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


Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
10.1093/mnras/staa276

Publication date:
2020

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC

Citation for published version (APA):
The Grism Lens-Amplified Survey from Space (GLASS) – XIII. G800L optical spectra from the parallel fields

L. E. Abramson, G. B. Brammer, K. B. Schmidt, T. Treu, T. Morishita, X. Wang, B. Vulcani, and A. Henry

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 ∼ 0.5, to quasars at z ∼ 4, to the star formation histories of hundreds of galaxies in between.

The spectroscopic catalogue has a median redshift of ⟨z⟩ = 0.60 with a median uncertainty of ∆z/(1 + z) ≲ 2 per cent at F814W ≲ 23 AB. Robust continuum detections reach a magnitude fainter. The 5σ limiting line flux is f_{lim} ≈ 5 × 10^{-17} erg s^{-1} cm^{-2} and half of all sources have 50 per cent of pixels contaminated at ≲ 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 – 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 ≲ 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
of its portfolio with the forthcoming operation of JWST, Euclid, and WFIRST. Hence, building intuition, applications, and tools to handle these data is prudent. JWST’s less widely used optical disperser – the ACS G800L grism – is powerful in this context. Covering $\lambda = 0.5$–$1.0\,\mu m$ with $\sim 2.5 \times$ 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$\alpha$ 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 $(H_0, \Omega_m, \Omega_\Lambda) = (70, 0.3, 0.7)$ 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 (Lotz et al. 2017) and four more from CLASH (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$\alpha$ emitters. G102 and G141 were exposed there $\sim 220$ sq. arcmin to 7-orbit ($\sim 2$ Mpc) away from the cluster centres.

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

A F814W pre-image was taken with each grism exposure (2 per orbit per PA). These data are necessary for source detection, spectral

## 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,VA’.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Abell 2744</td>
<td>j0014m3030</td>
<td>0:14:20:9</td>
<td>-30:29:46</td>
<td>225</td>
<td>1066</td>
<td>16</td>
<td>10 076</td>
<td>HFP, Rodney</td>
<td></td>
</tr>
<tr>
<td>Abell 370</td>
<td>j0240m0132</td>
<td>2:39:31:6</td>
<td>-1:31:42</td>
<td>343</td>
<td>1028</td>
<td>31</td>
<td>10 076</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Abell 370</td>
<td>j0240m0140</td>
<td>2:39:44:6</td>
<td>-1:40:08</td>
<td>245</td>
<td>965</td>
<td>32</td>
<td>10 344</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>MACS0416</td>
<td>j0416m2402</td>
<td>4:15:44:7</td>
<td>-24:02:04</td>
<td>337</td>
<td>808</td>
<td>15</td>
<td>10 076</td>
<td>CLASH, Rodney</td>
<td></td>
</tr>
<tr>
<td>MACS0416</td>
<td>j0416m2410</td>
<td>4:15:56:0</td>
<td>-24:09:42</td>
<td>254</td>
<td>887</td>
<td>24</td>
<td>10 186</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>MACS0717</td>
<td>j0717m3750</td>
<td>7:17:16:3</td>
<td>37:49:53</td>
<td>10</td>
<td>1123</td>
<td>46</td>
<td>9927</td>
<td>HFF, CLASH, Siana, Ebeling</td>
<td></td>
</tr>
<tr>
<td>MACS0717</td>
<td>j0717m3747</td>
<td>7:18:01:6</td>
<td>37:47:26</td>
<td>110</td>
<td>1186</td>
<td>37</td>
<td>10 061</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>MACS0744</td>
<td>j0745m3922</td>
<td>7:45:08:6</td>
<td>39:22:16</td>
<td>194</td>
<td>853</td>
<td>37</td>
<td>9562</td>
<td>CLASH</td>
<td></td>
</tr>
<tr>
<td>MACS0744</td>
<td>j0745m3930</td>
<td>7:45:20:6</td>
<td>39:30:10</td>
<td>109</td>
<td>934</td>
<td>35</td>
<td>10 061</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>MACS1149</td>
<td>j1150p2218</td>
<td>11:49:40:3</td>
<td>22:18:02</td>
<td>215</td>
<td>1117</td>
<td>37</td>
<td>10 084</td>
<td>HFF, CLASH$^4$, Siana</td>
<td></td>
</tr>
<tr>
<td>MACS1149</td>
<td>j1150p2225</td>
<td>11:50:00:9</td>
<td>22:25:25</td>
<td>122</td>
<td>1090</td>
<td>47</td>
<td>9854</td>
<td>CLASH</td>
<td></td>
</tr>
<tr>
<td>RXJ1347</td>
<td>j1347m1140</td>
<td>13:47:19:7</td>
<td>-11:30:11</td>
<td>13</td>
<td>1049</td>
<td>31</td>
<td>10 076</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>RXJ1347</td>
<td>j1347m1147</td>
<td>13:47:09:4</td>
<td>-11:37:36</td>
<td>293</td>
<td>954</td>
<td>20</td>
<td>10 344</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>MACS1423</td>
<td>j1424p2401</td>
<td>14:24:06:8</td>
<td>24:00:38</td>
<td>178</td>
<td>813</td>
<td>40</td>
<td>10 076</td>
<td>CLASH</td>
<td></td>
</tr>
<tr>
<td>MACS1423</td>
<td>j1424p2408</td>
<td>14:24:09:4</td>
<td>24:08:19</td>
<td>98</td>
<td>939</td>
<td>33</td>
<td>10 344</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>MACS2129</td>
<td>j2130m0736</td>
<td>21:29:31:5</td>
<td>-7:35:40</td>
<td>58</td>
<td>1131</td>
<td>39</td>
<td>10 061</td>
<td>Rodney</td>
<td></td>
</tr>
<tr>
<td>MACS2129</td>
<td>j2130m0742</td>
<td>21:29:54:2</td>
<td>-7:41:45</td>
<td>136</td>
<td>390</td>
<td>127</td>
<td>18 789</td>
<td>CLASH, Riess</td>
<td></td>
</tr>
<tr>
<td>RXJ2248</td>
<td>j2249m4433</td>
<td>22:49:18:0</td>
<td>-44:32:37</td>
<td>143</td>
<td>906</td>
<td>16</td>
<td>9372</td>
<td>HFF, CLASH, Riess</td>
<td></td>
</tr>
</tbody>
</table>

---

$^a$All objects with $z_{cl} - z_q/(1 + z_{cl})^{-1} < 2000$ km s$^{-1}$ irrespective of $z_q$ where $z_q$ is the GLASS cluster redshift from table 1 of Treu et al. (2015).

$^b$Covering band other than F814W to support spectral energy distribution SED (spectral energy distribution) fitting.


$^d$One band.

$^e$Grim data incorporated from GO12099 (PI Riess).

$^f$James Webb Space Telescope.

$^g$Wide-Field Infrared Survey Telescope.

$^h$Advanced Camera for Surveys

$^i$Hubble Frontier Fields.

$^j$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’ F814W 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 F814W 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 astrometric alignment including both a fine relative alignment anchored.

The former serves as the reference frame to which the latter is anchored.

GRIZLI first identifies cosmic rays in the F814W 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 F814W direct and G800L grism exposures, the alignment of the former is applied directly to the latter. Finally, we combine the F814W direct exposures into a rectified mosaic with ASTRODRIZLE 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 \(f_\lambda\), flux density with a normalization set by its observed flux in the F814W 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 F814W 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 F814W 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 \(\chi^2(\hat{z})\) 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 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 `.beam` 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 F814W...
direct image. All are rectified such that the spectral dispersion aligns with the data’s \( x \)-axis. Unlike in the .beam files, this rectification (and drizzling) necessitates a resampling of the original detector
emission lines, the rectified, drizzled F\( \text{its} \), the best-fitting high-resolution template SED with and without
information.

\( (iv) \ A'\ .\ _\text{full} \ . \ _\text{fits} \) 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\( \text{its} \) 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.

\( (v) \ A'\ .\ _\text{1D} \ . \ _\text{fits} \) table for each object comprising the wavelength array, contamination-subtracted source spectrum (e\(^{-}\) s\(^{-1}\)), sensitivity curve (e\(^{-}\) s\(^{-1}\) erg s\(^{-1}\) cm\(^{-2}\) Å\(^{-1}\)), 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 \( m_{B,14} = 26 \) – 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 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\( \sigma \) \( f_{\text{L}} \) 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 \( \log f/\text{erg s}^{-1}\text{ cm}^{-2} \gtrsim -16.5 \) for point sources, rising to \( \log f \approx -16.0 \) 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 \( \log f \approx -16.4 \) to \( \log f \approx -15.4 \) across that size range (turquoise).

Figure 2. Random examples of redshift estimation quality in the four quartiles of the redshift quality parameter, \( z_{Q} \), based on the \( m_{B,14} < 24 \) or \( S/N \gtrsim 5 \) in any optical strong line quality cut. Variation in the PDFs is noticeable.

(i) All non-point sources with \( m_{B,14} \leq 24 \) or \( S/N_{\text{line}} \geq 5 \) (‘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, \( S/N_{\text{line}} \) refers to the \( S/N \) in any of the following features as identified by GRIZLI: H\( \alpha \), H\( \beta \), [O II], [O III], [S II], [Mg II], or Ly\( \alpha \). The magnitude cut corresponds roughly to the inferred 5\( \sigma \) continuum sensitivity – \( \sim 5 \times 10^{-19} \) erg s\(^{-1}\) cm\(^{-2}\) Å\(^{-1}\) (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 \( \Delta z \equiv 84^{\text{th}} - 16^{\text{th}} \text{P}(z) \) percentile. Fig. 2 shows \text{P}(z) PDFs for each quartile of the \( z_{Q} \) distribution.

There are low risk sources with no detected emission lines. These are \( \sim 1 \) 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 \( \langle z \rangle \approx 0.60 \), with the low risk sample at \( \langle z \rangle \approx 0.58 \). 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).

\( m_{814} = 22.55 \) 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 \([\text{OII}] \lambda 3727\) 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 \( z \approx 0.60 \), 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 \( \sigma \) 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 \( \sigma \approx 0.01 \) below \( z = 1 \) (Fig. 4; data from Grillo et al. 2016; Caminha et al. 2017; Karman et al. 2017; Monna et al. 2017).\(^\text{11}\) Fig. 5 illustrates that this quality is typical at \( m_{814} \lesssim 22 \) \((S/N \sim 20 \text{ per spectral pixel})\), 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 \( m_{814} \).

\(^\text{11}\)Available at https://sites.google.com/site/vltclashpublic/data-release.
Colour coding is as in Fig. 3. Sources in the low risk sample fainter than the full sample (grey) and the baseline quality cut (main GLASS catalogue limits (Morishita et al. 2017). We caution, log $M_{\text{stellar mass}}$ (∗ probably high-EW line emitters. Of course, all of the above depends fundamentally on the grism resolution of $\sim$ 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 $\log M_*/M_\odot \sim 7$ at $z \sim 0.6$ to $\log M_*/M_\odot \sim 9$ 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 resolution and therefore may change significantly when ancillary photometry is included in the inference (Section 4.2).

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 number of lines detected at $S/N_{\text{line}} \geq 5$ and quality cut. The corresponding line IDs for the low risk samples are shown at right. Equation (1) characterizes the magnitude and wavelength dependence of the fifth percentile line flux for these objects:

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

where $\Lambda \equiv \lambda/7200\, \text{Å}$ and $M = m_{814} - 23$ 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 $S/N_{\text{line}}$, these IDs should be treated with care: visual inspection suggests $\sim$ 20–40 per cent of sources at fixed $N_{\text{lines}} \in [1, 3]$ 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] $\lambda 3727$ 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 H$\alpha$. The source on the right is a high-EW [O iii] emitter at $z \sim 0.65$ with quite different oxygen emissions. not an impediment; half the sample has $> 50$ per cent of pixels free of any contamination. Alternatively, half of sources have $\leq 20$ 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.
3.3 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 Lyα, carbon lines, and point-like morphology, the source at left is a z ~ 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 ρ ≤ 150 (300) kpc impact parameter from the quasar suitable for characterizing HI 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 ∼ 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 ~ 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 m814 = 23.8, this object shows that GLASS’ data reach continuum levels comparable to those from surveys on 10-m class ground-based

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.
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 $\sim 300$ bright, $S/N \geq 5$ [O III] emitters in the GLASS G800L data base are shown in Appendix D.

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 $\sim 20$–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 $\Delta m = 0.1$ mag (such that HFF fluxes are brighter) with a scatter of 0.25 mag about that offset. The grism-derived continuum SED model extrapolated well outside the G800L bandpass can be predictive to similar levels – low risk median $m_{GRIZLI} - m_{HFF} = [0.19, 0.25, 0.25, 0.25]$ 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.
More examples of low risk objects in the HFF footprints are shown in quantities (Dressler et al. 2016; Abramsone et al. 2018; Dressler, in prep) in inferring the SFHs of galaxies at $0.3 < z < 1$ with broad-band and UV can obviously occur. Scatters around these offsets – but larger disagreements in the IR – which we also give here in Appendix B. The GLASS data will support similar analyses at $z$ $\approx$ $0.6$ with median uncertainties of $\Delta z/(1 + z) \lesssim 0.02$ at $m_{814} \lesssim 23$ (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 $\lesssim 5 \sigma$ ($m_{814} \lesssim 24$; $f_\lambda \approx 5 \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$ Å$^{-1}$) or S/N $> 5$ in at least one spectral line ($f_{\text{lin}} \approx 5 \times 10^{-17}$ 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).

**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 H$\alpha$ 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 H$\beta$ 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] $\propto$ $z$) 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 H$\alpha$ 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 ($\sim$10 ks) coverage over $\sim 220$ sq. arcmin. Sources have a median redshift of $z = 0.6$ with median uncertainties of $\Delta z/(1 + z) \lesssim 0.02$ at $m_{814} \lesssim 23$ (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 $\lesssim 5 \sigma$ ($m_{814} \lesssim 24$; $f_\lambda \approx 5 \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$ Å$^{-1}$) or S/N $> 5$ in at least one spectral line ($f_{\text{lin}} \approx 5 \times 10^{-17}$ 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-0653(A).

REFERENCES

Abramson L. E. et al., 2018, AJ, 156, 29
Brammer G., 2019, Astrophysics Source Code Library, record, ascl:1905.001
Jones T. et al., 2015, AJ, 149, 107
Tanaka M. et al., 2018, PASJ, 70, S9

APPENDIX A: FILE NAMES, FORMATS, AND CONTENTS

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

\[ \text{glasacs-2019.05.10.fits} \]

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:

\[ \text{j} \times \text{RA}, \text{DE} \times \text{DE} \]

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

13A 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:

```plaintext
folder = mastercat[i].ROOT
sourceID = string(mastercat[i].ID, f = '\(\{(\FS)\}\)')
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: 

\(j\)RA\(\alpha\)DE\(\delta\)O\(\ID\):

(i) ‘\(\_{\text{stack}}.\text{fits}\)’ – 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 – \(F814W\) 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) ‘\(\_{\text{full}}.\text{fits}\)’ – 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\(\chi\) 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\(\chi\) 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 \(0.02 < \lambda/\mu m < 2 \times 10^4\) (WAVE);
(d) DSCI – \(F814W\) source direct image sampled at 0.1
arcsec pix\(^{-1}\);
(e) DNHT – \(F814W\) 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 in
lines in this file’s header. Otherwise, e.g. \(H\alpha\), \(H\beta\), \([O\text{ II}]\), and \([O\text{ III}]\), \([S\text{ II}]\) have names ‘\(H\alpha\), \(H\beta\), \(H\gamma\), \(H\delta\), \([O\text{ II}]\), \([O\text{ III}]\), \([S\text{ II}]\),’ 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) ‘\(\_{\text{fit}}.\text{fits}\)’ – A six-column 1D FITS table containing:

(a) WAVE – Wavelength array spanning only the G800L
bandpass (\(5450 < \lambda/\AA < 10170\));
(b) FLUX – Optimally extracted, contamination-subtracted 1D
source spectrum (\(e^{-1}\) s\(^{-1}\));
(c) ERR – \(1\) \(\sigma\) noise on FLUX (\(e^{-1}\) s\(^{-1}\));
(d) FLAT – Sensitivity curve needed to transform the above
into \(s^{-1}\) erg s\(^{-1}\) cm\(^{-2}\) \(\AA^{-1}\);
(e) LINE – Best-fitting galaxy SED template, including emis-

sion 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 (=\(z_{\text{MAP}}\); float)
7 NUMLINES – Number of potential lines in source spectrum
(int; 0–10)
8 HASLINES – Emission-line IDs (string; space-separated)
9 CHI2POLY – \(\chi^2\) 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 \(\chi^2\) of the SED template fit on a redshift
grid (float)
12 CHIMAX – Maximum \(\chi^2\) of the SED template fit on a redshift
grid (float)
13 BIC\_POLY – ’Bayesian Information Criterion’ of the polynomial
fit; ln(DoF) \(\cdot k + (\text{CHI2POLY-CHIMIN})\), where \(k = 4 + N_\ell\)
and \(N_\ell\) the number of spectra in the .beam file, since an
additive background component is fit for each.
14 BIC\_TEMP – Bayesian Information Criterion of the template
fit, where \(k = N_\ell + N_\ell\) and \(N_\ell\) 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 – \(2\) \(\sigma\) redshift lower limit (float)
17 Z16 – \(1\) \(\sigma\) redshift lower limit (float)
18 Z50 – 50 per cent confidence redshift estimate (float)
19 Z84 – \(1\) \(\sigma\) redshift upper limit (float)
20 Z97 – \(2\) \(\sigma\) redshift upper limit (float)
21 ZWIDTH1 = 1 \(\sigma\) \(\Delta z\) (84 – 16; float)
22 ZWIDTH2 = 2 \(\sigma\) \(\Delta z\) (97 – 2; float)
 Various quantities inferred from the template SED fits (M/L, 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:

```idl
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_GII ge snlim OR $ mastercat.SN_GIII 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, NOT(hiQ): nhiQ, $ NOT(loRisk: LoRisk, HIRISK: hiRisk)
RETURN, indices end
```

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

```idl
function loadHFFPhot, hffPhotFile $ 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
  flagks = redflag.KS_HAWKI
endif else begin
  fks = f.KS_MOSFIRE
  eks = e.KS_MOSFIRE
  flagks = 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.M160W], [fks], [f.CHI], [f.CH2])
errs = transpose([e.F435W], [e.F606W], [e.F814W], [e.F105W], [e.F125W], [e.M160W], [eks], [e.CHI], [e.CH2])
flags = transpose([redflag.F435W], [redflag.F606W], [redflag.F814W], [redflag.F105W], [redflag.F125W], [redflag.F140W], [redflag.F160W], [flagks], [redflag.CH1], [redflag.CH2])
```
:: Add the MW dust and zeropoint corrections
:: (should they be necessary; key names differ
:: slightly between catalogues)

mncorr = transpose([[mوء_F44F5W], [mوء_F69F6W],
                   [mوء_F125W], [mوء_F120W], [mوء_F14W],
                   [mوء_F160W], [1.0], [1.0], [1.0])]

if N_ELEMENTS(zpcorr_F44F5W) gt 0 then $
  zpcorr = transpose([[zpcorr_F44F5W], [zpcorr_F69F6W],
                      [zpcorr_F125W], [zpcorr_F120W],
                      [zpcorr_F14W], [zpcorr_F160W], [1.0], [1.0],
                      [1.0]]) $ 
else $
  zpcorr = transpose([[zpcorr_F44F5W], [zpcorr_F69F6W],
                      [zpcorr_F125W], [zpcorr_F120W],
                      [zpcorr_F14W], [zpcorr_F160W], [1.0], [1.0],
                      [1.0]]) $ 
endif 

:: Set up the output

nbands = n_elements(Fluxes[*][0])
results = [LAMBDA: [4350., 6560., 8140., 10590., 12500., 14000., 16000., 22500., 35000., 46000.],
          SED : fltarr(nbands), $
          ESED : fltarr(nbands), $
          FLAGS : bytarr(nbands), $
          RA : 0.d, $
          DEC : 0.d, $
          MCORR : fltarr((nbands, 1)), $
          ZPCORR : fltarr((nbands, 1))]
results = replicate(results, ninds)

:: Fill the output structure
:: get fluxes to erg/s/cm2/Hz

minds = n_elements(ra)
for ii = 0, minds - 1 do begin
  results[ii].SED = fluxes[*][ii] / 10. ^ 29.44
  results[ii].ESED = errors[*][ii] / 10. ^ 29.44
  results[ii].FLAGS = flags[*][ii]
  results[ii].RA = ra[ii]
  results[ii].DEC = dec[ii]
  results[ii].MCORR = mncorr
  results[ii].ZPCORR = zpcorr
endfor
RETURN, results
end

:: Get sources emitting $N_{\text{line}}$ lines at high confidence:

function getHighSNLines, mastercat, $ 
      SNCUT = snout 
      
if NOT keyword_set(SNCUT) then snout = 5.
:: Set up the output storage
:: allowing for up to 10 high-sn lines
output = [NUM_LINES: 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].NUM_LINES gt 0 then begin
    tsn = fltarr(mastercat[ii].NUM_LINES)
    for jj = 0, mastercat[ii].NUM_LINES - 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
        'Ly' : tsn[jj] = mastercat[ii].SN_LyA
      endcase
    endfor
    :: Count and store lines about S/N threshold
    foo = where(tsn ge snout, nlines)
    output[ii].NUM_LINES = 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

pro getSpecPhot, mastercat, field

if field eq 'Abell 2744' AND $ 
  field eq 'MACS0717' AND $ 
  field eq 'MACS1149' AND $ 
  field eq 'RXJ2248' then $ 
print, 'NO HFF ANTIQUITY PHOTOMETRY AVAILABLE' $ 
else begin
  :: Load spectral database
  data = mrdfits(mastercat, 1)
  :: Load photometry
  phot = loadhfpht(field)
  :: Match with a 1'' radius
  spheredmatch, phot.RA, phot.DEC, $
B4 Get all of the data file names for a given source:

```aloha
function getfname, root, id

    fname = root+'/'+'root'+ '_'+string(id, f = '(10s)')+'.'+XXX_file
    stack = repstr(fname, 'XXX', 'stack')
    full = repstr(fname, 'XXX', 'full')
    oneD = repstr(fname, 'XXX', '1D')
    srcName = root+'_' +string(id, f = '(10s)')

    fnames = {NONE: 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 ($\sigma/\sqrt{2}\langle\text{err}\rangle > 1$). Nevertheless, this analysis supports our conclusion that the $\sim 2200$ 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) ($z_{\text{phot}}$) against GLASS G800L redshifts ($z_{\text{spec}}$). 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 $S/N_{[O III]} \geq 5$ 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
This paper has been typeset from a TeX/\LaTeX{} file prepared by the author.