Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **493,** 952–972 (2020)
Advance Access publication 2020 February 3

doi:10.1093/mnras/staa276

The Grism Lens-Amplified Survey from Space (GLASS) – XIII. G800L optical spectra from the parallel fields

L. E. Abramson [1,2,3★] G. B. Brammer,[4,5] K. B. Schmidt,[6] T. Treu,[1] T. Morishita,[4] X. Wang,[1] B. Vulcani [7] and A. Henry[4]

[1]*UCLA, 430 Portola Plaza, Los Angeles, CA 90095, USA*
[2]*Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA*
[3]*Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA*
[4]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[5]*Cosmic Dawn Centre, Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, DK-2100 Copenhagen, Denmark*
[6]*Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*
[7]*INAF – Osservatorio Astronomico di Padova, Vicolo Osservatorio 5, I-35122 Padova, Italy*

Accepted 2020 January 27. Received 2019 December 23; in original form 2019 May 31

ABSTRACT

We present a catalogue of 22 755 objects with slitless, optical, *Hubble Space Telescope* (*HST*) spectroscopy from the Grism Lens-Amplified Survey from Space (GLASS). The data cover ∼220 sq. arcmin to 7-orbit (∼10 ks) depth in 20 parallel pointings of the Advanced Camera for Survey's G800L grism. The fields are located 6 arcmin away from 10 massive galaxy clusters in the HFF and CLASH footprints. 13 of the fields have ancillary *HST* imaging from these or other programs to facilitate a large number of applications, from studying metal distributions at $z \sim 0.5$, to quasars at $z \sim 4$, to the star formation histories of hundreds of galaxies in between. The spectroscopic catalogue has a median redshift of $\langle z \rangle = 0.60$ with a median uncertainty of $\Delta z/(1+z) \lesssim 2$ per cent at $F814W \lesssim 23$ AB. Robust continuum detections reach a magnitude fainter. The $5\,\sigma$ limiting line flux is $f_{\rm lim} \approx 5 \times 10^{-17}$ erg s^{-1} cm^{-2} and half of all sources have 50 per cent of pixels contaminated at $\lesssim$1 per cent. All sources have 1D and 2D spectra, line fluxes/uncertainties and identifications, redshift probability distributions, spectral models, and derived narrow-band emission-line maps from the Grism Redshift and Line Analysis tool (GRIZLI). We provide other basic sample characterizations, show data examples, and describe sources and potential investigations of interest. All data and products will be available online along with software to facilitate their use.

Key words: techniques: spectroscopic – catalogues – galaxies: evolution.

1 INTRODUCTION

The *Hubble Space Telescope* (*HST*) provides some of the best ultraviolet (UV) to near-infrared (NIR) imaging available. These data underpin most of our knowledge of the distant Universe. Ground-based spectra – from e.g. VVDS (Le Fèvre et al. 2005), MOSDEF (Kriek et al. 2015), and LEGA-C (van der Wel et al. 2016) – are critical, but the atmosphere limits continuum measurements to the most massive galaxies and blurs all spatial information. Using space-based data reduces these restrictions and yields a clearer picture of galaxy evolution.

HST's slitless grisms have played key roles here. They produce maps of every source at every wavelength at spatial scales inaccessible from Earth without adaptive optics. The two NIR grisms – G102 and G141 on WFC3[1] – have proven especially effective, with e.g. the WISP (Atek et al. 2010) and 3D-*HST* surveys (Momcheva et al. 2016), and our own Grism Lens-Amplified Survey from Space (GLASS; Schmidt et al. 2014; Treu et al. 2015) providing rest-optical $z > 1$ spectroscopy over scores to hundreds of sq. arcmin without the need for potentially biasing photometric pre-selection or slit masks. Paired with *HST* imaging, these data have extended our knowledge of the ages and star formation histories (SFHs) of galaxies at early cosmic times in ways comparable to ground-based results at $z \lesssim 1$ (e.g. Whitaker et al. 2012; Newman et al. 2014; Nelson et al. 2016; Wang et al. 2017, 2019; Abramson et al. 2018; Morishita et al. 2018, 2019).

Beyond their scientific utility, *HST*'s grisms play a pathfinding role: the astronomical community has decided that space-based, wide-field, slitless spectroscopy will be an increasingly large part

★ E-mail: labramson@carnegiescience.edu

[1]Wide Field Camera 3

Table 1. Basic survey information for the 20 GLASS ACS parallels. Field names reflect GLASS central pointing/IR grism (cluster) IDs and ROOT is the corresponding catalogue keyword (Appendix A2). PAs are 'PA_V3.'.

Field	ROOT	RA [J2000]	Dec. [J2000]	PA [deg]	$N_{\rm srcs}$	$N_{\rm infall}$[a]	G800L Exptime [s]	*HST* Imaging [program or PI][b, c]
Abell 2744	j0014m3023	0:13:53.9	– 30:22:51	323	976	26	10 076	HFF, Siana, Rodney
Abell 2744	j0014m3030	0:14:20.9	– 30:29:46	225	1066	16	10 076	HFF[d], Rodney[d]
Abell 370	j0240m0132	2:39:31.6	– 1:31:42	343	1028	31	10 076	–
Abell 370	j0240m0140	2:39:44.6	– 1:40:08	245	965	32	10 344	–
MACS0416	j0416m2402	4:15:44.7	– 24:02:04	337	808	15	10 076	CLASH, Rodney
MACS0416	j0416m2410	4:15:56.0	– 24:09:42	254	887	24	10 186	–
MACS0717	j0717p3750	7:17:16.3	37:49:53	10	1123	46	9927	HFF, CLASH, Siana, Ebeling
MACS0717	j0718p3747	7:18:01.6	37:47:26	110	1186	37	10 061	–
MACS0744	j0745p3922	7:45:08.6	39:22:16	194	853	37	9562	CLASH
MACS0744	j0745p3930	7:45:20.6	39:30:10	109	934	35	10 061	–
MACS1149	j1150p2218	11:49:40.3	22:18:02	215	1117	37	10 084	HFF, CLASH[d], Siana
MACS1149	j1150p2225	11:50:00.9	22:25:25	122	1090	47	9854	CLASH
RXJ1347	j1347m1140	13:47:19.7	– 11:40:11	13	1049	31	10 076	CLASH
RXJ1347	j1347m1147	13:47:09.4	– 11:47:36	293	954	20	10 344	–
MACS1423	j1424p2401	14:24:06.8	24:00:38	178	813	40	10 076	CLASH
MACS1423	j1424p2408	14:24:09.4	24:08:19	98	939	33	10 344	–
MACS2129	j2130m0736	21:29:31.5	– 7:35:40	58	1131	39	10 061	Rodney
MACS2129[e]	j2130m0742	21:29:54.4	– 7:41:45	136	3910	127	18 789	CLASH, Riess
RXJ2248	j2249m4433	22:49:18.0	– 44:32:37	143	906	16	9372	HFF, CLASH, Riess
RXJ2248	j2249m4438	22:48:46.0	– 44:37:39	223	1020	27	10 344	CLASH[d]

[a]All objects with $c|z - z_{\rm cl}|(1 + z_{\rm cl})^{-1} < 2000\ {\rm km\,s^{-1}}$ irrespective of z_Q where $z_{\rm cl}$ is the GLASS cluster redshift from table 1 of Treu et al. (2015).
[b]Covering band(s) other than *F*814*W* to support spectral energy distribution SED (spectral energy distribution) fitting.
[c]CLASH – Postman et al. (2012); http://archive.stsci.edu/proposal_search.php?mission=hst&id = 12067; HFF – Lotz et al. (2017); http://archive.stsci.edu/proposal_search.php?mission=hst&id = 13498; Ebeling – GO10420; http://archive.stsci.edu/proposal_search.php?mission=hst&id = 10420; Riess – GO13063; http://archive.stsci.edu/proposal_search.php?mission=hst&id = 13063; Rodney – GO13386; http://archive.stsci.edu/proposal_search.php?mission=hst&id = 13386; Siana – GO13389; http://archive.stsci.edu/proposal_search.php?mission=hst&id = 13389.
[d]One band.
[e]Grism data incorporated from GO12099 (PI Riess).

of its portfolio with the forthcoming operation of *JWST*,[2] *Euclid*, and *WFIRST*.[3] Hence, building intuition, applications, and tools to handle these data is prudent.

HST's less widely used optical disperser – the ACS[4] G800L grism – is powerful in this context. Covering $\lambda = 0.5$–$1.0\,\mu$m with ∼2.5 × WFC3's footprint, G800L extends space-based grism surveys to lower redshifts and more than doubles their area. The instrument's capability has been exploited by the GRAPES (GO9793, PI Malhotra; Pirzkal et al. 2004) and PEARS (GO10530, PI Malhotra; Straughn et al. 2008, 2009) surveys, and 3D-*HST*, with all results obtained prior to 3D-*HST* compiled by Kümmel et al. (2011).

Here, we describe GLASS' use of the G800L grism in parallel observing mode and present a source catalogue covering the 20 fields this survey comprises to 7-orbit depth. These data may be useful to anyone interested in a range of scientific questions – from infalling cluster galaxies to $z \gtrsim 4$ Lyman α emitters – and illustrate what could be done with higher signal-to-noise ratio (S/N) or spectral resolution data from future facilities.

Below, Section 2 describes data acquisition, field locations, reduction procedures, and output products. Section 3 describes the catalogue's redshift, magnitude, and contamination distributions, and highlights some notable sources for further study. Section 4 proposes some uses of the data, and Section 5 summarizes. We provide file descriptions and some useful pieces of code in the appendix. Throughout, we take $(H_0\ {\rm km^{-1}\ s\ Mpc}, \Omega_{\rm m}, \Omega_\Lambda) = (70, 0.3, 0.7)$ and quote AB magnitudes (Oke 1974). All data/products are available upon request and will be made public.

2 DATA

2.1 Acquisition

GLASS covers 10 $z \sim 0.4$–0.6 cluster sightlines: six from HFF[5] (Lotz et al. 2017) and four more from CLASH[6] (Postman et al. 2012). WFC3 was placed on the cluster centre at each sightline where gravitational lensing maximizes the chances of detecting high-redshift Lyman α emitters. G102 and G141 were exposed there in series for 5 + 2 orbits, respectively. This process was repeated in two visits at roughly orthogonal position angles (PAs) to facilitate source deblending. A 2 × 2 dither pattern with sub-pixel offsets was used to support image interlacing (Brammer et al. 2012), the format of the GLASS IR coadds.

In parallel with the NIR grism observations, G800L was exposed to a region ∼6 arcmin (∼2 Mpc) away from the cluster centres. Given *HST*'s rotation between visits, this strategy yielded 20, non-overlapping G800L survey fields; i.e. a spectroscopic data base covering ∼220 sq. arcmin to 7-orbit (∼10 ks) depth. Data were acquired between 2013 December 24 and 2015 January 5. Table 1 provides basic information.

A *F*814*W* pre-image was taken with each grism exposure (2 per orbit per PA). These data are necessary for source detection, spectral

[2]*James Webb Space Telescope.*
[3]*Wide-Field Infrared Survey Telescope.*
[4]Advanced Camera for Surveys
[5]Hubble Frontier Fields.
[6]Cluster Lensing and Supernova Survey with Hubble. Due to CLASH/HFF overlap, eight GLASS sightlines have CLASH coverage.

ID assignment, and wavelength calibration. Where possible, one PA was aligned with planned or actual ancillary imaging. This was achieved in 13 fields, where GLASS' *F*814*W* pre-images complement or supplement other *HST* photometry (Table 1). GLASS pre-images are the only *HST* images in the remainder (e.g. MACS0717 `j0718p3747`). We provide example spectrophotometry in Section 4.2 and catalogue matching instructions in Appendix B, but all analyses here *including redshift estimation* use no additional photometry unless stated.

2.2 Reduction

Although the GLASS NIR data (Schmidt et al. 2014; Treu et al. 2015) were reduced using a modified version of the 3D-*HST* pipeline (Brammer et al. 2012; Momcheva et al. 2016), the ACS parallel data were extracted using a PYTHON package with improved capabilities: GRIZLI – the Grism Redshift and Line Analysis tool (see Wang et al. 2017, 2019, Brammer 2019).[7] As with the original pipeline, the basic ingredients are two full-field ACS frames: one direct image and one containing the dispersed 2D spectra of all sources therein. The former serves as the reference frame to which the latter is anchored.

GRIZLI first identifies cosmic rays in the *F*814*W* and G800L exposures using standard settings with the ASTRODRIZZLE software (Gonzaga & et al. 2012), and fits and removes a two-dimensional master sky image from the grism exposures.[8] We refine the astrometric alignment including both a fine relative alignment between exposures and a global alignment to an absolute reference frame, here defined by the Gaia DR2 catalogue (Gaia Collaboration 2018). As there are no telescope offsets between the paired *F*814*W* direct and G800L grism exposures, the alignment of the former is applied directly to the latter. Finally, we combine the *F*814*W* direct exposures into a rectified mosaic with ASTRODRIZZLE at a final sampling of 0.03 arcsec pix^{-1} and generate a source catalogue and segmentation map from this mosaic using the SEP software package.[9]

The global contamination model for a given field is constructed in two passes. In the first, we assume that each object has a spectrum flat in units of f_λ flux density with a normalization set by its observed flux in the *F*814*W* direct image (integrated over the object's segmentation polygon). In the second pass, we step through objects in the catalogue sorted by increasing magnitude and compute a third-order polynomial fit to each spectrum after subtracting the contamination model from any neighbours (i.e. either the flat spectra of fainter sources or the polynomial models of brighter sources). This refinement is iterated three times.

In both passes, the *spatial* template of the dispersed 2D model spectra is taken from the observed *F*814*W* cut-out within an object's segmentation polygon. Note that this approach neglects morphological variation across the G800L bandpass; e.g. the PSF, continuum colour gradients, and morphological differences between line- and continuum-emitting regions. However, it provides a spatial model with much higher fidelity than any simple parametrized representation (e.g. Gaussian or Sérsic approximations).

The above process automatically yields a source and contamination model for each galaxy detected in the reference image. These can be used to 'decontaminate' source spectra – i.e. remove overlapping light from neighbouring sources – to provide more accurate line and continuum characterizations. This was critical in GLASS' crowded cluster G102 and G141 pointings, but the ACS parallels are sparse enough that source contamination is often negligible (Section 3). Note that no spectra are modelled or extracted for objects not detected in the reference frame. As such, the catalogue presented here is truncated at a *F*814*W* flux of $m_{814} = 26$ with various subsamples defined by brighter flux limits (Section 3).

There are two main differences between GRIZLI and the previous GLASS pipeline. First, GRIZLI automatically handles *HST* grism spectra taken at one or more telescope roll angles that translate into different spectral dispersion PAs. This is implemented in that the 2D spectral models are computed in the pixel space of each separate grism exposure. This is the space in which the field-dependent grism dispersion configuration is measured and defined (e.g. Kümmel et al. 2009). This approach allows for evaluating model fits in the space of the original detector pixels with a well-calibrated noise model. Recent works have demonstrated the strengths of this approach (see Morishita et al. 2018a,b; Wang et al. 2019), though note that the GLASS G800L data described here are observed at a single PA. This implementation causes the ACS data format to differ from the WFC3IR data (Schmidt et al. 2014; Treu et al. 2015) in that each field's 14 subframes are drizzle- and not interlace-combined. All weight maps account for this fact and contain the appropriate pixel-level covariances.

The second difference is that GRIZLI provides automatic redshift estimates of each source, along with its uncertainty, and 'risk factor' (Tanaka et al. 2018; proportional to $1/\chi^2$; Section 3). This is done by evaluating fits of galaxy templates similar to those used by the EAZY code (Brammer, van Dokkum & Coppi 2008). At the best redshift – typically where $\chi^2(z)$ is minimized but with an option to include a separate redshift prior – GRIZLI produces measurements and uncertainties of emission-line fluxes and equivalent widths, as well as continuum-subtracted 2D narrow-band maps around the emission-line species. We do not use all of the GRIZLI outputs here, but they may be of value to other investigations.

Note that all of the above processing is based exclusively on the G800L spectra; no ancillary photometry is employed. This caveat should be kept in mind for analyses based solely on our catalogue entries as it affects any quantity that may depend substantially on information outside the grism bandpass, such as stellar mass.

2.3 Products

The GLASS ACS GRIZLI reductions yield a suite of derived data products beyond the reference and spectral data frames:

(i) A master FITS catalogue containing IDs and summary metrics for all extracted sources. Spatial metrics are in pixels at the detection image plate scale of 0.03 arcsec pix^{-1}.

(ii) A '`.beams.`' FITS file containing 2D cut-outs straight from each grism exposure covering a given object, and the metadata necessary for using them to generate spectral models with GRIZLI.

(iii) A '`.stack.`' FITS table for each object containing its 2D spectrum, error map, contamination and source models, and *F*814*W*

[7] Available at https://github.com/gbrammer/grizli.

[8] We fit smooth functions to the G800L configuration files provided at http://www.stsci.edu/hst/instrumentation/acs/performance/prism--grism/wfc-g800l, which are in turn available at https://s3.amazonaws.com/grizli/CONF/ACS.WFC.sky.tar.gz.

[9] SEP is a PYTHON implementation of the SEXTRACTOR software (Bertin & Arnouts 1996) designed to exactly replicate its functionality; https://github.com/kbarbary/sep/

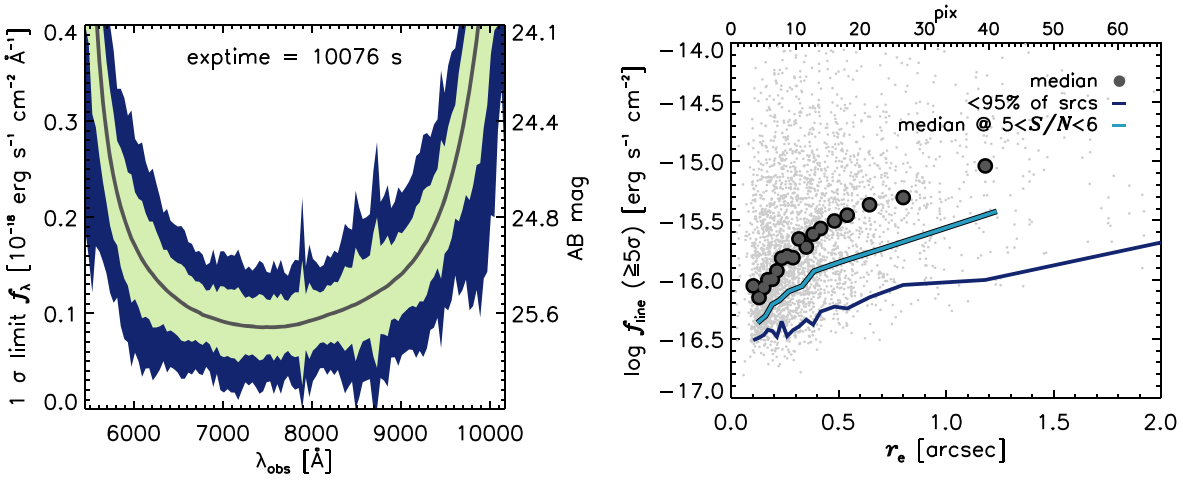

Figure 1. Left: GLASS G800L sensitivity based on a 1 arcsec spatial extraction aperture. This plot describes Abell 2744 (at PA 323) but is typical of all pointings. Thresholds are the median $1\,\sigma$ f_λ limit (grey) and that value's $1\,\sigma$ and $2\,\sigma$ scatter (shaded). Right: Line fluxes as a function of observed size in the 'high-*S/N* + lines' sample (Section 3). 95 per cent of lines have $\log f/\mathrm{erg\,s^{-1}\,cm^{-2}} \gtrsim -16.5$ for point sources, rising to $\log f \simeq -16.0$ in those with half-light radii of $r_e \sim 1.2$ arcsec (dark blue). Similarly, the median $\sim 5\,\sigma$ line flux limit rises from $\log f \simeq -16.4$ to $\log f \simeq -15.4$ across that size range (turquoise).

direct image. All are rectified such that the spectral dispersion aligns with the data's *x*-axis. Unlike in the `.beams.` files, this rectification (and drizzling) necessitates a resampling of the original detector pixels to 0.1 arcsec pix^{-1}. This file's header contains a wavelength solution, exposure time, source location, and ancillary information.

(iv) A '`.full.`' FITS table for each object containing redshift solutions, the covariance matrix and $\chi^2(z)$ of the template SED fits, the best-fitting high-resolution template SED with and without emission lines, the rectified, drizzled *F*814*W* direct image and error map, and rectified, drizzled narrow-band images at wavelengths surrounding possible emissions lines (Section 4.2). These stamps comprise a line map, continuum estimate, contamination, and error array for each line listed in the FITS header. Spatial sampling is again 0.1 arcsec pix^{-1}. The header also contains a wavelength solution, exposure time, source location, line fluxes, and ancillary information.

(v) A '`.1D.`' FITS table for each object comprising the wavelength array, contamination-subtracted source spectrum ($e^-\,s^{-1}$), sensitivity curve [$e^-\,s^{-1}/(\mathrm{erg\,s^{-1}\,cm^{-2}\,\AA^{-1}})$], RMS error ($e^-\,s^{-1}$), and line emission and continuum models. These extractions use an optimal weighting (Horne 1986) based on an object's spatial profile derived from the direct image itself as described above.

All of the above files are detailed in Appendix A, are available upon request, and will be published on GLASS' MAST website.[10] Useful software tools for selecting sources and matching to external catalogues will also be released (Appendix B).

3 SAMPLE CHARACTERISTICS

3.1 Counts, quality levels, and depth

GLASS' G800L spectral data base contains 22 755 objects extracted to a limiting magnitude of $m_{814} = 26$ – roughly 1000 sources per pointing. For the purposes of this characterization, we split these into three, somewhat overlapping, quality-based subsets defined by continuum and line S/N. These are:

[10] https://archive.stsci.edu/prepds/glass/.

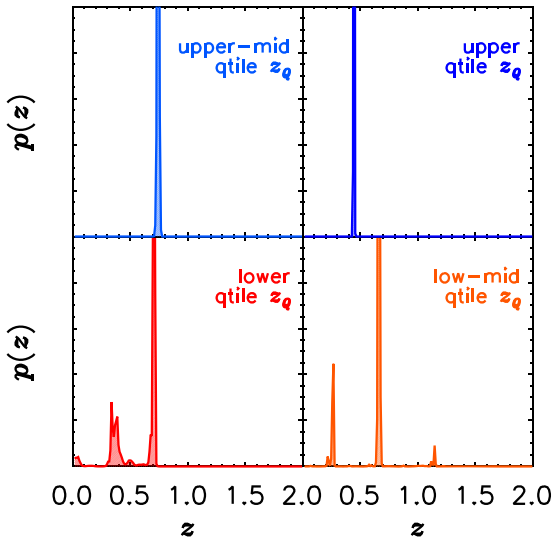

Figure 2. Random examples of redshift estimation quality in the four quartiles of the redshift quality parameter, z_Q based on the $m_{814} < 24$ *or* $S/N \geq 5$ in any optical strong line quality cut. Variation in the PDFs is noticeable.

(i) All non-point sources with $m_{814} \leq 24$ *or* $S/N_{\mathrm{line}} \geq 5$ ('high-S/N + lines'; $N = 4636$);

(ii) The subset of such galaxies below that subsample's median GRIZLI z_Q ('low risk'; $N = 2319$);

(iii) Everything not in (ii) ('high risk'; $N = 20436$).

Above, S/N_{line} refers to the *S/N* in any of the following features as identified by GRIZLI: Hα, Hβ, [O II], [O III], [S II], [Mg II], or Lyα. The magnitude cut corresponds roughly to the inferred $5\,\sigma$ continuum sensitivity – $\sim 5 \times 10^{-19}\,\mathrm{erg\,s^{-1}\,cm^{-2}\,\AA^{-1}}$ (Fig. 1) – as measured in a 1 arcsec spatial aperture. 'z_Q' is the GRIZLI-characterized redshift quality/risk (lower is better), which correlates well with redshift errors estimated as $\Delta z \equiv 84\mathrm{th}-16\mathrm{th}\ P(z)$ percentile. Fig. 2 shows $P(z)$ PDFs for each quartile of the z_Q distribution.

There are low risk sources with no detected emission lines. These are ∼1 mag brighter on average than the full low risk sample –

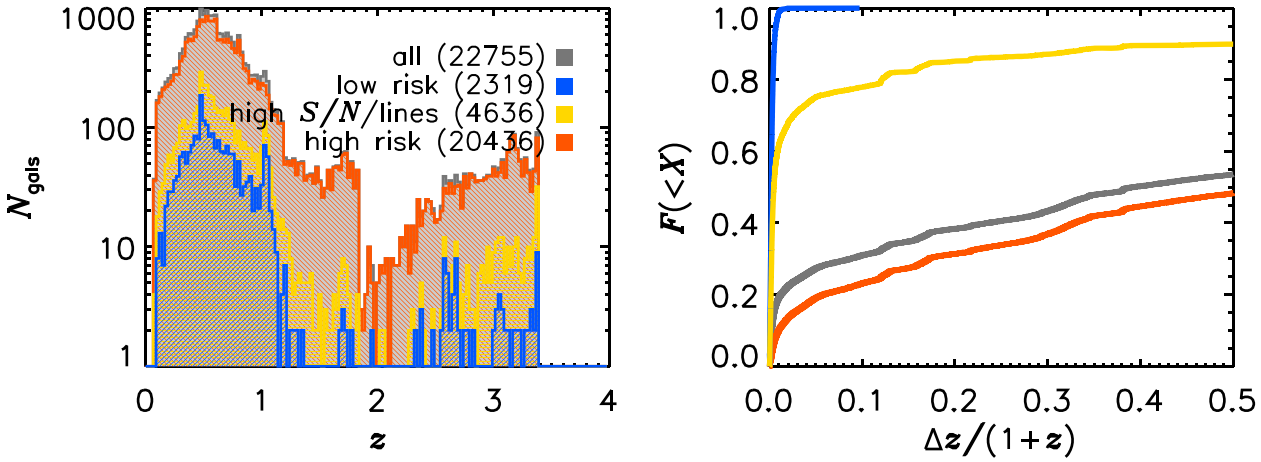

Figure 3. The redshift (left) and redshift uncertainty (right) distributions for the GLASS G800L sample. Here and in all following plots, the 'high' and 'low risk' samples are plotted in red and blue, respectively. These are defined based on the median z_Q of all sources with $m_{814} < 24$ *or* $S/N \geq 5$, whose distributions are plotted in gold. The median redshift for those sources is $\langle z \rangle \simeq 0.60$, with the low risk sample at $\langle z \rangle \simeq 0.58$. Median formal uncertainties rise from < 1 per cent for low risk to nearly 50 per cent for high risk objects. Half of the full sample has $\Delta z/(1+z) \lesssim 0.35$. These uncertainties would be reduced by the addition of archival photometry where available, especially at $m_{814} \gtrsim 23$ (Table 1, Fig. 5).

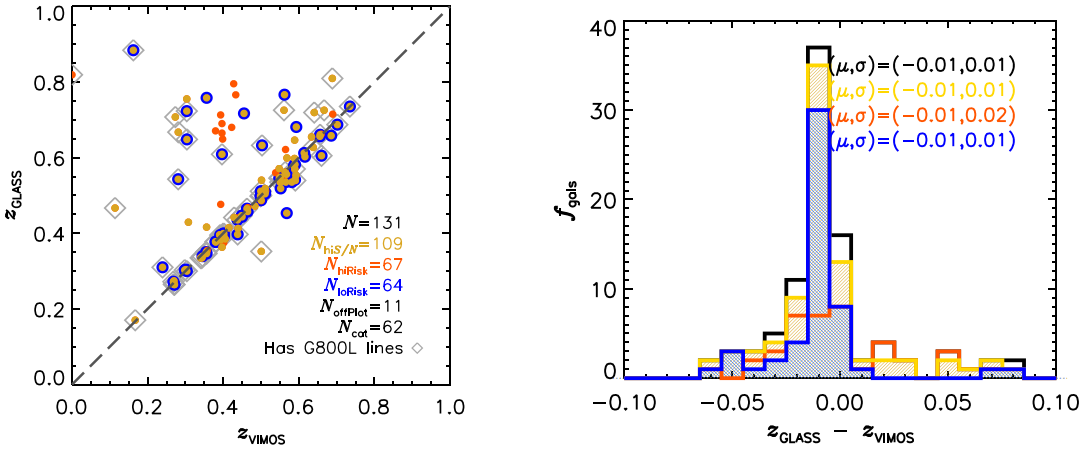

Figure 4. Left: GLASS G800L GRIZLI redshifts compared to those from VLT VIMOS as taken by various authors (see the text). Agreement is quite good at $z < 1$, with the global sample biased to lower redshifts by just $\mu(\Delta z) = -0.01$ with a scatter, σ_z, of the same order (right-hand panel). A similar comparison with HFF photo-zs is presented in the appendix. There is one galaxy at $z_{\rm VIMOS} = 3.1$ that is at similarly high $z_{\rm GLASS} = 2.9$, but all other off-plot sources are catastrophic outliers. Based on this comparison, the catastrophic outlier fraction is $62/131 = 0.47 \pm 0.09$ and $21/56 = 0.38 \pm 0.13$ for all sources and those with GLASS-detected emission lines, respectively. 'Catastrophic' is defined as $\Delta z > 3\,\sigma_z$ for the full sample (black histogram at right).

$m_{814} = 22.55$ versus 23.50, respectively – corresponding to the mean increase in continuum quality needed to achieve the redshift precision otherwise supplied by lines. Of course, 'low risk' does not imply *no* risk, and the spectra of any (set of) high-interest object(s) should be vetted depending on usage.

Fig. 3 shows the sources' inferred redshift distribution (left) and redshift error distribution (right; $\equiv \Delta z/(1+z)$). Note that, due to the inclusion of template lines spanning [O II] λ3727 and redward, there is an artefact in the redshift distribution at $z \sim 1.7$ where that feature drops out of the G800L bandpass. The sample's median redshift is $z \simeq 0.60$, irrespective of quality cuts. Compared to the high risk subsample, the enhanced quality of the other objects' redshifts is clear, with the ∼2200 low risk systems having estimates formally precise to <1 per cent. Visual inspection of high-quality objects suggests these redshifts are systematically accurate to within the $2\,\sigma$ formal uncertainties. The $R \sim 40$ spectral resolution makes more quantitative cross-checks difficult, but comparisons to much higher resolution VLT VIMOS spectra for 131 common objects suggest agreement is at the $\Delta z = 0.01$ level in the median, with a scatter of $\sigma_z = 0.01$ below $z = 1$ (Fig. 4; data from Grillo et al. 2016; Caminha et al. 2017; Karman et al. 2017; Monna et al. 2017).[11] Fig. 5 illustrates that this quality is typical at $m_{814} \lesssim 22$ ($S/N \sim 20$ per spectral pixel), and true for at least 50 per cent of objects up to a magnitude fainter.

The left-hand panel of Fig. 6 shows the $F814W$ apparent magnitude distribution for the above samples. Half of the full sample is brighter than $m_{814} \approx 25$, though it is dominated by insecure redshifts at that limit. The distribution of 'high-S/N + lines' objects cuts off sharply at $m_{814} = 24$ by construction (see definition above), though counts in the low risk subsample remain relatively flat beyond this point. Comparing these histograms amplifies Fig. 5's results, suggesting that robust automated redshift estimations probably require ancillary SED information or spectral lines beyond m_{814}

[11] Available at https://sites.google.com/site/vltclashpublic/data-release.

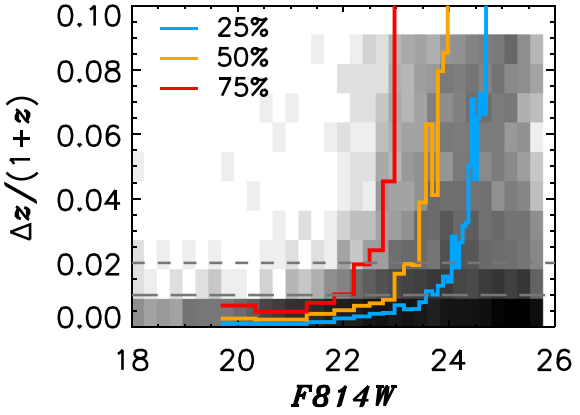

Figure 5. GLASS G800L GRIZLI redshifts are precise to < 1 per cent at $m_{814} \lesssim 22$ for most sources but degrade rapidly by $m_{814} \sim 24$, where ancillary SED coverage or strong emission lines are needed to ensure fidelity. The coloured lines show redshift uncertainty quartiles as a function of flux.

~ 22 in these data, so catalogue values should be used cautiously. Inverted, this statement provides a selection criterion, such that GLASS sources with $m_{814} > 22$ and tight redshift distributions are probably high-EW line emitters.

The right-hand panel of Fig. 6 presents the samples' rough stellar mass (M_*) limit. For the low risk objects, this extends from $\log M_*/\mathrm{M}_\odot \sim 7$ at $z \sim 0.6$ to $\log M_*/\mathrm{M}_\odot \sim 9$ at $z \sim 1$ – similar to the main GLASS catalogue limits (Morishita et al. 2017). We caution, however, that these estimates are based only on the information in the G800L spectrum and therefore may change significantly when ancillary photometry is included in the inference (Section 4.2).

Of course, all of the above depends fundamentally on the grism S/N. The top panel of Fig. 7 shows the median per-pixel S/N for the full sample (grey) and the baseline quality cut ($m_{814} \leq 24$ or any line detected at $>5\,\sigma$). In the case of the former, half of the spectra have $\langle \mathrm{S/N} \rangle > 2\,\mathrm{pix}^{-1}$, and 25 per cent have at least twice that. The high-confidence sample has a median and upper quartile of $\sim$7 and $\sim$20, respectively, more than triple the S/N of the general population. We therefore recommend basing most general analyses on at least our 'high-S/N + lines' quality cut, and only proceeding beyond this after more rigorous exploration of the data. Fortunately, Fig. 7's middle and bottom rows reveal that contamination is often not an impediment: half the sample has >50 per cent of pixels free of any contamination. Alternatively, half of sources have $\lesssim$20 per cent of pixels contaminated at >20 per cent. Contamination (by either metric) is slightly higher for the higher quality data subset. This is likely due in part to their typically $\sim$30 per cent larger half-light radii (Fig. 8), which increases the probability of spectral collisions.

3.2 Higher level outputs: spectral line maps

The GLASS G800L catalogue contains automated spectral line identifications. The left-hand panel of Fig. 9 shows the distribution of galaxy fractional counts as a function of the number of lines detected at $\mathrm{S/N_{line}} \geq 5$ and quality cut. The corresponding line IDs for the low risk sample are shown at right. Equation (1) characterizes the magnitude and wavelength dependence of the fifth percentile line flux for these objects:

$$\log f_{0.05} = -9.78 - 12.81\,\Lambda + 6.28\,\Lambda^2 - 0.04\,M - 0.06\,M\Lambda + 0.01\,M^2, \quad (1)$$

where $\Lambda \equiv \lambda/7200$ Å and $M \equiv m_{814} - 23$ for $5000 \leq \lambda/\text{Å} \leq 10\,000$ and $20 \leq m_{814} \leq 26$.

As expected, the higher quality samples typically have one to two well detected lines while the general population has zero. Of detected lines, the most common are [O III], [O II], and the strong Balmer lines. However, even at moderate $\mathrm{S/N_{line}}$, these IDs should be treated with care: visual inspection suggests $\sim$20–40 per cent of sources at fixed $N_{\mathrm{lines}} \in [1, 3]$ may reflect spurious characterizations. Furthermore, because [Mg II] is not explicitly included in the fitting templates – fluxes are simply measured at the relevant wavelength given the redshift inferred from the continuum + [O II] λ3727 and redder emission lines – objects identified as [Mg II] emitters mainly reflect reduction errors or noise spikes. Hence, these sources should not be taken *prima facie* as e.g. AGN or Lyman continuum leaker candidates (Finley et al. 2017; Henry et al. 2018). Appendix B contains example code to extract high confidence line emitters.

In addition to the total fluxes and errors for every line in every source, GRIZLI outputs 2D spectral cut-outs in the relevant wavelength regions. Fig. 10 shows examples of these line maps for two high-quality sources with their direct images and full 1D and 2D spectra. The source on the left is a typical $z \sim 0.2$ star-forming galaxy with prominent and extended Hα. The source on the right is a high-EW [O III] emitter at $z \sim 0.65$ with quite different oxygen

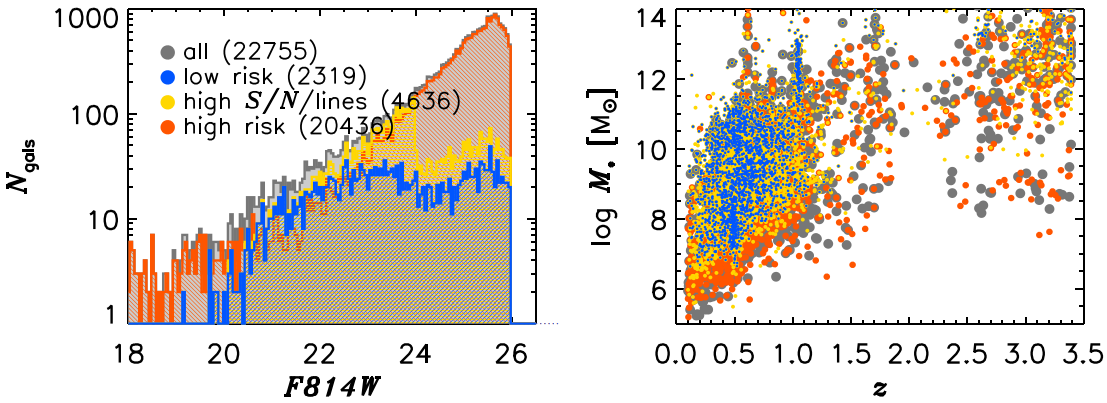

Figure 6. *F*814*W* magnitude distributions (left) and estimated stellar mass coverage as a function of redshift (right) for the full and quality-defined samples. Colour coding is as in Fig. 3. Sources in the low risk sample fainter than $m_{814} = 22$ tend to be high-EW line emitters. Lines therefore form the basis of redshifts in that regime, with the continuum shape playing a larger role for brighter objects. High-quality sources reach $\log M_*/\mathrm{M}_\odot \sim 7$ and $\sim$9 at $z = 0.60$ and 1.0, respectively, though these numbers are based on the relatively small G800L bandpass.

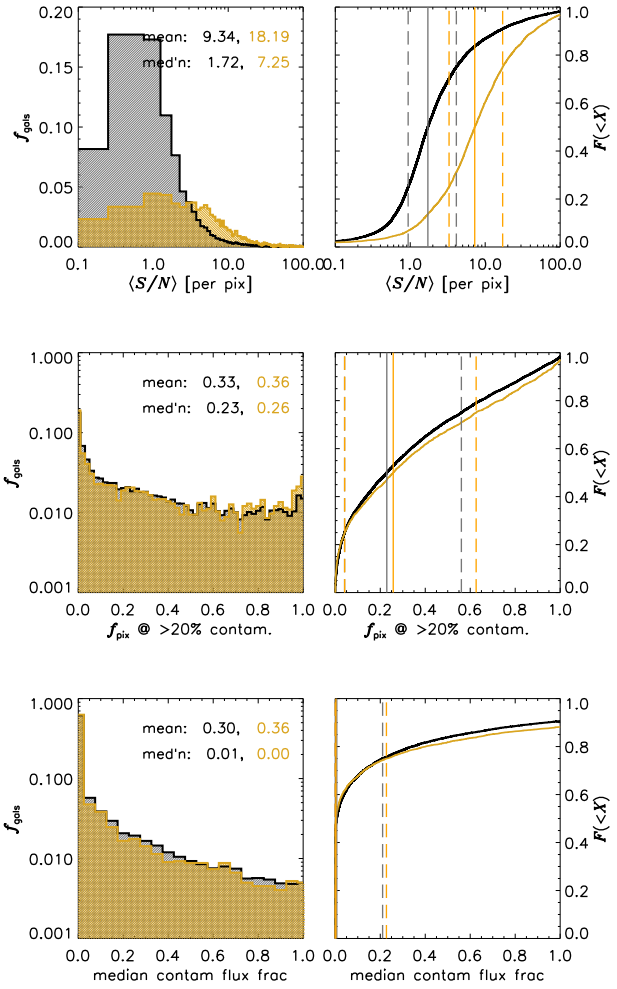

Figure 7. Top: median S/N per spectral pixel from the optimal 1D extractions for all galaxies (grey) and those in the high-S/N + lines subsample (gold). The latter has about 3× the S/N as the former. Middle: the fraction of source 2D spectral pixels whose flux is contaminated by overlapping traces to at least 20 per cent. Unlike in the WFC3 data, fully half of G800L GLASS spectra have fewer than 20 per cent substantially contaminated pixels. Bottom: median contaminating flux per pixel per source; another metric of contamination. Cumulative distributions for all quantities are shown at right: 75 per cent of objects have a median per-pixel contamination of <20 per cent; 75 per cent of high-S/N + lines objects have typical continuum S/N > 4 pix^{-1}.

and H β morphologies, perhaps suggesting the presence of an active galactic nucleus (though H β is poorly resolved). We discuss these maps further in Section 4.3.

3.3 Notable sources

Slitless spectroscopy avoids the need for photometric pre-selection. As such, it maximizes the chance of serendipitously capturing interesting sources. Fig. 11 shows two of these in the GLASS G800L footprint that are also known redshift misclassifications (due to the modelling of lines only at and blueward of [O II] λ3727 in the template spectra; see Section 3.1). As revealed by its prominent Ly α, carbon lines, and point-like morphology, the source at left is a $z \sim 4$ quasar. Of course, GLASS data simultaneously provide a redshift survey along this QSO's line of sight, and thus immediately suggest 61 (164) foreground candidates with a $\rho \leq 150$ (300) kpc

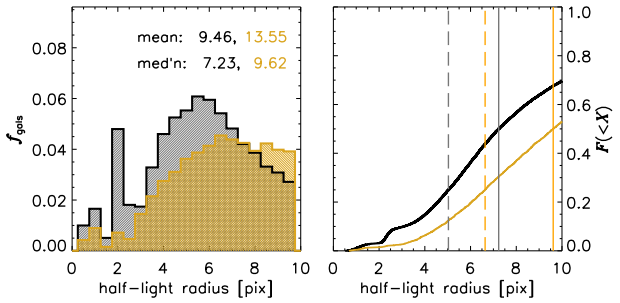

Figure 8. Sample size distribution. The high-S/N + lines sample comprises larger sources on average, contributing to their slightly higher contamination levels.

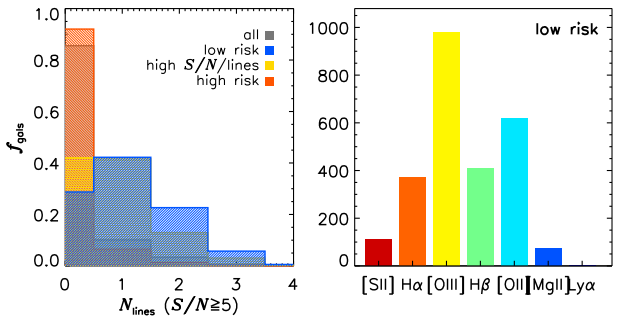

Figure 9. Line detection statistics for all galaxies and quality subsamples split as in Fig. 3. Line IDs from the low risk sample cut (blue histogram at left) are shown in the right-hand panel. All lines have integrated S/N $\geq$ 5. Unsurprisingly, higher quality subsamples tend to have more high-confidence line detections, with the mode in the full and high-risk samples being zero lines. Of detections, the most common is [O III]. Note that low risk sources with no detected lines are on average $\sim$1 mag brighter than the full low risk sample ($m_{814} = 22.55$ versus $m_{814} = 23.50$, respectively), corresponding intuitively to the fact that their continuua must be better measured to ensure such a high-quality redshift solution. Note also that, since the line is not explicitly modelled, sources identified as [Mg II] emitters typically reflect reduction errors or noise spikes.

impact parameter from the quasar suitable for characterizing H I and low-ionization species such as Mg II (Chen et al. 2010; Prochaska et al. 2011; Rudie et al. 2012; Johnson, Chen & Mulchaey 2015). As such, this source could support higher resolution spectroscopic follow-up to learn about the circumgalactic media (CGM) of scores of galaxies (though 30-m class facilities will be needed for $R \sim$ 40 000 studies). Combined with their ISM maps (Fig. 10), the CGM enrichment levels/temperatures/ionization states derived from such follow-up could provide powerful empirical constraints on metal transport, and therefore evolutionary models (e.g. Davé, Finlator & Oppenheimer 2011; Peeples et al. 2014; Muratov et al. 2015; Davé et al. 2017). Both of these studies are enabled by slitless spectroscopy.

The right-hand panel of Fig. 11 shows another $z \sim 4$ source. As opposed to the QSO, this object's lack of Ly α emission, strong Ly α break, and extended morphology reveal it to be a Lyman break galaxy (LBG). Large samples of such objects at these redshifts exist (e.g. Bouwens et al. 2015), but spectroscopy remains difficult. The low backgrounds in space and avoidance of slit losses, however, make *HST* efficient at spectroscopically identifying LBGs. Indeed, at $m_{814} = 23.8$, this object shows that GLASS' data reach continuum levels comparable to those from surveys on 10-m class ground-based

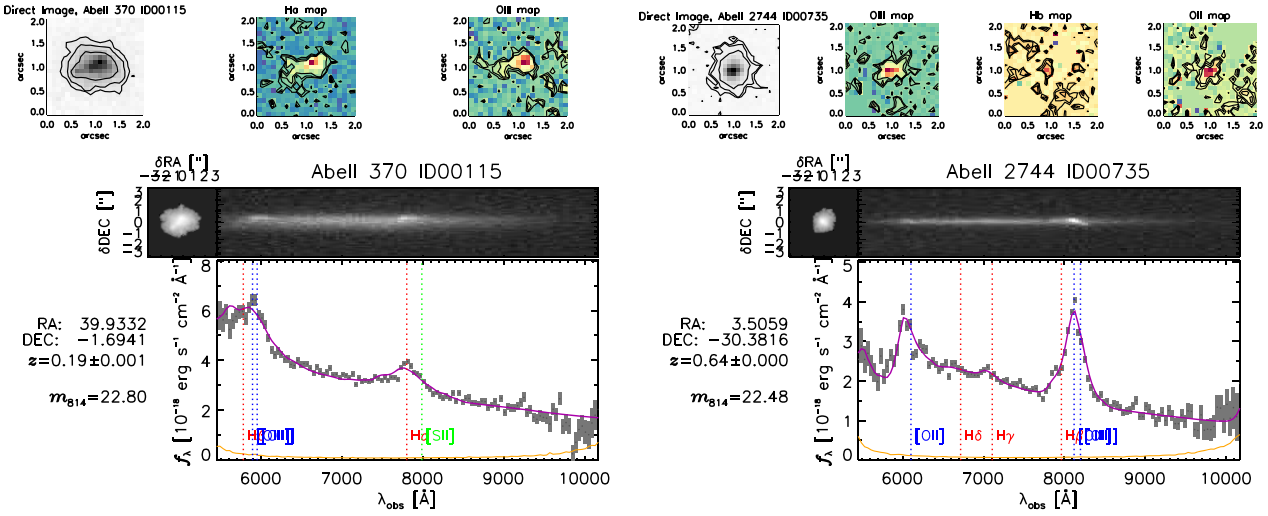

Figure 10. Top: line maps for two low risk subsample sources. Contours show 70, 80, 90 per cent flux contours. Bottom: the sources' direct images and 2D spectra above their 1D optimal extractions with GRIZLI model (purple) and line IDs overlaid (blue = oxygen, red = hydrogen, green = sulphur). The source on the right shows a typical redshift systematic offset for these sources. Ancillary source details are printed to the left of both spectra (coordinates, redshift, flux).

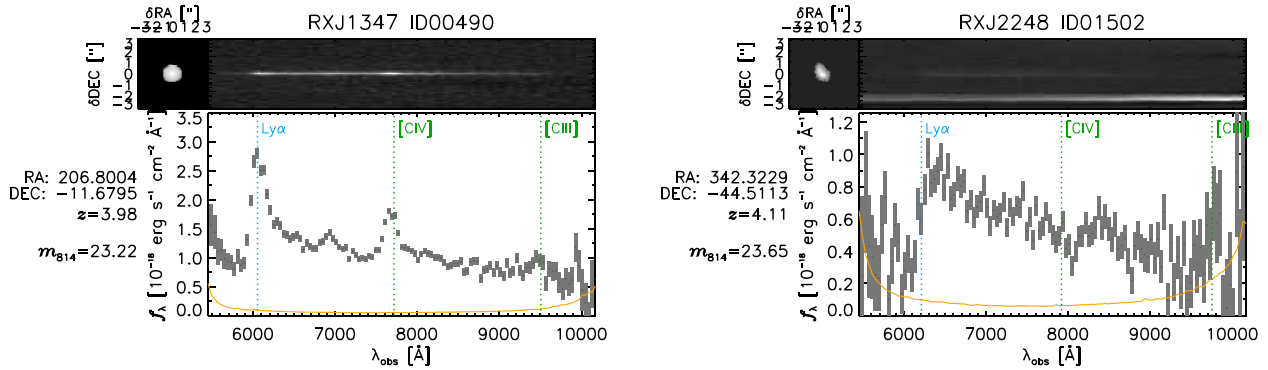

Figure 11. Examples of two known redshift misclassifications. Both the QSO (left) and LBG (right) are at $z \sim 4$ but were initially placed at $z = 0.89$ and $z = 0.70$, respectively. More examples of the ∼300 bright, $S/N \geq 5$ [O III] emitters in the GLASS G800L data base are shown in Appendix D.

telescopes for similar integration times, albeit at lower spectral resolution (Steidel et al. 1999; Shapley et al. 2003).

4 DISCUSSION AND FURTHER APPLICATIONS

The GLASS G800L data base will support a range of investigations. We note three potentially fruitful avenues below beyond the QSO and LBG follow-up studies just discussed.

4.1 Infalling cluster members

As shown in Table 1, each pointing in the catalogue contains ∼20–40 objects within 2000 km s^{-1} of the HFF or CLASH cluster redshifts. These are likely galaxies either falling into the cluster potential for the first time, or 'splashing back' after one or more crossings through the clusters (More, Diemer & Kravtsov 2015; Baxter et al. 2017). Both cases present opportunities to study the forces affecting galaxies in the densest environments in the Universe. We encourage anyone interested in ram pressure stripping (Gunn & Gott 1972) or pre-processing (Zabludoff & Mulchaey 1998) to explore these objects. Analogous analyses to Vulcani et al. (2015, 2016, 2017, see Section 4.3) but based on [O III] may, for example, prove powerful.

4.2 Joint analyses with *HST* photometry

Over half of the sources in the data base lie in regions covered by CLASH, HFF, or other *HST* imaging. Given the limited bandpass of the G800L grism, incorporating these data into any further SED fitting is valuable, especially for analyses that rely on accurate stellar mass or SFR inferences. Fig. 12 shows the GLASS spectra overlaid on HFF photometry for three of the 859 common GLASS/Shipley et al. (2018) sources in the Abell 2744 parallel field. For low risk sources, fluxing between the grism data and HFF photometry is good to $\Delta m_{814} = 0.1$ mag (such that HFF fluxes are brighter) with a scatter of 0.25 mag about that offset.[12] The grism-derived continuum SED model extrapolated well outside the G800L bandpass can be predictive to similar levels – low risk median $m_{GRIZLI} - m_{\rm HFF} =$ [0.19, 0.25, 0.25, 0.25] in [F105,125,140,160W] with >0.5 mag

[12]This measurement accounts for zero-point and Milky Way extinction, but not aperture corrections, to the HFF data.

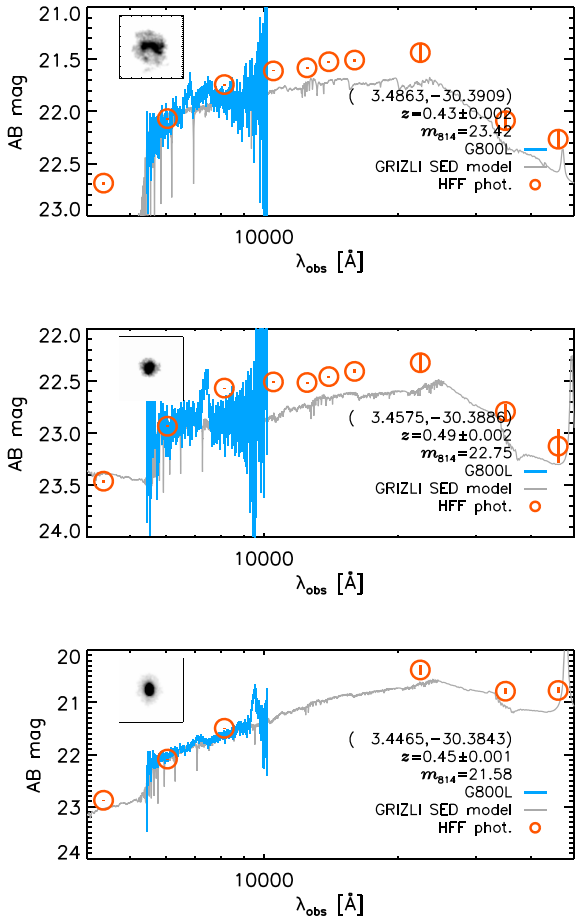

Figure 12. Examples of GLASS G800L spectra (blue) and GRIZLI inferred continuum SEDs (grey) shown with matched source photometry from the HFF campaign (Shipley et al. 2018, orange circles). Some inferences agree remarkably well with photometry lying well outside the grism bandpass despite its limited coverage; others are less consistent. Such offsets – typically at the ∼0.25 mag level (Section 4.2) – will affect stellar mass and SFR estimates. However, the consistent location of observed and inferred features (e.g. the 1.6 μm bump) support the accuracy of these sources' grism-determined redshifts (see also Fig. C1). These data are available to anyone wishing to perform joint spectrophotometric SED fitting on the 13 fields where *HST* photometry exist (Table 1). More examples of low risk objects in the HFF footprints are shown in Appendix D.

scatters around these offsets – but larger disagreements in the IR and UV can obviously occur.

Moreover, the combination of low-resolution grism spectroscopy covering, e.g. the Balmer or 4000 Å breaks with broad-band photometry to the red and blue is now being used to great effect in inferring the SFHs of galaxies at $0.3 < z < 3$, not just observed quantities (Dressler et al. 2016; Abramson et al. 2018; Dressler, Kelson & Abramson 2018; Morishita et al. 2018b). Fig. 12 shows the GLASS data will support similar analyses at $z \sim 0.4$–1.25. Critically, given their high spatial resolution and low contamination, these data could support spatially resolved SED analyses to constrain individual galaxies' joint mass and structural evolution over at least the past ∼1–2 Gyr. These empirical inferences can be compared to simulations to provide direct, longitudinal tests of numerical physical prescriptions, not just bulk predictions for the galaxy population at large. Abramson et al. (2018) performed such an analysis, but were limited to just four systems due to the high contamination rates in GLASS' central WFC3 pointings. The G800L data do not suffer from this issue, so spatially resolved spectrophotometric SFH reconstructions based on functional [e.g. by PYSPECFIT (Newman et al. 2014), or other means (Iyer & Gawiser 2017)] or free-form inferential techniques (Pacifici et al. 2012; Kelson et al. 2014; Leja et al. 2017; Morishita et al. 2018b) should yield a valuable data base of hundreds of high-quality mass, SFR, and structural histories over large ranges in *observed* mass, SFR, and structural parameters.

4.3 Emission-line mapping

As mentioned in Sections 1 and 3.2, Fig. 10 illustrates one of the key advantages of *HST* and future space-based slitless spectroscopy: the automatic production of spatially resolved spectral features mapped at the diffraction limit – 600 pc at $z = 0.6$ (FWHM = 0.09 arcsec). These maps can be used, for example, to infer galaxy metallicity distributions and even outflow patterns at sub-kpc scales to challenge numerical feedback and star formation models in new ways. For example, Vulcani et al. (2015, 2016, 2017) used the G102 Hα maps from GLASS' central pointings to study the connection between galaxy stellar and gas morphology, and associate this with various features of the local and global environment in clusters. Jones et al. (2015) and Wang et al. (2017) used oxygen and Hβ to produce gas-phase metallicity maps and gradients for sources at $z \sim 1$–3 and unprecedentedly low stellar masses ($\log M_* \lesssim 8$). Wang et al. (2019) identified two such sources with steeply positive gradients ([O/H] $\propto r$) and used simple models to produce outflow maps and mass loading factors as a function of underlying stellar mass density, showing that individual systems are not described well by *only* energy- or momentum-driven wind models. The GLASS G800L data will support similar studies based on [O II] through Hα lines at $z \lesssim 1$, both outside and in the infall regions around massive clusters. Further two-dimensional explorations of the ISM will no doubt be fruitful.

5 SUMMARY

We present a catalogue of 22 755 objects with ACS G800L slitless spectroscopy from the 20 GLASS parallel fields (Table 1). The catalogue extends to $m_{814} = 26$ with uniform 7-orbit (∼10 ks) coverage over ∼220 sq. arcmin. Sources have a median redshift of $z = 0.60$ with median uncertainties of $\Delta z/(1+z) \lesssim 0.02$ at $m_{814} \leq 23$ (Figs 3 and 5) and are typically contaminated at the 0 per cent–20 per cent level (Fig. 7). About a quarter of the sample either has continuum flux detected at $\geq 5\sigma$ ($m_{814} \lesssim 24$; $f_\lambda \approx 5 \times 10^{-19}$ erg s^{-1} cm^{-2} Å^{-1}) or S/N > 5 in at least one spectral line ($f_{\rm lim} \approx 5 \times 10^{-17}$ erg s^{-1} cm^{-2}) such that median redshift errors are <1 per cent. Incorporating photometry for the 13 fields that overlap with extant *HST* imaging from CLASH, HFF, or other programs will allow redshifts to be obtained to much greater depth and support a rich variety of spectrophotometric studies into galaxy SFHs at $z \sim 0.5$–1.5 (Section 4.2).

The full catalogue also contains full 2D spectra, contamination, and source models for each object, along with automated line identifications, fluxes, uncertainties, and 2D spatially resolved maps produced by GRIZLI. Optimal 1D extractions are also provided, as are full UV–sub-mm best-fitting SEDs and redshift PDFs. All data and derived products are available on request, and will be published on MAST along with code for performing various basic operations – which we also give here in Appendix B.

These data and products will support a wide range of investigations – from mapping galaxy outflows at $z < 0.8$ (Sections 3.2 and 4.3) to identifying QSOs and LBGs at $z \sim 4$ (Section 3.3) – and serve as useful intuition builders as the field prepares for the increasing ubiquity of slitless spectroscopy in the eras of *JWST* and *WFIRST*.

Facilities: HST ACS.

Software: IDL (Coyote libraries; http://www.idlcoyote.com/), PYTHON (GRIZLI).

ACKNOWLEDGEMENTS

LEA thanks Camilla Pacifici, Dan Kelson, and the anonymous reviewer. Data from the Grism Lens-Amplified Survey from Space (GLASS) derive from observations made with the National Aeronautics and Space Administration (NASA)/European Space Agency *Hubble Space Telescope*. GLASS is supported by NASA through grant number HST-GO-13459 from the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., under contract NAS 5-26555. This work uses data and catalogue products from HFF-DeepSpace, funded by the National Science Foundation and STScI. This work also uses catalogues derived from data from Very Large Telescope programmes 186.A-0798, 094.A-0115(B), 094.A-0525(A), 60.A-9345(A), and 095.A-0653(A).

REFERENCES

Abramson L. E. et al., 2018, AJ, 156, 29
Atek H. et al., 2010, ApJ, 723, 104
Baxter E. et al., 2017, ApJ, 841, 18
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bouwens R. J. et al., 2015, ApJ, 803, 34
Brammer G., 2019, Astrophysics Source Code Library, record, ascl:1905.001
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Brammer G. B. et al., 2012, ApJS, 200, 13
Caminha G. B. et al., 2017, A&A, 600, A90
Chen H.-W., Helsby J. E., Gauthier J.-R., Shectman S. A., Thompson I. B., Tinker J. L., 2010, ApJ, 714, 1521
Davé R., Finlator K., Oppenheimer B. D., 2011, MNRAS, 416, 1354
Davé R., Rafieferantsoa M. H., Thompson R. J., Hopkins P. F., 2017, MNRAS, 467, 115
Dressler A., Kelson D. D., Abramson L. E., 2018, ApJ, 869, 152
Dressler A. et al., 2016, ApJ, 833, 251
Finley H. et al., 2017, A&A, 608, A7
Gaia Collaboration, 2018, A&A, 616, A14
Gonzaga S., Hack W., Fruchter A., Mack J., 2012, eds., The DrizzlePac Handbook, Baltimore, STScI
Grillo C. et al., 2016, ApJ, 822, 78
Gunn J. E., Gott J. R. III, 1972, ApJ, 176, 1
Henry A., Berg D. A., Scarlata C., Verhamme A., Erb D., 2018, ApJ, 855, 96
Horne K., 1986, PASP, 98, 609
Iyer K., Gawiser E., 2017, ApJ, 838, 127
Johnson S. D., Chen H.-W., Mulchaey J. S., 2015, MNRAS, 449, 3263
Jones T. et al., 2015, AJ, 149, 107
Karman W. et al., 2017, A&A, 599, A28
Kelson D. D. et al., 2014, ApJ, 783, 110
Kriek M. et al., 2015, ApJS, 218, 15
Kümmel M., Walsh J. R., Pirzkal N., Kuntschner H., Pasquali A., 2009, PASP, 121, 59
Kümmel M. et al., 2011, A&A, 530, A86
Le Fèvre O. et al., 2005, A&A, 439, 845
Leja J., Johnson B. D., Conroy C., van Dokkum P. G., Byler N., 2017, ApJ, 837, 170
Lotz J. M. et al., 2017, ApJ, 837, 97
Momcheva I. G. et al., 2016, ApJS, 225, 27
Monna A. et al., 2017, MNRAS, 466, 4094
More S., Diemer B., Kravtsov A. V., 2015, ApJ, 810, 36
Morishita T. et al., 2017, ApJ, 835, 254
Morishita T. et al., 2018, ApJ, 856, L4
Morishita T. et al., 2019, ApJ, 877, 141
Muratov A. L., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., Quataert E., Murray N., 2015, MNRAS, 454, 2691
Nelson E. J. et al., 2016, ApJ, 828, 27
Newman A. B., Ellis R. S., Andreon S., Treu T., Raichoor A., Trinchieri G., 2014, ApJ, 788, 51
Oke J. B., 1974, ApJS, 27, 21
Pacifici C., Charlot S., Blaizot J., Brinchmann J., 2012, MNRAS, 421, 2002
Peeples M. S., Werk J. K., Tumlinson J., Oppenheimer B. D., Prochaska J. X., Katz N., Weinberg D. H., 2014, ApJ, 786, 54
Pirzkal N. et al., 2004, ApJS, 154, 501
Postman M. et al., 2012, ApJS, 199, 25
Prochaska J. X., Weiner B., Chen H.-W., Mulchaey J., Cooksey K., 2011, ApJ, 740, 91
Rudie G. C. et al., 2012, ApJ, 750, 67
Schmidt K. B. et al., 2014, ApJ, 782, L36
Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65
Shipley H. V. et al., 2018, ApJS, 235, 14
Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1
Straughn A. N. et al., 2008, AJ, 135, 1624
Straughn A. N. et al., 2009, AJ, 138, 1022
Tanaka M. et al., 2018, PASJ, 70, S9
Treu T. et al., 2015, ApJ, 812, 114
van der Wel A. et al., 2016, ApJS, 223, 29
Vulcani B. et al., 2015, ApJ, 814, 161
Vulcani B. et al., 2016, ApJ, 833, 178
Vulcani B. et al., 2017, ApJ, 837, 126
Wang X. et al., 2017, ApJ, 837, 89
Wang X. et al., 2019, ApJ, 882, 94
Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, ApJ, 754, L29
Zabludoff A. I., Mulchaey J. S., 1998, ApJ, 496, 39

APPENDIX A: FILE NAMES, FORMATS, AND CONTENTS

The core component of the GLASS G800L data base is the catalogue master FITS table:

1 `glass-acs-2019.05.10.fits`.

This table contains field names and IDs for all extracted sources – from which their data filenames can be constructed – along with a wealth of summary metrics described at the end of this appendix.

The master FITS file is designed for the data base's organization into 20 folders – one for each ACS pointing. These are named by their field centroid coordinates in the format:

1 `jRARA`$^{p}_{m}$`DEDE`,

such that '`j0014m3023`' corresponds to the Abell 2744 parallel at PA 323.[13] These correspond to the `ROOT` column in the master catalogue. Within each folder, all objects are identified by a source

[13] A program to translate these coordinates to a cluster field name is available on request and will be posted to the MAST website.

number, corresponding to the `ID` column. As such, the master catalogue can be used to point to the `ii`-th object's data files by writing:

```
folder = mastercat[ii].ROOT
sourceID = string(mastercat[ii].ID, f = '(FI05)')
sourceFileBase = folder+'/'+folder+'_'+sourceID
```

which will construct the base filename from which a source's data files are built (Appendix B).

A1 Individual source files

As mentioned in Section 2.3, each source is associated with multiple data files. These are defined by the following suffixes appended to their base ID `jRARA`$^{p}_{m}$`DEDE_0IDID`:

(i) '`.stack.fits`' – A nine layer 2D FITS table containing:

(a) `SCI` – A source's rectified, drizzled 2D spectral cut-out;
(b) `WHT` – inverse variance array;
(c) `CONTAM` – contamination model;
(d) `MODEL` – source model;
(e) `KERNEL` – *F*814*W* direct image used to convolve the 2D trace to a common spatial sampling.

Layers 1, 2, 4, and 5 are then repeated with the source resampled to a common PA. Since each GLASS source is observed only once, this PA is set to the PA of observation, such that the layers are identical (to <1 per cent). This file's header also contains the wavelength solution, exposure time, RA/Dec., and other ancillary information.

(ii) '`.full.fits`' – An $N \geq 5$ layer 2D FITS table containing:

(a) `ZFIT_STACK` – A six-column table containing redshift estimation information. This includes the `ZGRID` redshift array, $P(z)$ distribution (`PDF`), redshift `RISK` and `CHI2` goodness of fit distributions, and SED template/redshift covariance matrix (`COVAR`);
(b) `COVAR` – Covariance matrix of the template fit coefficients at the best-fitting redshift (taken to be where $P(z)$ is maximized from the previous extension).
(c) `TEMPL` – A three-column table containing the best-fitting SED (`FULL`), continuum-only model (`CONTINUUM`), and wavelength array spanning $0.02 < \lambda/\mu\mathrm{m} < 2 \times 10^4$ (`WAVE`);
(d) `DSCI` – *F*814*W* source direct image sampled at 0.1 arcsec pix^{-1};
(e) `DWHT` – *F*814*W* image weight map sampled at 0.1 arcsec pix^{-1};
(f) `LINE001` – Map of `LINE001` in the file header;
(g) `CONTINUUM` – Continuum near `LINE001` in the file header;
(h) `CONTAM` – Contamination near `LINE001` in the file header;
(i) `LINEWHT` – Weight map near `LINE001` in the file header;
(j) `LINE002` – As above, but for `LINE002` in the file header;

Layers 6–*N* will not be present in an object with no entries in the master catalogue's `HASLINES` column (see below) or no lines in this file's header. Otherwise, e.g. H$\alpha\beta\gamma\delta$, [O II] and [O III], [S II] have names 'Ha, Hb, Hg, Hd, O II, O III, S II,' etc. There is one set of four layers for each line. The header also contains the wavelength solution, exposure time, RA/Dec., line fluxes, and other ancillary information.

(iii) '`.1D.fits`' – A six-column 1D FITS table containing:

(a) `WAVE` – Wavelength array spanning only the G800L bandpass ($5450 < \lambda/\text{Å} < 10\,170$);
(b) `FLUX` – Optimally extracted, contamination-subtracted 1D source spectrum (e$^-$ s^{-1});
(c) `ERR` – 1 σ noise on `FLUX` (e$^-$ s^{-1});
(d) `FLAT` – Sensitivity curve needed to transform the above into f_λ (e$^-$ s^{-1}/erg s^{-1} cm^{-2} Å^{-1});
(e) `LINE` – Best-fitting galaxy SED template, including emission lines.
(f) `CONT` – Best-fitting galaxy SED template, continuum only.

Note that the spatial sampling in these files is 0.1 arcsec pix^{-1} rather than the detection images' 0.03 arcsec pix^{-1}. The `getfname` function in Appendix B returns the filenames for the above FITS tables given a source's `ROOT` and `ID` entry in the master catalogue.

A2 Master catalogue column definitions

The following entries are found in the GLASS G800L master catalogue FITS table:

1 `ROOT` – ACS field (string)
2 `ID` – Source ID in `ROOT` (int)
3 `RA` – Source RA (J2000, decimal deg.; double)
4 `DEC` – Source DEC (J2000, decimal deg.; double)
5 `NINPUT` – *N* frames in source's G800L coadd (int; 1–14)
6 `REDSHIFT` – Source redshift estimate (=z_{MAP}; float)
7 `NUMLINES` – Number of potential lines in source spectrum (int; 0–10)
8 `HASLINES` – Emission-line IDs (string; space-separated)
9 `CHI2POLY` – χ^2 of a third-order polynomial fit to the spectrum (float)
10 `DOF` – Approximate spectral degrees of freedom given by the total number of unmasked pixels in the 'beams' spectra extracted from the grism exposures (float)
11 `CHIMIN` – Minimum χ^2 of the SED template fit on a redshift grid (float)
12 `CHIMAX` – Maximum χ^2 of the SED template fit on a redshift grid (float)
13 `BIC_POLY` – 'Bayesian Information Criterion' of the polynomial fit; $\ln(\mathrm{DOF}) \cdot k + (\mathrm{CHI2POLY\text{-}CHIMIN})$, where $k = 4 + N_\mathrm{b}$ and N_b the number of spectra in the `.beams.` file, since an additive background component is fit for each.
14 `BIC_TEMP` – Bayesian Information Criterion of the template fit, where $k = N_\mathrm{c} + N_\mathrm{b}$ and N_c is the number of non-zero template fit coefficients.
15 `BIC_DIFF` – Difference `BIC_POLY` − `BIC_TEMP`. Large values generally correspond to cases where the galaxy SED templates much better explain the observed spectrum than the trivial polynomial fit.
16 `Z02` – 2 σ redshift lower limit (float)
17 `Z16` – 1 σ redshift lower limit (float)
18 `Z50` – 50 per cent confidence redshift estimate (float)
19 `Z84` – 1 σ redshift upper limit (float)
20 `Z97` – 2 σ redshift upper limit (float)
21 `ZWIDTH1` – 1 σ Δz (84 − 16; float)
22 `ZWIDTH2` – 2 σ Δz (97 − 2; float)

23 Z_MAP – Maximum likelihood redshift (same as REDSHIFT; float)
24 Z_RISK – Redshift where the 'risk' is minimized, where risk is defined as in Tanaka et al. (2018) (float)
25 MIN_RISK – Risk estimate at Z_RISK (float)
26 FLUX_{LINE} – Line flux (erg s^{-1} cm^{-2} Å^{-1}; float)
27 ERR_{LINE} – Line error (erg s^{-1} cm^{-2} Å^{-1}; float)
28 EW50_{LINE} – Median line equivalent width drawn from the line-plus-continuum covariance matrix (observed-frame Å; float)
29 EWHW_{LINE} – Half-width between the 16th and 84th percentiles of the EW estimates for a given line (observed-frame Å; float)
30 SN_{LINE} – Line $S/N = FLUX/ERR$ (float)
31 CHINU – Spectral model fit reduced χ^2 ($CHIMIN/DOF$; float)
32 ZQ – EAZY redshift risk assessment (float)
33 IDX – Vizier search query URL for this source (string)
34 ELLIPTICITY – SExtractor ELLIPTICITY measure (float)
35 MAG_AUTO – SExtractor MAG_AUTO ($F814W$ AB mag; float)
36 FLUX_RADIUS – SExtractor half-light radius (0.03 arcsec pix; float)
37 A_IMAGE – SExtractor semimajor axis length (0.03 arcsec pix; float)
38 T_G800L – Total G800L exposure time (seconds; float)
39 IS_POINT – Point source flag (bit)

where '{LINE}' refers to each of the following:

1 He I λ1083, He II λ1640;
2 Lyman α;
3 N V λ1240, N IV λ1487, N III] λ1750;
4 C IV λ1549, C III] λ1908;
5 [O III] λ1663, [O II] λ3727, [O III] λ4363, [O III] λλ4959, 5007, [O I] λ6302;
6 [Mg II];
7 [Ne V] λ3346, [Ne VI] λ3426, [Ne III] λ3867;
8 Hδ, Hγ, Hβ, Hα + [N II];
9 [S II] λλ6716, 6731, [S III] λλ9069, 9532.

Various quantities inferred from the template SED fits (M/L, M_*, sSFR, and implied fluxes in various filters) are also given, but – as noted in the main text – these estimates are not typically reliable and so caution against their use (Section 3.1).

APPENDIX B: USEFUL ALGORITHMS

Below are some useful IDL routines to perform some basic data operations mentioned in the main text.

B1 Return the high S/N + lines and low/high risk subsamples:

```
function getGoldSample, mastercat, $
                    MAGLIM = maglim, $
                    SNLIM = snlim

 if NOT keyword_set(MAGLIM) then maglim = 24.
 if NOT keyword_set(SNLIM) then snlim = 5.

 ;; Define 'high S/N + lines' objects
 hiQ = where(mastercat.MAG_AUTO le maglim OR $
            (mastercat.SN_HA ge snlim OR $
             mastercat.SN_HB ge snlim OR $
             mastercat.SN_OII ge snlim OR $
             mastercat.SN_OIII ge snlim OR $
             mastercat.SN_SII ge snlim OR $
             mastercat.SN_MGII ge snlim OR $
             mastercat.SN_LYA ge snlim) AND NOT $
             mastercat.IS_POINT, compl = nhiQ)

 ;; Use highQ to define 'low' and 'high risk'
 ;; after nulling the not-hiQ qualities
 mastercat[nhiQ].ZQ = 0
 medrisk = median(mastercat[highQ].ZQ)
 loRisk = where(mastercat.ZQ le medrisk, $
                           compl = hiRisk)
 indices = {HIQ: hiQ, NOTHIQ: nhiQ, $
            LORISK: loRisk, HIRISK: hiRisk}

 RETURN, indices
end
```

B2 Match to HFF photometry from Shipley et al. (2018):

```
function loadHffPhot, hffPhotFile

  ;; Restore the Shipley et al. .save catlaog
  restore, hffPhotFile
  ninds = n_elements(ra)

  ;; Find the NIR intrument
  if n_elements(f_KS_HAWKI) gt 0 then begin
     fks = f_KS_HAWKI
     eks = e_KS_HAWKI
     flgks = redflag_KS_HAWKI
  endif else begin
     fks = f_KS_MOSFIRE
     eks = e_KS_MOSFIRE
     flgks = redflag_KS_MOSFIRE
  endelse

  ;; Create the SED, error, and flag arrays
  ;; (F435W--Spitzer 4.5 um)
  fluxes = transpose([[f_F435W], [f_F606W], [f_F814W],
       [f_F105W], [f_F125W], [f_F140W], [f_F160W],
       [fks], [f_CH1], [f_CH2]])
  errs = transpose([[e_F435W], [e_F606W], [e_F814W],
       [e_F105W], [e_F125W], [e_F140W], [e_F160W],
       [eks], [e_CH1], [e_CH2]])
  flags = transpose([[redflag_F435W], [redflag_F606W],
       [redflag_F814W], [redflag_F105W],
       [redflag_F125W], [redflag_F140W],
       [redflag_F160W], [flgks], [redflag_CH1],
       [redflag_CH2]])
```

```
 ;; Add the MW dust and zeropoint corrections
 ;; (should they be necessary; key names differ
 ;; slightly between catalogues)
 mwcorr = transpose([[mw_F435W], [mw_F606W],
      [mw_F814W], [mw_F105W], [mw_F125W], [mw_F140W],
      [mw_F160W], [1.0], [1.0], [1.0]])
 if N_ELEMENTS(zpcorr_f814w) gt 0 then $
    zpcor = transpose([[zpcorr_F435W], [zpcorr_F606W],
         [zpcorr_F814W], [zpcorr_F105W], [zpcorr_F125W],
         [zpcorr_F140W], [zpcorr_F160W], [1.0], [1.0],
         [1.0]]) $
 else $
    zpcor = transpose([[zpcor_F435W], [zpcor_F606W],
         [zpcor_F814W], [zpcor_F105W], [zpcor_F125W],
         [zpcor_F140W], [zpcor_F160W], [1.0], [1.0],
         [1.0]])

 ;; Set up the output
 nbands = n_elements(fluxes[*,0])
 results = {LAMBDA: [4350., 6060., 8140., $
                  10500., 12500., 14000., 16000., $
                  22500., 35000., 46000.], $
           SED   : fltarr(nbands), $
           ESED  : fltarr(nbands), $
           FLAGS: bytarr(nbands), $
           RA    : 0.d, $
           DEC   : 0.d, $
           MWCORR: fltarr(nbands), $
           ZPCORR: fltarr(nbands)}
 results = replicate(results, ninds)

 ;; Fill the output structure
 ;; get fluxes to erg/s/cm2/Hz
 ninds = n_elements(ra)
 for ii = 0, ninds - 1 do begin
    results[ii].SED = fluxes[*,ii]/10.^29.44
    results[ii].ESED = errs[*,ii]/10.^29.44
    results[ii].FLAGS = flags[*,ii]
    results[ii].RA = ra[ii]
    results[ii].DEC = dec[ii]
    results[ii].MWCORR = mwcorr
    results[ii].ZPCORR = zpcor
 endfor

 RETURN, results
end

;;
;;
;;

pro getSpecPhot, mastercat, field

 if field ne 'Abell 2744' AND $
    field ne 'MACS0717' AND $
    field ne 'MACS1149' AND $
    field ne 'RXJ2248' then $
      print, 'NO HFF ANCILLARY PHOTOMETRY AVAILABLE' $
 else begin

     ;; Load spectral database
    data = mrdfits(mastercat, 1)

    ;; Load photometry
    phot = loadhffphot(field)

    ;; Match with a 1'' radius
    spherematch, phot.RA, phot.DEC, $
          data.RA, data.DEC, 1./3600, $
          minds, sinds, len

    ;; Trim and align databases
    data = data[sinds]
    phot = phot[minds]

 endelse

 ... ;; Operations go here
end
```

B3 Get sources emitting $N_{\rm lines}$ lines at high confidence:

```
function getHighSnLines, mastercat, $
                    SNCUT = sncut

 if NOT keyword_set(SNCUT) then sncut = 5.

 ;; Set up the output storage
 ;; allowing for up to 10 high-sn lines
 output = {NUMLINES: 0, $
          IDS: strarr(10), $
          SNS: fltarr(10)}
 output = replicate(output, n_elements(mastercat))

 ;; Go through the lines and get their S/N
 for ii = 0, n_elements(mastercat) - 1 do begin

    lines = mastercat[ii].HASLINES
    lines = strsplit(lines, ' ', /extr)

    if mastercat[ii].NUMLINES gt 0 then begin

       tsn = fltarr(mastercat[ii].NUMLINES)

       for jj = 0, mastercat[ii].NUMLINES - 1 do begin
          tl = lines[jj]
          case tl of
             'Ha': tsn[jj] = mastercat[ii].SN_HA
             'Hb': tsn[jj] = mastercat[ii].SN_HB
             'OII': tsn[jj] = mastercat[ii].SN_OII
             'OIII': tsn[jj] = mastercat[ii].SN_OIII
             'SII': tsn[jj] = mastercat[ii].SN_SII
             'MgII': tsn[jj] = mastercat[ii].SN_MGII
             'Lya': tsn[jj] = mastercat[ii].SN_LYA
             ... ;; More line names here.
          endcase
       endfor

       ;; Count and store lines about S/N threshold
       foo = where(tsn ge sncut, nlines)
       output[ii].NUMLINES = nlines
       if nlines gt 0 then begin
          output[ii].IDS[0:nlines-1] = lines[foo]
          output[ii].SNS[0:nlines-1] = tsn[foo]
       endif
    endif
 endfor

 RETURN, answer
end
```

B4 Get all of the data file names for a given source:

```
function getfname, root, id

  fname = root+'/'+root+'_'+string(id, f =
       '(I05)')+'.XXX.fits'
  stack = repstr(fname, 'XXX', 'stack')
  full = repstr(fname, 'XXX', 'full')
  oneD = repstr(fname, 'XXX', '1D')
  srcName = root+'_'+string(id, f = '(I05)')

  fnames = {ONED: oneD, $
            FULL: full, $
            STACK: stack, $
            SRCNAME: srcName}

  RETURN, fnames
end
```

APPENDIX C: COMPARISON TO EXTANT PHOTOMETRIC REDSHIFTS

Shipley et al. (2018) provide photometric redshifts for GLASS G800L sources in the HFF footprint. Fig. C1 compares these to the grism-only redshifts. Agreement is quite good for low risk sources, and unbiased for the full high-$S/N+$ lines sample. The low risk sample is consistent with Gaussian errors, though errors may be slightly underestimated for the other samples ($\sigma/\sqrt{2}\langle \mathrm{err}\rangle > 1$). Nevertheless, this analysis supports our conclusion that the ∼2200 low risk objects should be immediately useful to analyses 'straight out of the box.'

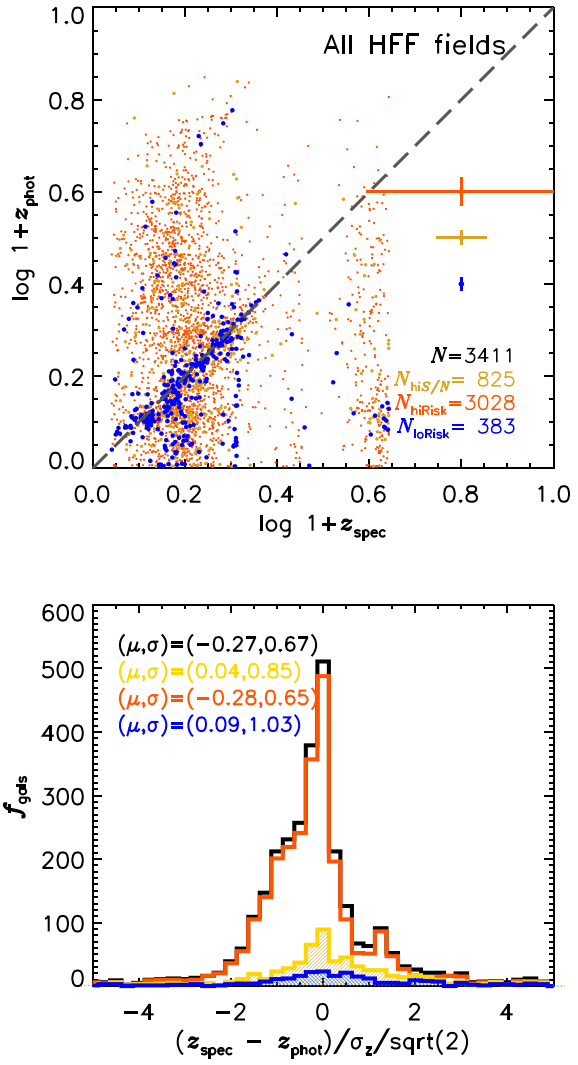

Figure C1. Similar to Fig. 4 but plotting HFF photometric redshifts from Shipley et al. (2018) ($z_{\rm phot}$) against GLASS G800L redshifts ($z_{\rm spec}$). Distributions are much wider, but agreement is still fair, with low risk sources having photo-z errors consistent with Gaussian noise. All samples show slight biases, but those for high-$S/N+$ lines and low-risk sources are small in the mean.

APPENDIX D: FURTHER DATA EXAMPLES

Below are more examples of GLASS G800L spectra and combined spectrophotometry for select sources in the HFF footprint. All galaxies shown Fig. D1 have $S/N_{[\mathrm{O\,III}]} \geq 5$ and are above the median brightness in the low risk sample. There are 304 such sources in the GLASS ACS data base. All galaxies in Fig. D2 are in the low risk sample with matching HFF photometry. There are 383 such objects in the GLASS ACS data base (3411 in total with HFF overlap).

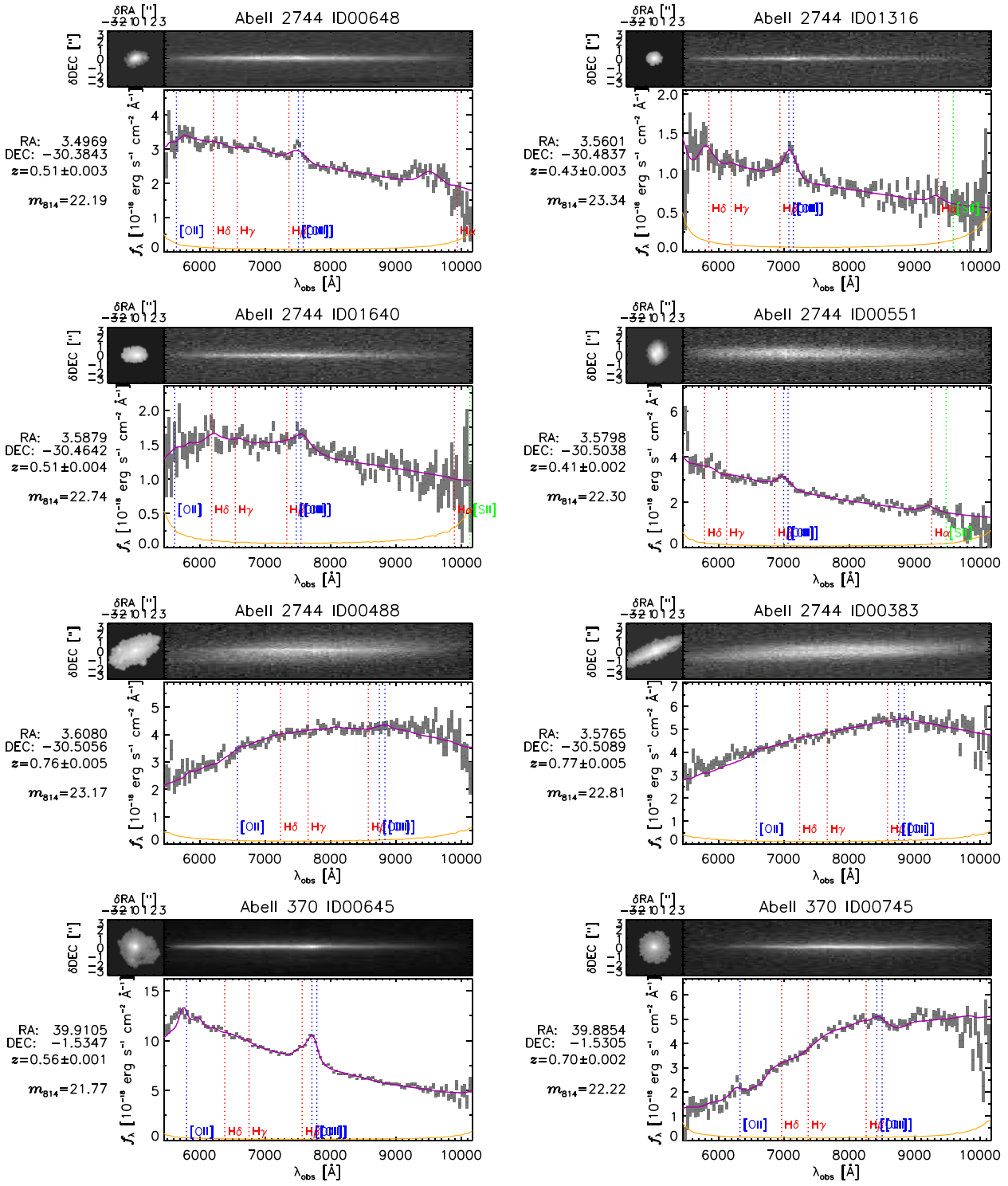

Figure D1. Examples of the ∼300 bright, $S/N \geq 5$ [O III] emitters in the glass data base.

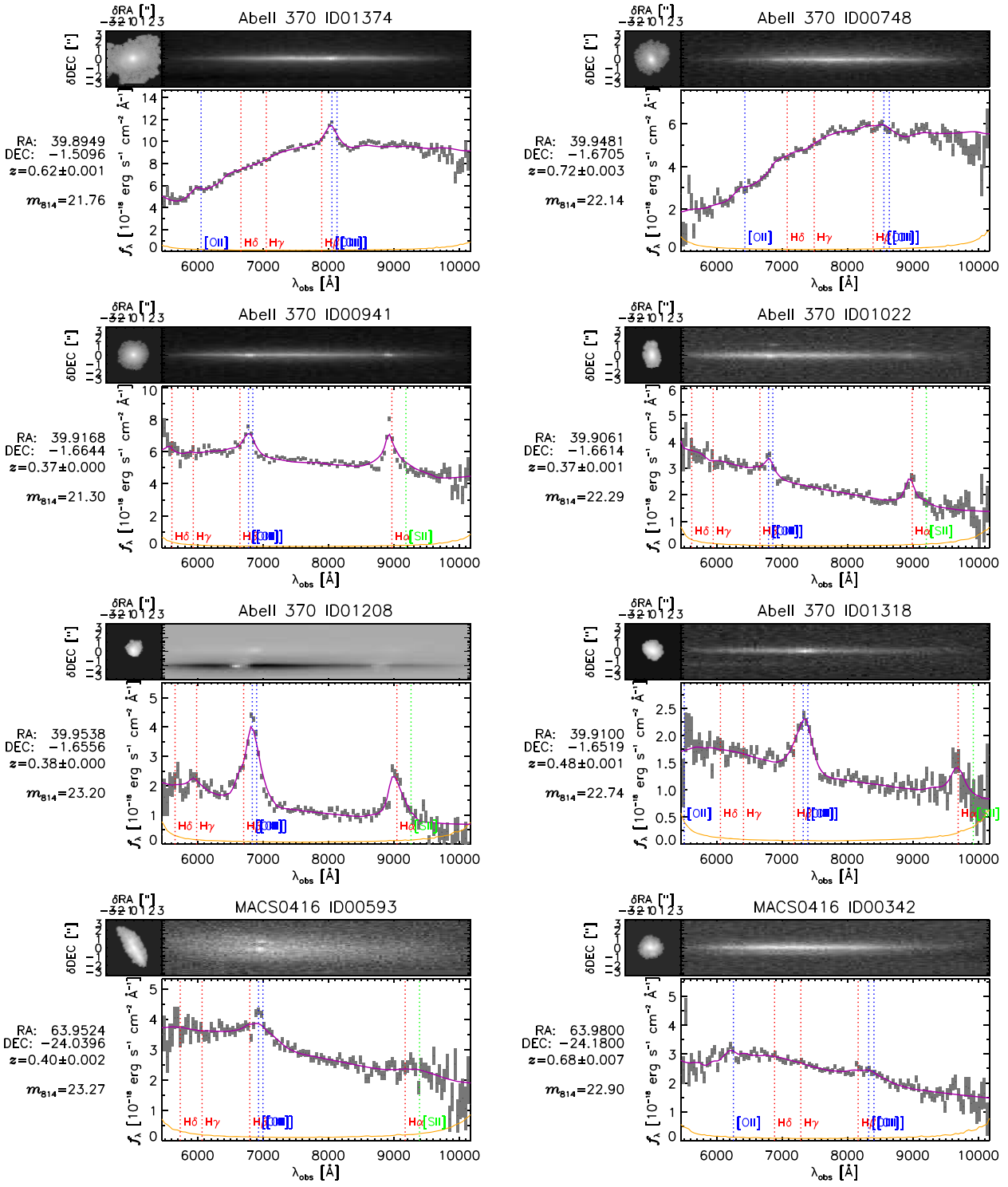

Figure D1 *–continued*

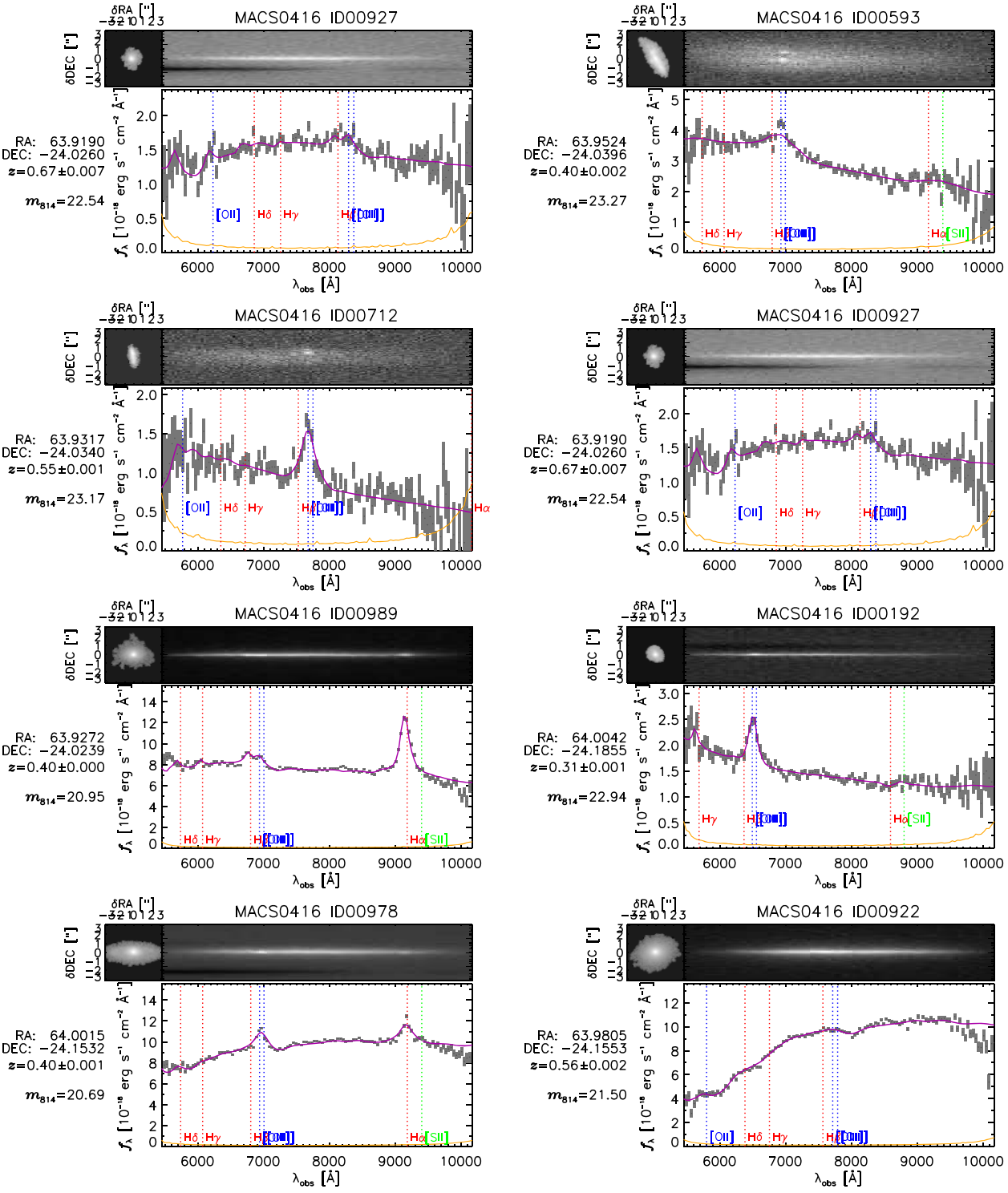

Figure D1 *–continued*

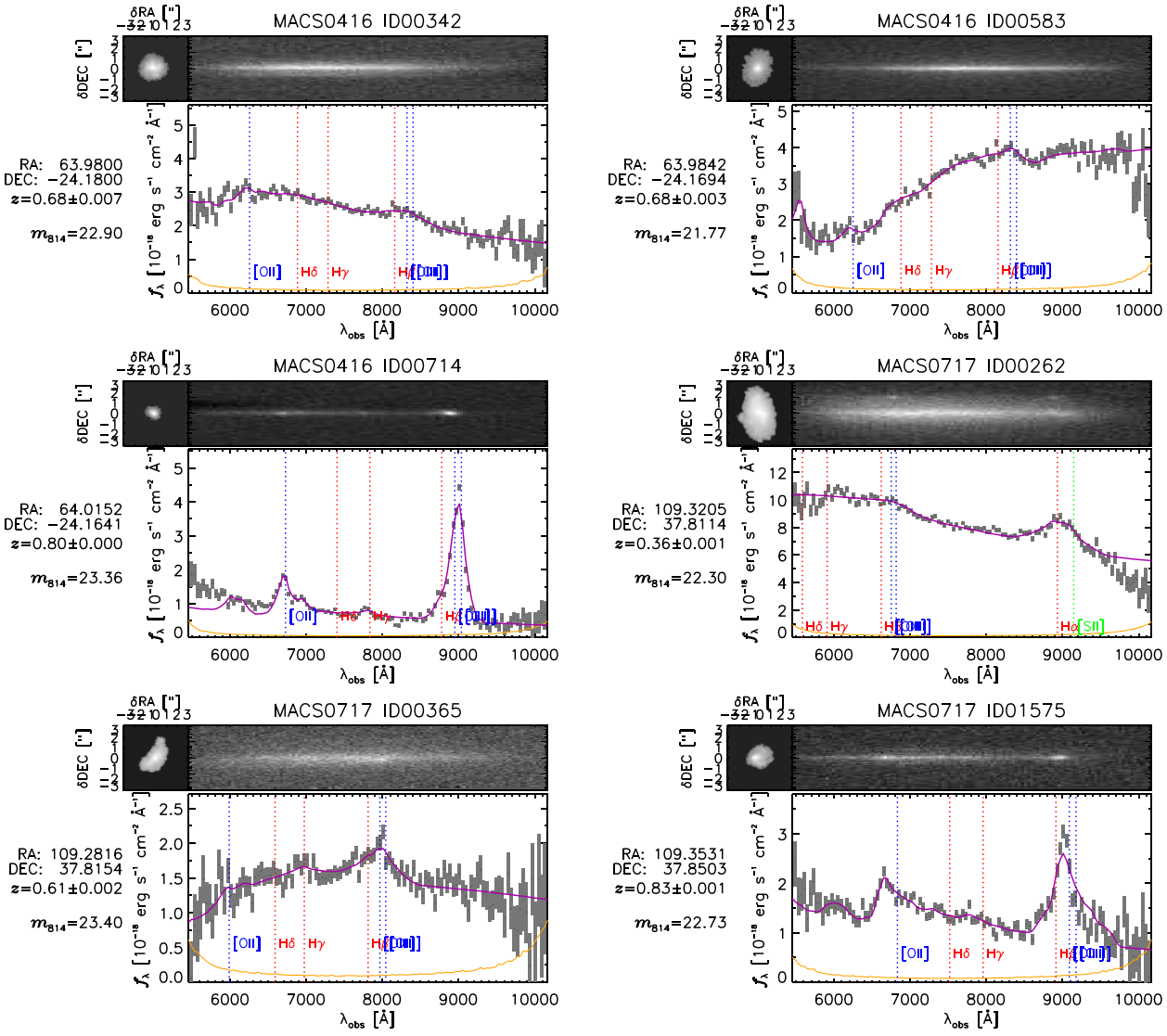

Figure D1 *–continued*

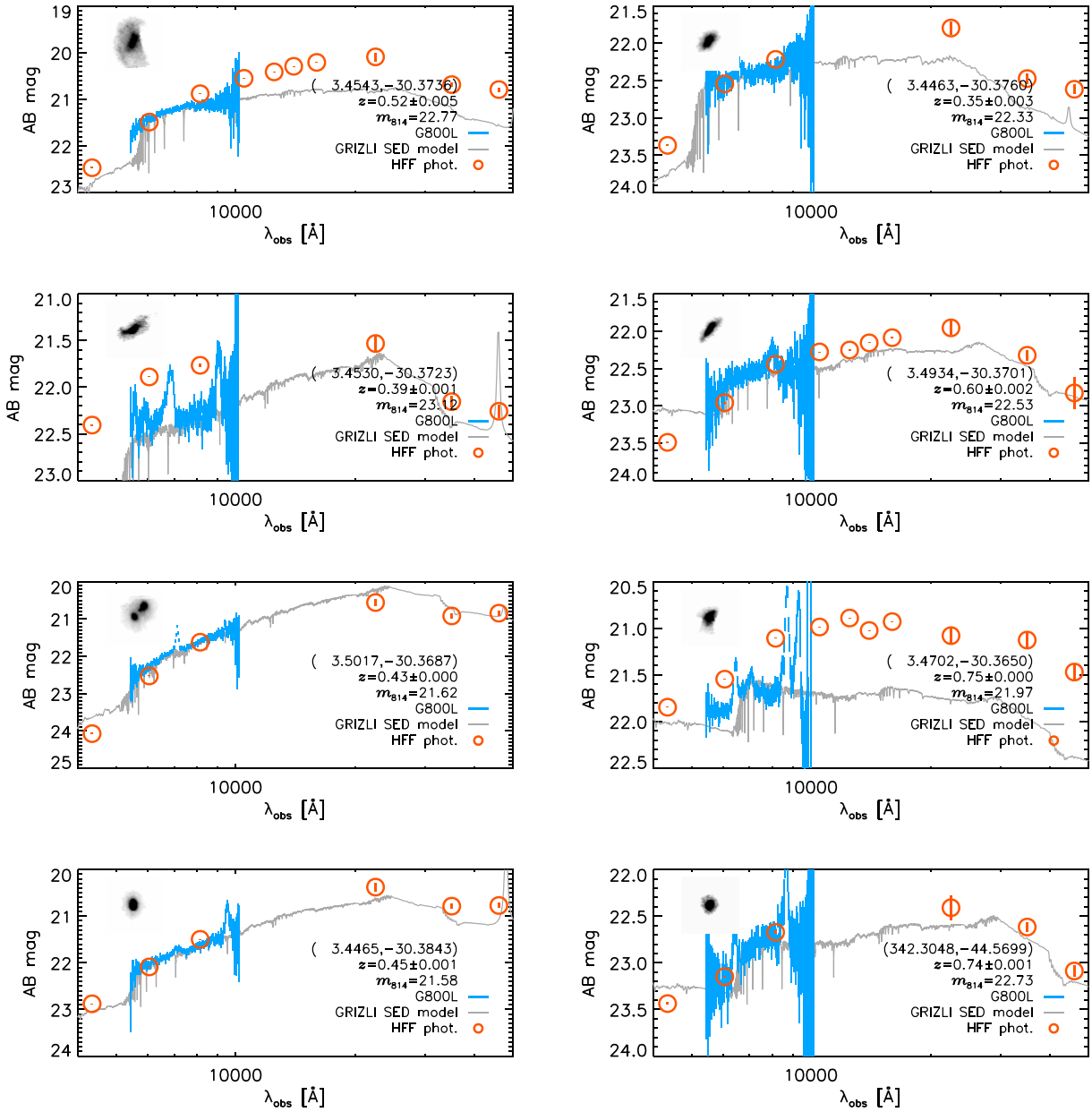

Figure D2. GLASS G800L sources in the Abell 2744 and RXJ2248 parallel fields matched to HFF photometry from Shipley et al. (2018).

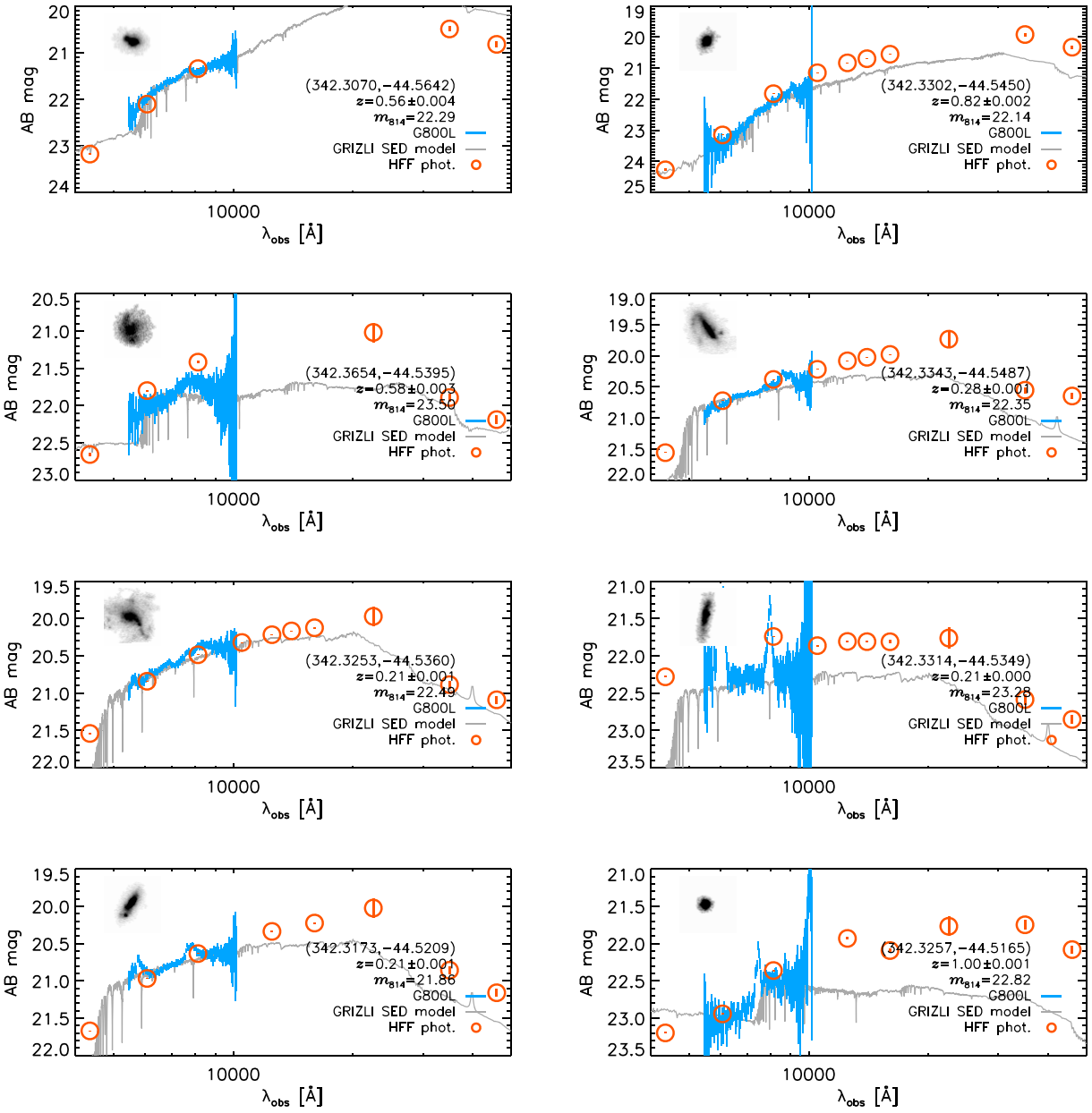

Figure D2 *–continued*

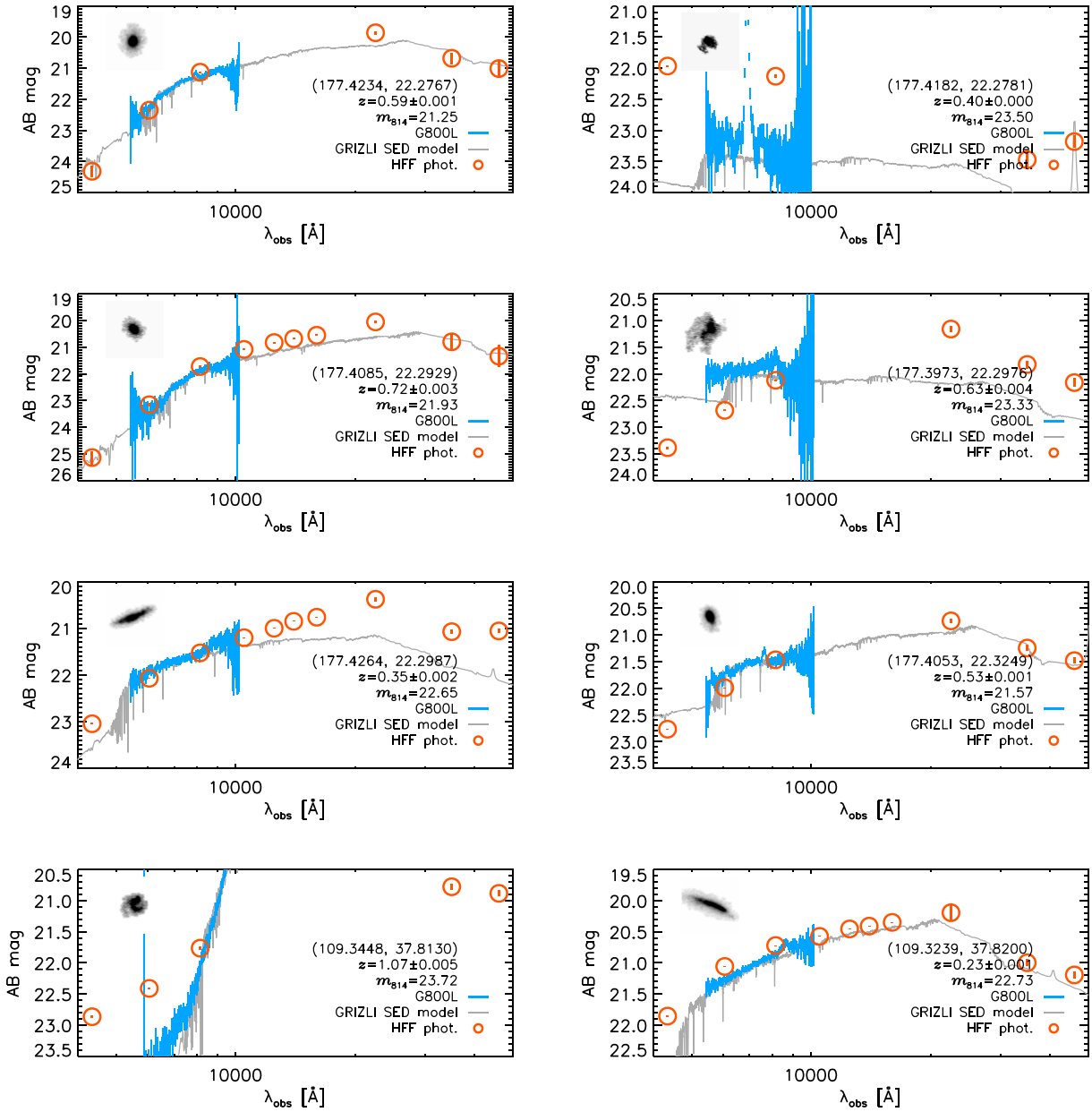

Figure D2 *–continued*

This paper has been typeset from a TEX/LATEX file prepared by the author.