the same as in Figure 13 and described therein. Top left: Index 99 in MUNICS. The best counterpart is 11720. Top right: Index 126 in MUNICS. The best counterpart is 11315. This source is also a high EW LAE. Middle left: Index 162 in COSMOS. The best counterpart is 27975. This source has an X-ray detection. Middle right: Index 164 in COSMOS. The best counterpart is 23399. Bottom: Index 261 in COSMOS. The best counterpart is 68206. This source has an X-ray detection.

Observatory and was only possible with the support and patience of the director, Dr. David Lambert. Salary and travel support for students was provided by a Texas Norman Hackerman Advanced Research Program under grants ARP 003658-0005-2006 and 003658-0295-2007. J.J.A. acknowledges the support of a National Science Foundation Graduate Research Fellowship and the University of Texas at Austin's William S. Livingston Award, David Bruton Jr. Fellowship, and Donald P. Harrington Fellowship. This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. Some of the data presented in this paper were obtained from the Multimission Archive at the Space Telescope Science Institute (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-*HST* data is provided by the NASA Office of Space Science via grant NNX09AF08G and by other grants and contracts.

Facility: Smith (VIRUS-P)

APPENDIX

ATMOSPHERIC DIFFERENTIAL REFRACTION ASTROMETRY CORRECTION

Atmospheric differential refraction (ADR) effects over this data set's wavelength and airmass ranges are of the order of the astrometric solution errors, so we have made an astrometry correction to the guider-based positions considering the emission-line source wavelength. There are two ADR effects on the observed fiber positions at any given wavelength: the atmosphere's wavelength-dependent index of refraction at a fixed airmass and the different airmasses between the science and guider FOVs. As we have stated in Section 2.1, the guider's effective wavelength is 5000 Å, and we ignore color corrections for different guide stars. In order to retain the ability to stack exposures taken at the same dither position, we average the positional differences over the N exposures and apply Equation (A1) from Smart & Green (1977) as the average positional corrections for an emission-line source at wavelength λ due to ADR where ϕ is the site latitude, δ is the declination, θ_g is the distance angle between the guider and IFU centers, H is the hour angle at the middle of the frame's exposure, and k is the constant of mean refraction calculated (Filippenko 1982) for average 2 km altitude conditions (Allen 1973) and related to the atmosphere's index of refraction. Common corrections derived this way are 0″–2″ with a median of 0.″3 for our sample:

$$\Delta\alpha = \sum_{i=1}^{N} \frac{k(\lambda)\sec^2\delta\sin H}{N\times(\tan\delta\tan\phi+\cos H)} - \frac{k(5000\,\text{Å})\sec^2(\delta+\theta_g)\sin H}{N\times(\tan(\delta+\theta_g)\tan\phi+\cos H)}$$

$$\Delta\delta = \sum_{i=1}^{N} \frac{k(\lambda)\times(\tan\phi-\tan\delta\cos H)}{N\times(\tan\delta\tan\phi+\cos H)} - \frac{k(5000\,\text{Å})\times(\tan\phi-\tan(\delta+\theta_g)\cos H)}{N\times(\tan(\delta+\theta_g)\tan\phi+\cos H)}. \quad \text{(A1)}$$

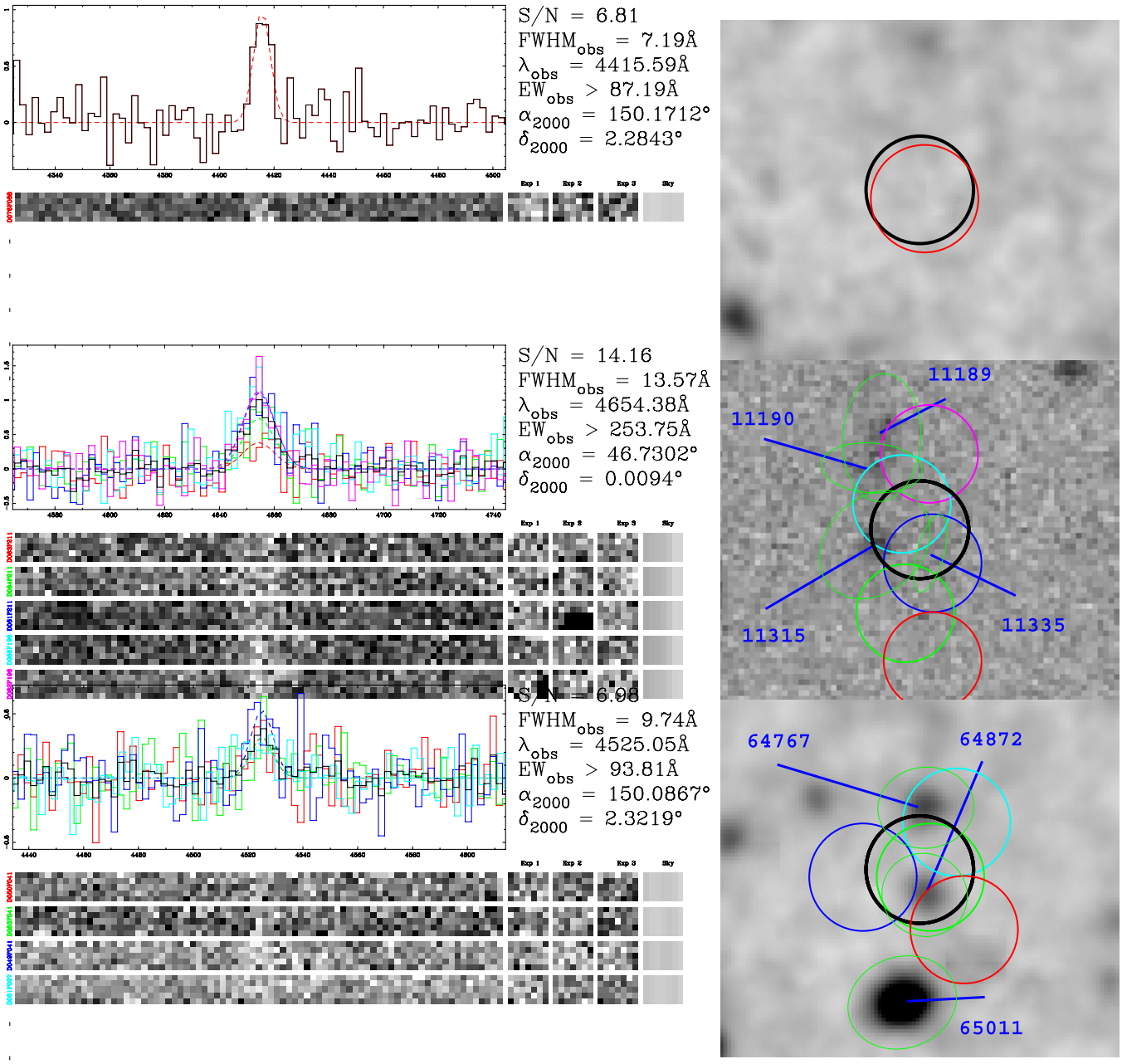

Figure 36. Three high-significance LAEs with EW$_0$ >240 Å. Left: the spectral detection figures. Right: the *R*-band images. Top: Index 314 in the COSMOS field. No continuum counterpart is found. Middle: Index 126 in the MUNICS field. The best counterpart is listed as 113115. The emission line is also significantly spatially extended. Bottom: Index 231 in the COSMOS field. The best counterpart is listed as 64872.

(A color version of this figure is available in the online journal.)

REFERENCES

Adams, J. J., Hill, G. J., & MacQueen, P. J. 2008, Proc. SPIE, 7014, 232
Adelman-McCarthy, J. K., et al. 2008, ApJS, 175, 297
Alexander, D. M., et al. 2003, AJ, 126, 539
Allen, C. W. (ed.) 1973, Astrophysical Quantities (3rd ed.; London: Athlone Press)
Barden, S. C., Sawyer, D. G., & Honeycutt, R. K. 1998, Proc. SPIE, 3355, 892
Barger, A. J., Cowie, L. L., & Wang, W. 2008, ApJ, 689, 687
Becker, T. 2001, Dissertation, Univ. Potsdam
Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Blanc, G., et al. 2010, ApJ, submitted
Bond, N. A., Feldmeier, J. J., Matković, A., Gronwall, C., Ciardullo, R., & Gawiser, E. 2010, ApJ, 716, L200
Bongiovanni, É., et al. 2010, A&A, 519, L4
Bower, R. G., et al. 2004, MNRAS, 351, 63
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Capak, P., et al. 2004, AJ, 127, 180
Capak, P., et al. 2007, ApJS, 172, 99
Cappelluti, N., et al. 2007, ApJS, 172, 341
Cappelluti, N., et al. 2009, A&A, 497, 635
Cassata, P., et al. 2010, A&A, in press (arXiv:1003.3480)
Charlot, S., & Fall, S. M. 1993, ApJ, 415, 580
Cowie, L. L., & Hu, E. M. 1998, AJ, 115, 1319
Daddi, E., Cimatti, A., Renzini, A., Fontana, A., Mignoli, M., Pozzetti, L., Tozzi, P., & Zamorani, G. 2004, ApJ, 617, 746
Dickinson, M., & Giavalisco, M. 2003, in The Mass of Galaxies at Low and High Redshift, ed. R. Bender & A. Renzini (Berlin: Springer), 324
Dierckx, P. 1993, Curve and Surface Fitting with Splines (Monographs on Numerical Analysis; Oxford: Clarendon)
Drory, N., Feulner, G., Bender, R., Botzler, C. S., Hopp, U., Maraston, C., Mendes de Oliveira, C., & Snigula, J. 2001, MNRAS, 325, 550
Eddington, A. S. 1913, MNRAS, 73, 359

Eddington, A. S. 1940, MNRAS, 100, 354
Elvis, M., et al. 2009, ApJS, 184, 158
Fernández-Soto, A., Lanzetta, K. M., & Yahil, A. 1999, ApJ, 513, 34
Filippenko, A. V. 1982, PASP, 94, 715
Finkelstein, S. L., Rhoads, J. E., Malhotra, S., & Grogin, N. 2009, ApJ, 691, 465
Francis, P. J., et al. 2001, ApJ, 554, 1001
Furusawa, H., et al. 2008, ApJS, 176, 1
Gallego, J., García-Dabó, C. E., Zamorano, J., Aragón-Salamanca, A., & Rego, M. 2002, ApJ, 570, L1
Gawiser, E., et al. 2006, ApJ, 642, L13
Gawiser, E., et al. 2007, ApJ, 671, 278
Giavalisco, M., et al. 2004, ApJ, 600, L93
Gilbank, D. G., et al. 2010, MNRAS, 405, 2419
Greisen, E. W., & Calabretta, M. 1993, BAAS, 25, 808
Gronwall, C., et al. 2007, ApJ, 667, 79
Guaita, L., et al. 2010, ApJ, 714, 255
Hammer, F., et al. 1997, ApJ, 481, 49
Hayashino, T., et al. 2004, AJ, 128, 2073
Hill, G. J., Gebhardt, K., Komatsu, E., & MacQueen, P. J. 2004, in AIP Conf. Ser. 743, The New Cosmology: Conference on Strings and Cosmology, ed. R. E. Allen, D. V. Nanopoulos, & C. N. Pope (Melville, NY: AIP), 224
Hill, G. J., et al. 2008a, in ASP Conf. Ser. 399, ed. T. Kodama, T. Yamada, & K. Aoki (San Francisco, CA: ASP), 115
Hill, G. J., et al. 2008b, Proc. SPIE, 7014, 231
Hill, G. J., et al. 2010, Proc. SPIE, 7735, 21
Hogg, D. W., Cohen, J. G., Blandford, R., & Pahre, M. A. 1998, ApJ, 504, 622
Hu, E. M., Cowie, L. L., Capak, P., McMahon, R. G., Hayashino, T., & Komiyama, Y. 2004, AJ, 127, 563
Hu, E. M., Cowie, L. L., Kakazu, Y., & Barger, A. J. 2009, ApJ, 698, 2014
Hu, E. M., & McMahon, R. G. 1996, Nature, 382, 231
Ilbert, O., et al. 2009, ApJ, 690, 1236
Jeong, D., & Komatsu, E. 2006, ApJ, 651, 619
Jeong, D., & Komatsu, E. 2009, ApJ, 691, 569
Kakazu, Y., Cowie, L. L., & Hu, E. M. 2007, ApJ, 668, 853
Kelson, D. D. 2003, PASP, 115, 688
Kennicutt, R. C., Jr. 1992, ApJ, 388, 310
Koehler, R. S., Schuecker, P., & Gebhardt, K. 2007, A&A, 462, 7
Koester, B. P., et al. 2007, ApJ, 660, 239
Kornei, K. A., Shapley, A. E., Erb, D. K., Steidel, C. C., Reddy, N. A., Pettini, M., & Bogosavljević, M. 2010, ApJ, 711, 693
Kron, R. G. 1980, ApJS, 43, 305
Kurtz, M. J., & Mink, D. J. 1998, PASP, 110, 934
Lemaux, B. C., et al. 2009, ApJ, 700, 20
Lilly, S. J., et al. 2009, ApJS, 184, 218
Lowenthal, J. D., et al. 1997, ApJ, 481, 673
Luo, B., et al. 2010, ApJS, 187, 560
Ly, C., et al. 2007, ApJ, 657, 738
Maraston, C. 2005, MNRAS, 362, 799
Martin, C. L., Sawicki, M., Dressler, A., & McCarthy, P. 2008, ApJ, 679, 942
Martin, D. C., et al. 2005, ApJ, 619, L1
Massey, P., Strobel, K., Barnes, J. V., & Anderson, E. 1988, ApJ, 328, 315
Matsuda, Y., Yamada, T., Hayashino, T., Yamauchi, R., & Nakamura, Y. 2006, ApJ, 640, L123
Matsuda, Y., et al. 2004, AJ, 128, 569
Meurer, G. R., et al. 2007, AJ, 134, 77
Monet, D. G., et al. 2003, AJ, 125, 984
Murphy, J. D., MacQueen, P. J., Hill, G. J., Grupp, F., Kelz, A., Palunas, P., Roth, M., & Fry, A. 2008, Proc. SPIE, 7018, 2
Nilsson, K. K., Östlin, G., Møller, P., Möller-Nilsson, O., Tapken, C., Freudling, W., & Fynbo, J. P. U. 2010, arXiv:1009.0007
Nilsson, K. K., Tapken, C., Møller, P., Freudling, W., Fynbo, J. P. U., Meisenheimer, K., Laursen, P., & Östlin, G. 2009, A&A, 498, 13
Nilsson, K. K., et al. 2007, A&A, 471, 71
O'Donnell, J. E. 1994, ApJ, 422, 158
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Ouchi, M., et al. 2003, ApJ, 582, 60
Ouchi, M., et al. 2008, ApJS, 176, 301
Palunas, P., Teplitz, H. I., Francis, P. J., Williger, G. M., & Woodgate, B. E. 2004, ApJ, 602, 545
Pickles, A. J. 1998, PASP, 110, 863
Pierre, M., et al. 2004, J. Cosmol. Astropart. Phys., JCAP09(2004)011
Pirzkal, N., et al. 2004, ApJS, 154, 501
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (ed.) 1992, in Numerical Recipes in FORTRAN. The Art of Scientific Computing (Cambridge: Cambridge Univ. Press)
Ramsey, L. W., et al. 1998, Proc. SPIE, 3352, 34
Rauch, M., et al. 2008, ApJ, 681, 856
Rhoads, J. E., Malhotra, S., Dey, A., Stern, D., Spinrad, H., & Jannuzi, B. T. 2000, ApJ, 545, L85
Saito, T., Shimasaku, K., Okamura, S., Ouchi, M., Akiyama, M., Yoshida, M., & Ueda, Y. 2008, ApJ, 675, 1076
Sánchez, S. F. 2006, Astron. Nachr., 327, 850
Sandin, C., Becker, T., Roth, M. M., Gerssen, J., Monreal-Ibero, A., Böhm, P., & Weilbacher, P. 2010, A&A, 515, A35
Santos, M. R., Ellis, R. S., Kneib, J., Richard, J., & Kuijken, K. 2004, ApJ, 606, 683
Savage, R., et al. 2010, Proc. SPIE, 7733, 149
Sawicki, M., et al. 2008, ApJ, 687, 884
Schechter, P. 1976, ApJ, 203, 297
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Scoville, N., et al. 2007, ApJS, 172, 1
Severgnini, P., et al. 2005, A&A, 431, 87
Shoji, M., Jeong, D., & Komatsu, E. 2009, ApJ, 693, 1404
Smart, W. M., & Green, R. M. (ed.) 1977, Textbook on Spherical Astronomy (Cambridge: Cambridge Univ. Press)
Smith, D. J. B., & Jarvis, M. J. 2007, MNRAS, 378, L49
Smith, M. P., Hill, G. J., MacQueen, P. J., Altmann, W., Goertz, J. A., Good, J. M., Segura, P. R., & Wesley, G. L. 2008, Proc. SPIE, 7014, 233
Somerville, R. S., Lee, K., Ferguson, H. C., Gardner, J. P., Moustakas, L. A., & Giavalisco, M. 2004, ApJ, 600, L171
Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, ApJ, 519, 1
Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., & Giavalisco, M. 2000, ApJ, 532, 170
Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., & Giavalisco, M. 2003, ApJ, 592, 728
Steidel, C. C., Giavalisco, M., Dickinson, M., & Adelberger, K. L. 1996, AJ, 112, 352
Stern, D., Bunker, A., Spinrad, H., & Dey, A. 2000, ApJ, 537, 73
Stetson, P. B. 1990, PASP, 102, 932
Sullivan, M., Treyer, M. A., Ellis, R. S., Bridges, T. J., Milliard, B., & Donas, J. 2000, MNRAS, 312, 442
Sutherland, W., & Saunders, W. 1992, MNRAS, 259, 413
Teplitz, H. I., Collins, N. R., Gardner, J. P., Hill, R. S., & Rhodes, J. 2003, ApJ, 589, 704
Tilvi, V., et al. 2010, ApJ, 721, 1853
Treyer, M. A., Ellis, R. S., Milliard, B., Donas, J., & Bridges, T. J. 1998, MNRAS, 300, 303
Tufts, J. R., MacQueen, P. J., Smith, M. P., Segura, P. R., Hill, G. J., & Edmonston, R. D. 2008, Proc. SPIE, 7021, 9
Tull, R. G., MacQueen, P. J., Sneden, C., & Lambert, D. L. 1995, PASP, 107, 251
Turner, J. E. H., Miller, B. W., Beck, T. L., Song, I., Cooke, A. J., Seaman, R. L., & Valdés, F. G. 2006, New Astron. Rev., 49, 655
Valdes, F. 1992, in ASP Conf. Ser. 25, Astronomical Data Analysis Software and Systems I, ed. D. M. Worrall, C. Biemesderfer, & J. Barnes (San Francisco, CA: ASP), 417
van Breukelen, C., Jarvis, M. J., & Venemans, B. P. 2005, MNRAS, 359, 895
Venemans, B. P., et al. 2005, A&A, 431, 793
Watson, M. G., et al. 2009, A&A, 493, 339
Williams, R. E., et al. 1996, AJ, 112, 1335
Wirth, G. D., et al. 2004, AJ, 127, 3121
Xu, C., et al. 2007, AJ, 134, 169
Yang, Y., Zabludoff, A., Eisenstein, D., & Davé, R. 2010, ApJ, 719, 1654
Zacharias, N., Monet, D. G., Levine, S. E., Urban, S. E., Gaume, R., & Wycoff, G. L. 2005, VizieR Online Data Catalog, 1297, 0
Zanichelli, A., et al. 2005, PASP, 117, 1271