WFC3 Infrared Spectroscopic Parallel (WISP) survey: photometric and emission-line data release

A.J Battisti, M.B Bagley, M Rafelski, I Baronchelli, Y.S Dai, A.L Henry, H Atek, J Colbert, M.A Malkan, P.J Mccarthy, et al.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **530,** 894–928 (2024)
Advance Access publication 2024 April 1

https://doi.org/10.1093/mnras/stae911

WFC3 Infrared Spectroscopic Parallel (WISP) survey: photometric and emission-line data release

A. J. Battisti [1,2★] M. B. Bagley [3★] M. Rafelski [4,5] I. Baronchelli [6,7] Y.S. Dai [8] A. L. Henry [4,5] H. Atek [9] J. Colbert,[10] M. A. Malkan [11] P. J. McCarthy,[12] C. Scarlata [13] B. Siana [14] H. I. Teplitz [10] A. Alavi [10] K. Boyett [2,15] A. J. Bunker [16] J. P. Gardner [17] N. P. Hathi [4] D. Masters [10] V. Mehta [10] M. Rutkowski [18] K. Shahinyan [13] B. Sunnquist [4] and X. Wang [19,20,21]

Affiliations are listed at the end of the paper

Accepted 2024 March 22. Received 2024 March 9; in original form 2023 May 1

ABSTRACT

We present reduced images and catalogues of photometric and emission-line data (∼230 000 and ∼8000 sources, respectively) for the WFC3 (Wide Field Camera 3) Infrared Spectroscopic Parallel (WISP) survey. These data are made publicly available on the *Mikulski Archive for Space Telescopes* and include reduced images from various facilities: ground-based *ugri*, *Hubble Space Telescope* (*HST*) WFC3, and *Spitzer* IRAC (Infrared Array Camera). Coverage in at least one additional filter beyond the WFC3/IR data are available for roughly half of the fields (227 out of 483), with ∼20 per cent (86) having coverage in six or more filters from *u* band to IRAC 3.6 μm (0.35–3.6 μm). For the lower spatial resolution (and shallower) ground-based and IRAC data, we perform PSF (point spread function)-matched, prior-based, deconfusion photometry (i.e. forced-photometry) using the TPHOT software to optimally extract measurements or upper limits. We present the methodology and software used for the WISP emission-line detection and visual inspection. The former adopts a continuous wavelet transformation that significantly reduces the number of spurious sources as candidates before the visual inspection stage. We combine both WISP catalogues and perform spectral energy distribution fitting on galaxies with reliable spectroscopic redshifts and multiband photometry to measure their stellar masses. We stack WISP spectra as functions of stellar mass and redshift and measure average emission-line fluxes and ratios. We find that WISP emission-line sources are typically 'normal' star-forming galaxies based on the mass–excitation diagram ([O III]/Hβ versus $M_\star$; $0.74 < z_{\rm grism} < 2.31$), the galaxy main sequence (SFR versus $M_\star$; $0.30 < z_{\rm grism} < 1.45$), S_{32} ratio versus $M_\star$ ($0.30 < z_{\rm grism} < 0.73$), and O_{32} and R_{23} ratios versus $M_\star$ ($1.27 < z_{\rm grism} < 1.45$).

Key words: catalogues – surveys – ISM: evolution – galaxies: evolution – galaxies: general – galaxies: photometry.

1 INTRODUCTION

Slitless spectroscopy provides simultaneous spectra for every source in an observed field of view (FOV). Such observations have the notable advantage that they can be used to efficiently perform blind spectroscopic surveys over large areas of the sky for the study of galaxy evolution, relative to slits, fibres, or integral field spectroscopy (IFS) that cover much smaller FOVs. This gain in mapping area from slitless spectroscopy usually comes at the expense of reduced sensitivity due to a higher background, more source confusion, and low to moderate spectral resolution ($R \sim 100$–1000) relative to slits or IFS ($R \gtrsim 2000$). The latter limits dynamical analyses and may result in lines blending together (e.g. Hα + [N II]).

Blind surveys provide a unique and unbiased view of galaxy evolution by avoiding issues of cosmic variance or photometric pre-selection. The grisms on the *Hubble Space Telescope* (*HST*) have been widely utilized to perform slitless spectroscopic surveys of thousands of galaxies at intermediate redshifts ($0.5 \lesssim z \lesssim 2.5$; e.g. APPLES (ACS Pure Parallel Lyα Emission Survey), Pasquali et al. 2003; GRAPES (Grism ACS Program for Extragalactic Science), Pirzkal et al. 2004; WISP, Atek et al. 2010; 3D-HST, Brammer et al. 2012; PEARS (Probing Evolution And Reionization Spectroscopically), Pirzkal et al. 2013; GLASS (Grism Lens Amplified Survey from Space), Treu et al. 2015; FIGS (Faint Infrared Grism Survey), Pirzkal et al. 2017; MAMMOTH–Grism (MApping the Most Massive Overdensity Through Hydrogen), Wang et al. 2022; CLEAR (CANDELS Lyα Emission at Reionization), Simons et al. 2023; MUDF (MUSE Ultra Deep Field), Revalski et al. 2023). Next generation facilities, including the *JWST*, *Euclid*, and the *Nancy Grace Roman Space Telescope* (*Roman*), will revolutionize this capability by providing slitless spectroscopy for millions of galaxies that span larger areas of the sky and a wider range in cosmic time (e.g. Bagley et al. 2020).

The WFC3 (Wide Field Camera 3) Infrared Spectroscopic Parallel (WISP, Atek et al. 2010) survey was the largest, multicycle *HST* pure-parallel grism program. WISP obtained slitless spectroscopy for thousands of galaxies in 483 pointings using up to two near-infrared (NIR) grisms, *G102* (0.80–1.15 μm, $R \sim 210$) and *G141* (1.08–1.69 μm, $R \sim 130$). Pure-parallel surveys, such as WISP,

★ E-mail: Andrew.Battisti@anu.edu.au (AJB); mbagley@utexas.edu (MBB)

have the benefit that the data are obtained 'for-free' for numerous random fields that are independent and uncorrelated (i.e. unbiased). The versatility of large slitless surveys like WISP are demonstrated by the range of science cases that can be explored, such as: characterizing the star-forming galaxy main sequence (MS, Atek et al. 2014), characterizing dust attenuation (Domínguez et al. 2013) and dust attenuation curves (Battisti et al. 2022), characterizing the mass–metallicity relation (Henry et al. 2013, 2021), characterizing massive, quenched galaxies (Bedregal et al. 2013), identifying galaxy pairs (Dai et al. 2021), identifying single spectral lines through machine learning (Baronchelli et al. 2020, 2021), identifying Lyman-α emitters (Bagley et al. 2017), identifying bright, rare galaxies (Bagley et al. 2024), predicting emission-line galaxy number counts for future surveys (Colbert et al. 2013; Mehta et al. 2015; Bagley et al. 2020), investigating the mass–size relation of passive galaxies (Zanella et al. 2016), crowd-sourced analysis of slitless spectroscopic data (Dickinson et al. 2018), nitrogen enhancement of star-forming galaxies (Masters et al. 2014), and the discovery of very faint, distant (400 pc) brown dwarfs (Masters et al. 2012).

A challenge for pure-parallel slitless surveys is that the random fields are sporadic in position and typically do not have deep ancillary photometric data available (e.g. compared to legacy deep fields). To remedy this, the WISP team has carried out several supplementary observing programs to obtain additional photometry with a variety of facilities, with priority given to the deepest ∼200 WISP fields. These include *HST* WFC3/UVIS optical, ground-based optical, and *Spitzer*/IRAC NIR data.

This paper describes the public data release of a self-consistent photometric catalogue, reduced images, and the emission-line catalogue of WISP fields. These data are hosted at the WISP website,[1] on the *Mikulski Archive for Space Telescopes* (*MAST*). Currently ∼50 per cent of the WISP fields have additional photometric data beyond WFC3/IR. These data provide a valuable galaxy sample for legacy science and can serve as a useful reference for ongoing and future grism surveys with *JWST*, *Euclid*, and *Roman*.

This paper is organized as follows: Section 2 describes the observational data, Section 3 describes the WISP photometric pipeline, Section 4 presents the WISP photometric catalogue, Section 5 describes the WISP spectroscopic pipeline, Section 6 presents the WISP emission-line catalogue, Section 7 shows our results from combining both catalogues to study galaxies at $0.3 < z_{\rm grism} < 2.3$, and Section 8 summarizes our main conclusions. Throughout this work, we adopt a lambda-cold dark matter cosmological model, with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$. All magnitudes are in the AB magnitude system (Oke & Gunn 1983).

2 OBSERVATIONS AND DATA REDUCTION

2.1 WISP survey – *HST* WFC3/IR grism spectroscopy and imaging

The WISP survey (PI: M. Malkan; Atek et al. 2010) is a multicycle (cycles 17–23) *HST* pure-parallel program that obtained WFC3/IR observations for 483 pointings in random fields (i.e. location depended on the primary observing target and its position angle). Of the 483 parallel pointings (which we denote as 'Par no.'; e.g. Par1, Par2, etc.), there are 40 fields with partial overlap (i.e. there are 443 unique fields). The position of the WISP pointings are shown in Fig. 1. The program IDs and details for the different *HST* cycles are summarized

[1] https://archive.stsci.edu/prepds/wisp/

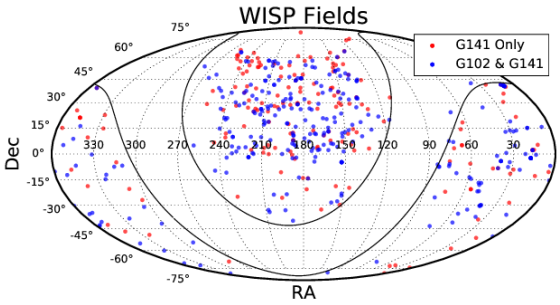


Figure 1. Locations of the 483 WISP fields. These fields are primarily located outside of the Galactic Plane. Most short parallel opportunities obtained WFC3/IR *G141* grism data along with one imaging filter (*F140W* or *F160W*) and long opportunities obtained *G102* + *G141* grisms and two imaging filters (*F110W* + *F140W* or *F110W* + *F160W*).

in Table 1. The effective grism area of each WISP pointing is ∼3.55 arcmin^2 in *G102* and ∼3.8 arcmin^2 in *G141*, relative to the full 4.6 arcmin^2 WFC3/IR FOV, due to area on both the left and right sides of each field being 'lost'. On the left, this is because sources are not covered in the direct images necessary for source identification and wavelength calibration. On the right, this is because emission lines cannot be distinguished from contaminating zero-order images (Bagley et al. 2017). In principle, high-resolution ancillary data with a larger FOV, such as the WFC3/UVIS data available for a subset of fields, could be used to locate sources outside the WFC3/IR FOV. However, such an approach is not implemented for the current data release. Thus, the effective imaging and grism areas of the full WISP survey are ∼2200 and ∼1600 arcmin^2, respectively (this also accounts for field overlap). An overview of the imaging and grism filters that are included in this data release are shown in Fig. 2.

Due to the nature of parallel observations, the integration times for each field was set by the primary target observations. We refer the reader to Atek et al. (2010) for a complete description of the observing strategy and data reduction. In brief, short opportunities (1–3 continuous orbits; 180 fields) usually obtained *G141* and one WFC3/IR filter, either *F140W* or *F160W*. However, a handful of short opportunities obtained *G102* and *F110W* instead (4 fields). Long opportunities (four or more continuous orbits; 299 fields) obtained *G102*+*G141* and two WFC3/IR filters, either *F110W*+*F140W* or *F110W*+*F160W*, except for Cycle 19 which used *F140W*-only. Cycle 20 and beyond are almost exclusively deep fields (*G102* + *G141*; See Table 1). In general, grism integration times are ∼6 × those for the direct images. For the long opportunities, the integration times in the two grisms were set to achieve approximately uniform sensitivity for an emission line of a given flux across the full wavelength range (∼5:2 for *G102*:*G141*). The median 5σ detection limit for emission lines fluxes (point source) in both grisms is $\sim 5 \times 10^{-17}$ erg s^{-1} cm^{-2} (Atek et al. 2010), but we stress that this varies considerably with the length of the opportunity and variation in background levels (zodiacal light, Earth limb brightening, and telescope thermal emission). The line detection limits for each individual WISP fields are presented in Fig. 3 and Table 2.

Due to *HST* only having a single dispersion direction for each grism, and position/orientation being tied to the Primary observing target, all grism data are single orientation. We note that multi-orientation pure-parallel observations are possible on *JWST* due to the availability of multidirectional grisms, and that this is already being utilized (e.g. Cycle 1 PASSAGE survey with NIRISS, GO 1571; PI: M. Malkan, Malkan et al. 2021). Multi-orientation grism

Table 1. Summary of WISP *HST* Survey (PI: M Malkan).

GO ID	Cycle	Observing period[a]	WFC3/UVIS[b]	WFC3/IR	WFC3/IR grism	Orbits
11696	17	2009 Nov 24– 2011 May 1	*F475X*, *F600LP*	*F110W*, *F140W*, *F160W*	*G102*, *G141*	250
12283	18	2010 Oct 6–2012 Mar 11	*F606W*, *F814W*[c]	*F110W*, *F160W*	*G102*, *G141*	280
12568	19	2011 Oct 29–2013 May 19	None	*F140W*	*G102*, *G141*	260
12902	20	2012 Oct 16–2014 Mar 20	*F606W*, *F814W*	*F110W*, *F160W*	*G102*, *G141*	260
13352	21	2013 Oct 31–2015 June 14	*F606W*, *F814W*	*F110W*, *F160W*	*G102*, *G141*	375
13517	21	2013 Dec 8–2015 Feb 12	*F606W*, *F814W*	*F110W*, *F160W*	*G102*, *G141*	200
14178	23	2015 Nov 30–2017 May 1	*F606W*, *F814W*	*F110W*, *F140W*, *F160W*	*G102*, *G141*	520
					Total:	2145

Notes. There are 184 short opportunity fields (one to three continuous orbits), 180 fields with *G141* and either *F140W* or *F160W*, and 4 fields with *G102* and *F110W* (see Section 2.1). There are 299 long opportunity fields (four or more continuous orbits); these obtained *G102*+*G141* and either *F110W*+*F140W* or *F110W* + *F160W*, except for Cycle 19 which used *F140W*-only.
[a]Observations often extended beyond the nominal *HST* cycle period.
[b]UVIS data were only obtained for a subset of the deep fields (155 fields; see Section 2.2).
[c]Most UVIS data in this cycle used 2 × 2 binning in the *F475X* and *F600LP* filters and those are not included in the current data release.

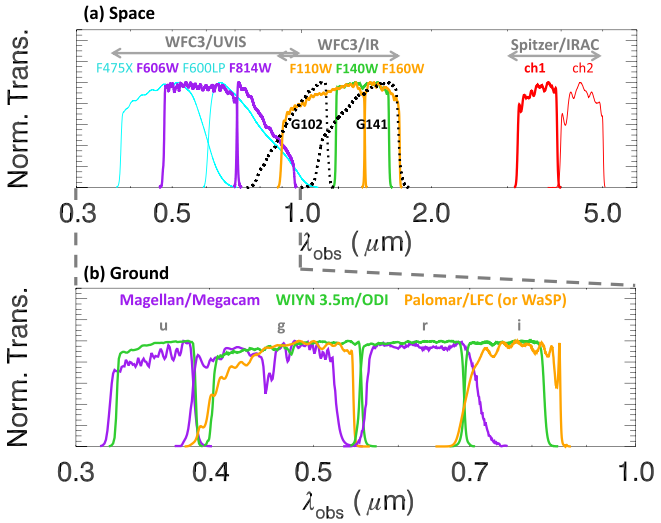


Figure 2. Filters used in the WISP survey and follow-up observations separated by (a) space- and (b) ground-based facilities. The WISP survey obtained WFC3/IR imaging together with the grism spectroscopy (dotted black curves) and all WISP fields have either *F140W* or *F160W* data, with deep fields also having *F110W* imaging. Additional photometry is inhomogeneous and comprises of WFC3/UVIS, *Spitzer*, and/or ground-based optical data (*ugri*) from Magellan, WIYN 3.5m, and Palomar. The *F475X*, *F600LP*, and IRAC ch2 curves are shown as thinner lines to highlight that they are available for <20 fields (other filters are >20 fields), with a breakdown shown in Table 5.

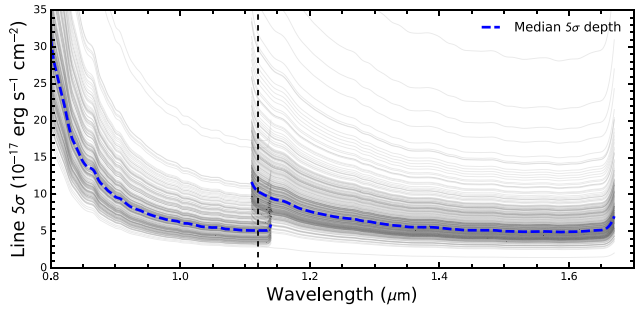


Figure 3. The 5σ emission-line sensitivities in the individual WISP fields (light grey curves). The blue dashed line indicates the median 5σ field depth. The depth varies significantly field-to-field due to variations in the exposure time and background levels in each field.

data can be used to construct spatially resolved emission-line maps (e.g. Pirzkal et al. 2018; Wang et al. 2022).

All WISP data are reduced and calibrated with the WFC3 pipeline `calwf3`, together with custom scripts to improve the calibration and account for the specific challenges of undithered, pure-parallel observations. The foundation of the WISP reduction pipeline is described in Atek et al. (2010). Here, we additionally implement a multiple component fit to the sky background in the grism images. While the 'input' spectrum of the sky can be considered uniform across the FOV, the sky emission is not uniform on the detector. Typically, a single master sky image is scaled and fit to remove the background in each grism image, under the assumption that the relative strength of the background components remain the same exposure to exposure. However, as noted by Brammer et al. (2012), the structure in the sky of WFC3 grism images is highly variable, with zodiacal light and a helium emission line at 1.083 μm contributing. While the zodiacal component should stay relatively constant throughout a full set of exposures taken on the same date, the helium emission comes from Earth's upper atmosphere and can vary on time-scales shorter than the length of an exposure as *HST* moves closer or farther from Earth and the telescope pointing changes. We model and subtract the background in each grism exposure with a linear combination of the zodiacal light and helium sky images created by Brammer, Ryan & Pirzkal (2015), calculating the amplitudes of each sky component with two iterations of a least-squares fit.

We use the `AstroDrizzle` and `TWEAKREG` routines of DRIZZLEPAC (Gonzaga et al. 2012; Hoffmann et al. 2021) to combine the individual exposures, correcting for astrometric distortions and any alignment issues between exposures. The IR direct and grism images are drizzled onto a 0.08 arcsec pixel^{-1} scale. Object detection in the IR direct images (*F110W*, *F140W*, and *F160W*) is performed with Source Extractor (version 2.5; Bertin & Arnouts 1996). For fields with imaging in two filters, we create a combined detection image and supplement the catalogue with sources detected individually in only one of the filters. We use the AXE software package (Kümmel et al. 2009) to extract and calibrate the spectra, using the Source Extractor catalogues as inputs.

2.2 *HST* WFC3/UVIS imaging

For 155 of the long opportunity fields, WFC3/UVIS (FOV = 2.7 arcmin×2.7 arcmin) imaging data were also obtained with some combination of *F475X*, *F606W*, *F600LP*, and/or *F814W* filters (only two of these at most for a single field). These observations were

Table 2. WISP emission-line depths for the 419 fields in the emission-line catalogue.

Par	RA	Dec.	1σ depth *G102*[a] (erg s^{-1} cm^{-2})	1σ depth *G141*[a] (erg s^{-1} cm^{-2})
1	01:06:35.29	+15:08:53.8	9.67×10^{-18}	9.57×10^{-18}
2	01:25:10.02	+21:39:13.7	...	1.17×10^{-17}
5	14:27:06.64	+57:51:36.2	7.72×10^{-18}	5.30×10^{-18}
6	01:50:17.18	+13:04:12.8	1.15×10^{-17}	8.76×10^{-18}
7	14:27:05.90	+57:53:33.7	1.54×10^{-17}	8.56×10^{-18}
			...	

Notes. A full ASCII version of this table is available online.
[a]Grism flux limits depend on wavelength (see Fig. 3). Values presented here are at $\lambda = 1.1$ μm for *G102* and 1.5 μm for *G141*.

all carried out as part of the main parallel survey (PI: M. Malkan; see Table 1 for cycles and GO IDs). The variety in UVIS photometric filters used is a result of a change in strategy between cycles. Specifically, in the first cycle of WISP observations (Cycle 17), the *F475X* and *F600LP* filters were utilized to maximize throughput. However, it was realized that the less-wide filters provided more information for spectral energy distribution (SED) fitting, and thus future observations switched to *F606W* and *F814W*. In Cycle 18, we switched to 2×2 binned observations for the UVIS imaging, although it was later realized that the combination of lack of binned calibrations and pixel based charge transfer efficiency corrections made the data less useful. In Cycle 19, we only obtained short parallel opportunities, and thus none included UVIS imaging. Starting with Cycle 20 and onwards, all UVIS observations were obtained in unbinned *F606W* and *F814W* data. In total, we have 70 and 43 fields with two and one unbinned UVIS filters, respectfully, for a total of 113 fields with *G102* + *G141* grism spectroscopy and excellent UVIS imaging.

2.2.1 HST *data reduction and co-addition*

The UVIS images are also reduced through `calwf3`, along with additional custom calibrations described in Rafelski et al. (2015) including custom dark calibration files to address hot pixel masking, background gradients, and blotchy background patterns. These calibrations only apply to unbinned WFC3 data and, as a result, the binned UVIS data in Cycle 18 are not included in this data release. We first applied a pixel-based Charge Transfer Efficiency (CTE) correction on all individual raw dark files. When post-flashed dark files were available, we used those. The Cycle 17 UVIS data (*F475X* and *F600LP*) had high background and minimal CTE degradation. All UVIS data since Cycle 20 (*F606W* and *F814W*) use post-flash or have sufficient background on their own based on the recommendations at the time. We then used the dark files from a 3–5 d window to create a super dark to identify the hot pixels. These hot pixels use a lower threshold to identify hot pixels otherwise lost due to CTE degradation. We also create a master superdark that stacks all dark observations within the same anneal cycle as the observations, which we subtract from the data within `calwf3` to reduce the blotchy patterns and gradients. The details of all these steps are outlined in the appendix of Rafelski et al. (2015). This dark calibration methodology was partly incorporated as part of UVIS 2.0 (Bourque & Baggett 2016) utilizing post-flashed CTE-corrected darks. Then later in 2020 the dark reference files moved to contemporaneous darks. The pipeline still does not include improvements for hot pixels. For future reductions, we point to the hot pixel treatment in Prichard et al. (2022) and Revalski et al. (2023), which also includes additional improvements to artefacts caused by CTE degradation such as matched amplifier backgrounds and read out cosmic rays.

The UVIS image mosaics were created with `AstroDrizzle` and `TWEAKREG`. The images where corrected for cosmic rays and aligned to the NIR images. The final image drizzle parameters are set to have `final_scale` of 0.04 arcsec pixel^{-1}, `final_pixfrac` = 0.75, and The `final_wht_type` = IVM (inverse-variance weighting map). We also create NIR and UVIS image mosaics with matched plate scales (`final_scale`) of 0.04, 0.08, and 0.13 arcsec pixel^{-1}. The use of `final_scale` of 0.08 or 0.13 arcsec pixel^{-1} for UVIS data produces significantly undersampled images, and we therefore use the 0.04 arcsec pixel^{-1} NIR and UVIS mosaics for matched photometry. These images are then convolved with a kernal to match the point spread-function (PSF) of all the images to that of the *F160W* filter. We used the IRAF task `PSFMATCH` to calculate the matching kernals and convolve the images. Additionally, RMS images are created by taking $1/\sqrt{(\mathrm{IVM})}$ and cleaned images of the UVIS data are generated by masking edges, chip gaps, and bad pixels with randomly generated Gaussian noise. The pixel-matched imaging data in both WFC3/UVIS and WFC3/IR were made available in previous data releases from the WISP team via the survey website on *MAST*.[2]

2.3 Ground-based imaging

The ground-based data were taken between the 2010A–2019A semesters from various PIs across several independent programs. The data were taken with Palomar/LFC (Large Format Camera; 31 nights), Palomar/WaSP (Wafer-Scale camera for Prime; 2 nights), WIYN/MiniMosaic (Wisconsin-Indiana-Yale-NOIRLab; 12 nights), WIYN/ODI (One Degree Imager; 9.5 nights), and Magellan/Megacam (3.5 nights). A summary of the programs, including their IDs, PIs, dates, and the number of nights, are listed in Table 3.

The observing campaign can be broadly divided into two categories: (1) programs with Palomar (LFC or WASP) and WIYN/MiniMosaic focused on observations in the *g* and *i* bands, where *i* band was only obtained for fields without *HST*/UVIS *F814W* or *F600LP* data (these filters have similar effective wavelengths), with the goal of improving stellar mass estimates for WISP galaxies; (2) programs with WIYN/ODI and Magellan/Megacam focused on observations in *u*, *g*, and *r* bands, where *r* band was only obtained for cases without *HST*/UVIS *F606W* data (similar in wavelength), with the goal of providing rest-frame UV measurements to constrain reddening from dust attenuation (Battisti et al. 2022). The

[2]https://archive.stsci.edu/prepds/wisp/

Table 3. Summary of ground-based imaging campaigns on WISP fields included in this release.

ID	PI(s)	Telescope/Instr.	Semester	Bands	Nights
P-23	B. Siana	Palomar/LFC	2010A	*gi*	4
H-05	B. Siana	Palomar/LFC	2010B	*gi*	2
J-06	C. Scarlata	Palomar/LFC	2010B	*gi*	3
0438	A. Henry	WIYN/MiniMo	2010B	*gri*	4
0160	A. Henry	WIYN/MiniMo	2011A	*gi*	5
0222	A. Henry	WIYN/MiniMo	2011B	*g*	3
J-18	H. Telpitz	Palomar/LFC	2012A	*gi*	3
J-01	M. Rafelski	Palomar/LFC	2013A	*gi*	2
J-10	M. Rafelski	Palomar/LFC	2013B	*gi*	3
J-23	M. Rafelski	Palomar/LFC	2014A	*gi*	2
J-05	Y. Dai	Palomar/LFC	2014B	*gi*	2
J-16	J. Colbert	Palomar/LFC	2015A	*gi*	3
J-12	Y. Dai	Palomar/LFC	2016A	*g*	1
J-14	Y. Dai	Palomar/LFC	2016B	*gi*	3
J-14	I. Baronchelli	Palomar/LFC	2017B	*gi*	3
0298	A. Battisti	WIYN/ODI	2017B	*ugr*	2.5
...	P. McCarthy and A. Battisti	Magellan/Megacam	2017B	*ugr*	1.5
0275	A. Battisti	WIYN/ODI	2018A	*ugr*	2.5
...	A. Battisti	Magellan/Megacam	2018A	*ugr*	2
0140	A. Battisti	WIYN/ODI	2018B/19A	*ugr*	4.5[a]
N-115	Y. Dai	Palomar/WASP	2019A	*g*	2

Notse. Columns list the (1) program ID, (2) principal investigator, (3) telescope and instrument used, (4) semester observed, (5) bands observed, and (6) number of nights where data were taken. [a]Three of these nights were provided as additional technical time.

WIYN/ODI and Magellan/Megacam instruments have large FOVs (40 arcmin × 48 arcmin and 24 arcmin × 24 arcmin, respectively), which are significantly larger than the *HST* WFC3/IR footprint (FOV = 2.3 arcmin × 2.2 arcmin). Whenever possible, observations were optimized to observe multiple WISP fields simultaneously.

The desired depths of the programs varied, but for WIYN/ODI and Magellan/Megacam they were AB mag = 26, 26, and 25.3 in *u*, *g*, and *r* (5σ point source), respectively. These depths were typically achieved for the Magellan runs but not for most of the WIYN runs, due to poor weather and/or observing conditions (e.g. lunar phase). A detailed summary on the depths for each facility is given in Section 4.2.

2.3.1 *Ground-based data reduction and co-addition*

Below we describe the data reduction process for each telescope individually, but note that they follow a similar overall procedure that includes bias, flat, and dark-subtraction. All images were astrometrically aligned to the World Coordinate System (WCS) using the `astrometry.net` software package, unless otherwise specified. Flux calibrations were also performed in similar ways, unless otherwise specified, and is summarized in Section 3.1.

(i) *Palomar/LFC* – the Palomar/LFC WISP observations spanned the largest number of different observers among our data sets. All raw images were re-reduced following the same procedure for self-consistency. The LFC consists of six chips (each 6.14 arcmin × 12.29 arcmin; referred to as `chip0` through `chip5`), covering a circular area with a ∼24 arcmin diameter, and the WISP field was always centred on `chip0`. Basic data reduction steps were only performed on the `chip0` data. We note that dark frames were not always available for subtraction, however, this has a minimal impact on the final photometric quality because the LFC has very little dark current.[3]

(ii) *Palomar/WaSP* – basic reduction steps were identical to Palomar/LFC, except that WaSP has four large chips (square FOV: 18.4 arcmin × 18.5 arcmin) and all chips were reduced.

(iii) *WIYN/Minimosaic* – the raw archival WIYN data were obtained via the NOAO Science Archive (now known as the NOIRLab Astro Data Archive).[4] Basic reduction steps were similar to Palomar, except that Minimosaic has two large chips (square FOV: 9.6 arcmin × 9.6 arcmin) and both chips were used.

(iv) *WIYN/ODI* – basic reduction, WCS-alignment, and photometric calibration was performed using the ODI Pipeline, Portal, and Archive (PPA) system.[5] PPA is a service provided by the WIYN Consortium, Inc., and hosted at Indiana University. The photometric calibration in the ODI PPA is based on measurements from the SDSS (Sloan Digital Sky Survey; *ugriz*) and Pan-STARRS (Panoramic Survey Telescope And Rapid Response System; *grizy*). We note that for a few fields outside the SDSS footprint, the *u*-band ODI data required manual photometric calibration, which is discussed in Section 3.1.

(v) *Magellan/Megacam* – images with basic reduction and WCS-alignment were provided by the OIR Telescope Data Center at Harvard's Center for Astrophysics (CfA), supported by the Smithsonian Astrophysical Observatory.

The process for co-adding images was the same for all ground-based data. First, images are normalized by exposure time. Second, normalized images are combined using the SWARP software (Bertin et al. 2002; Bertin 2010). During SWARP, images are background

[3]https://sites.astro.caltech.edu/palomar/observer/200inchResources/lfccookbook.html#dark
[4]https://astroarchive.noirlab.edu
[5]https://portal.odi.iu.edu/index/front

subtracted using a 128 pixel background mesh unless saturated stars bleed onto a significant portion of field. In such cases, a 32 pixel background is adopted to minimize the area affected by saturation. During this step, we also mask out satellite trails from individual exposures if they overlap with the WISP field (satellite trails outside the WISP area are ignored). For convenience, the publicly released images are cropped into 5 arcmin × 5 arcmin regions centred on the WISP fields. For reference, the *HST*/UVIS FOV is 2.70 arcmin × 2.70 arcmin and the *HST*/IR FOV is 2.05 arcmin × 2.27 arcmin.

2.4 *Spitzer* IRAC imaging

Spitzer IRAC/channel 1, 3.6 μm imaging (warm mission; PI: J Colbert) was obtained for ∼200 of the deepest WISP field over multiple cycles [GO 80 134 (Cycle 8), 90 230 (Cycle 9), 10 041 (Cycle 10), 12 093 (Cycle 12)], with the primary goal of providing accurate stellar masses for galaxies in the WISP survey. For a handful of WISP fields, IRAC/channel 2, 4.5 μm data were also obtained.

2.4.1 Spitzer *data reduction and co-addition*

Roughly half of the WISP parallels are in low background fields (∼0.08 MJy sr^{-1} at 3.6 μm), while the remainder are in fields with medium backgrounds (∼0.12 MJy sr^{-1}). To achieve similar sensitivity for all the targeted fields (5σ depth of 0.9 μJy), we split the observations into two integration times, 25 versus 35 min on-source. The variation of the background with observing window does not significantly change the required exposure time. For the observations, we use a medium dither pattern with 100 s frames and either 15 or 21 exposures, depending on the total integration time. For the small subset of fields observed with the 4.5 μm filter, we always used the 21 exposure (35 min) observation.

Individual calibrated IRAC exposures are referred to as Basic Calibrated Data, or BCDs. However, IRAC BCD data contain several artefacts, including mux-bleed, mux-stripe, column droop, and bright star ghosts.[6] The IRAC pipeline therefore also produces corrected BCD images (cBCD), which attempt to mitigate these artefacts. We start with the cBCDs for all of our data reduction and analysis.

All the IRAC mosaics are generated using the Spitzer MOsaicker and Point source EXtractor (MOPEX) package. Before generating mosaics, problematic cBCDs – those with unusually high noise, extreme saturation, or other unexplained large artefacts – are removed. Such problematic cBCDs only make up a small fraction (<5 per cent) of the input cBCDs and do not significantly affect total exposure times for any observations. The first frame of all IRAC Astronomical Observation Requests has decreased sensitivity and was not included in the mosaics. The MOPEX `Overlap` pipeline subtracts the estimated background from Zodiacal light from each cBCD and then matches the background level in all frames with an additive correction. The MOPEX `Mosaic` pipeline then performs outlier rejection to remove cosmic rays, moving objects, and other artefacts, before resampling the images onto a common reference frame. Finally, the images are combined to produce a weighted mean and median mosaic along with associated coverage, uncertainty, and standard-deviation maps.

[6] https://irsa.ipac.caltech.edu/data/SPITZER/docs/irac/iracinstrumenthandbook/35/

Table 4. Summary of SExtractor parameters used on the ground-based data for determining zero-points.

Parameter	Value
`DETECT_THRESH`	1.5
`ANALYSIS_THRESH`	1.5
`DEBLEND_NTHRESH`	32
`DEBLEND_MINCONT`	0.005
`DETECT_MINAREA`	6
`BACK_SIZE`	64
`BACK_FILTERSIZE`	3
`FILTER`	Y
`FILTER_NAME`	gauss_2.0_5 × 5.conv[a]

Notes. The photometry from SExtractor is used only for constraining magnitude zero-points for the ground-based data. We use TPHOT to measure photometry for the WISP catalogue (see Section 3.3). [a]5 × 5 convolution mask of a Gaussian PSF with FWHM = 2.0 pixels.

3 WISP PHOTOMETRIC PIPELINE

The methods to obtain zero-points and photometry differ between the *HST* data (high spatial resolution) and the ground-based and *Spitzer* data (both low spatial resolution). In this section, we describe the data processing workflow used to create the WISP photometric catalogue.

3.1 Photometric zero-points

3.1.1 HST *zero-points*

The *HST* WFC3/UVIS and WFC3/IR zero-points are taken from the Space Telescope Science Institute (STScI) calibration website.[7] We note that the WFC3 zero-point solutions change slightly over time and different zero-point values are applied to data from different *HST* cycles over the course of the program.

3.1.2 Ground-based data zero-points

For all ground-based imaging, the flux zero-points were determined, in order of priority, from: (1) direct comparison to SDSS (*ugri*; 110° ≲ RA ≲ 265° and Dec. > 0°), (2) direct comparison to Pan-STARRS (*gri*; full sky for Dec. > −30°), and (3) using standard fields [i.e. fields with (1) or (2)] at different airmasses (atmospheric extinction) for individual nights of each observation. Option (3) was necessary when neither SDSS or Pan-STARRS are available for calibration. This was required for many of the southern WISP fields observed with Magellan, particularly for the *u*-band imaging.

First, we extract `psfMag` values for stars (`PhotoType` = 6) in the SDSS DR14 and Pan-STARRS DR2 public catalogues to use as our reference values. Magnitudes in Pan-STARRS were converted into the SDSS system using the relationships in Tonry et al. (2012). Second, we run SExtractor (Bertin & Arnouts 1996) with similar parameters across facilities, for consistency, on our reduced, co-added images (Section 2.3.1). A summary of the most relevant SExtractor parameters are listed in Table 4. Third, magnitude zero-points are determined by cross-matching close sources and comparing `MAG_AUTO` values from SExtractor to the catalogue `psfMag` values (both represent total magnitudes). A distance threshold of ∼1 arcsec was typically adopted, but this varied with seeing conditions (maximum of 3 arcsec).

[7] https://www.stsci.edu/hst/instrumentation/wfc3/data-analysis/photometric-calibration

The filters for all of the ground-based data (see Fig. 2b) are similar to SDSS. As a result, colour differences tend to be small, however, this was still accounted for during the zeropoint (ZP) determination. In particular, the Magellan g and r and Palomar i filters show the largest deviations from SDSS. These colour corrections are based on fitting the zero-point magnitudes as a function of $u - g$, $g - r$, $g - r$, and, $r - i$ colours for the u, g, r, and i filters, respectively.

The typical precision of the ZPs are $\sigma_{\rm ZP} \sim 0.05$ mag for Magellan/Megacam (*ugr*), WIYN/ODI (*ugr*), WIYN/MiniMosaic (*gri*), and Palomar/WaSP (g). For Palomar/LFC (*gi*), the typical precision is slightly worse, $\sigma_{\rm ZP} \sim 0.1$ mag, due to the smaller FOV with fewer reference stars for calibration. We recommend adopting a minimum uncertainty of 0.1 mag for all ground-based optical data to account for potential systematic ZP offsets.

For rare instances of overlapping data (e.g. both Magellan and Palomar data are available in the g filter), we adopt the case with better seeing (typically also deeper), unless there is a dramatic difference in depth relative to the poorer seeing case ($\gtrsim 1$ mag difference). A few comparisons between overlapping data from different telescopes with roughly similar imaging quality (depth and seeing) were possible and used as a check on zero-points and colour-corrections. We find that cases of overlapping data agree to within 1σ for ∼60 per cent–70 per cent of sources, as expected. We find that the agreement gets poorer (larger scatter) when comparing data sets with larger quality differences (e.g. 1 arcsec versus 1.7 arcsec seeing), however the median ZP offsets remain consistent with zero.

3.1.3 Spitzer *zero-points*

Similar to *HST*, the *Spitzer* zero-points are taken from the IRAC instrument handbook website.[8] We note that the WISP data were obtained during the *Spitzer* warm mission, which had a different zero-point solution to the cold mission.

3.2 *HST* PSF- and aperture-matched isophotal photometry (high spatial resolution)

The methods used to obtain the *HST* photometry are described in Bagley et al. (2017) and Henry et al. (2021). In brief, a segmentation map is first generated from the *F110W* and *F160W* detections using SExtractor on the 0.08 arcsec pixel^{-1} mosaics and then regridded to the 0.04 arcsec pixel^{-1} scale mosaics. The photometry is determined using `photutils` in Astropy (Astropy Collaboration 2013, 2018) to derive isophotal fluxes in all *HST* bands using the regridded segmentation map and includes local sky subtraction. SExtractor photometry is also performed on WFC3/IR (0.08 arcsec pixel^{-1}) images (prior to PSF-matching), in order obtain the total magnitudes (`MAG_AUTO`).

For the WFC3/UVIS data, we determine aperture corrections, `APCOR`, using the difference between `MAG_AUTO` and the isophotal magnitude, `MAG_ISO`, in the *F160W HST* filter and list these values in our catalogue (i.e. `APCOR_UVIS = MAG_AUTO_F160W-MAG_ISO_F160W`). This is similar to the method used in Henry et al. (2021) and Battisti et al. (2022). We recommend total magnitudes to be estimated for UVIS filters using $m_{\rm tot}$ = `MAG_ISO_[UVISFILTER] + APCOR_UVIS`. For the WISP subsample used in our results (Secton 7), the median value of these aperture corrections (`APCOR_UVIS`) is $\Delta m = -0.11$ mag.

[8] https://irsa.ipac.caltech.edu/data/SPITZER/docs/irac/iracinstrumenthandbook/14/

3.3 Ground-based and *Spitzer* forced photometry (low spatial resolution)

We performed template-fitting or 'forced' photometric measurements on the lower resolution data (ground-based and IRAC; both data sets follow the same procedure) by exploiting the coordinates of the *HST*-detected sources as priors. Fig. 4 demonstrates an example of the photometric fitting from the TPHOT WISP pipeline on the ground-based data.

The core of this computation is performed using the TPHOT software (v2; Merlin et al. 2015, 2016), which we summarize below, and provide in more detail in Appendix A. In brief, TPHOT requires two images at different resolutions and a list of detected objects. The algorithm transforms the low-resolution image into a (simulated) high-resolution image. To do this, TPHOT 'distributes' the flux of the sources in the low-resolution image according to the flux distribution in the high-resolution image using a deconvolution kernel (provided by the user). Compared to aperture photometry (and closest-counterpart associations), this technique allows us to obtain more accurate photometric estimates, especially when multiple *HST* sources appear blended in the low-resolution images. For all of the *HST* sources undetected in the low-resolution bands, we set their fluxes as an upper limit to the 3σ detection limit.

4 WISP PHOTOMETRIC CATALOGUE

4.1 Photometric catalogue description

The WISP photometric catalogue contains ∼230 000 sources in 439 WISP fields. These sources include stars, galaxies, and active galactic nuclei (AGNs). Fields that are absent are cases where the WISP photometric pipeline failed, which most commonly occurred due to overcrowding, bright stars, or persistence. We note that the filter availability is quite inhomogeneous among WISP fields and entries for fields without data in a given filter have values of '−99' for their photometry. We provide a summary of the number of fields with individual filters in this data release in Table 5. The photometry in this catalogue are *not* corrected for foreground Milky Way (MW) extinction.

We present an explanation of the WISP photometric catalogue entries in Table 6. These include RA, Dec., sizes and position angles (based on *HST*), and Kron-like elliptical aperture magnitude (`MAG_AUTO`; only WFC3/IR filters) from SExtractor. There are two additional aperture measurements available for *HST* data: (1) isophotes (`ISO`) based on the *F110W* and *F160W* segmentation map and (2) fixed circular apertures (`APER`; radii $r = 0.2$, 0.5, 1.0, and 1.5 arcsec). A visual representation of these apertures is shown in Fig. 5. For the ground-based and *Spitzer* data, photometry based on TPHOT is provided. For consistent 'total' photometry, we recommend using `MAG_AUTO` for WFC3/IR, `MAG_ISO + APCOR_UVIS` for WFC3/UVIS, and `AB_MAG` (TPHOT values) for ground-based and *Spitzer* data.

There are also two quality flags in the photometric catalogue. A description of these flags is provided in Appendix B. These flags should be carefully considered when selecting sources from the catalogue to avoid use of unreliable photometry.

4.2 Summary of imaging depth and completeness

Since the segmentation maps of sources from WFC3/IR imaging data are used as priors for the shallower and/or lower resolution imaging data, we only formally characterize the depth and completeness

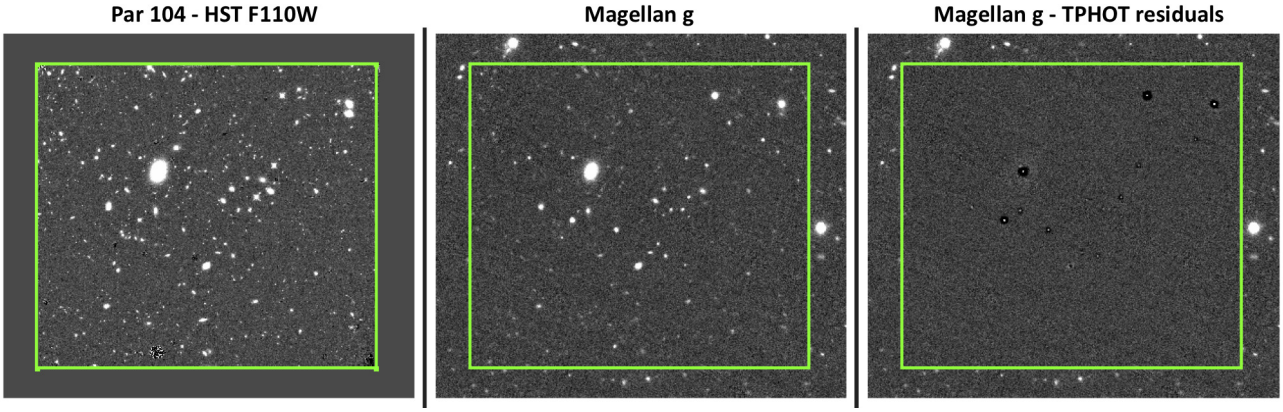


Figure 4. *HST* images and segmentation maps are used to perform 'forced' photometry on the lower resolution data (ground-based and *Spitzer*) by combining the PSFs of both *HST* and the lower resolution data (detailed in Section 3.3). As an example, we show data from Par 104: (left) the *HST F110W* image, (middle) the Magellan *g* image, and (right) the residual of the Magellan *g* image after running through TPHOT. Instances of oversubtracted residuals (black circles) are expected to occur for sources that are saturated in the *HST* image (mostly foreground stars); for these cases the *HST* (deeper) light-profile models are less accurate relative to the same sources in the unsaturated (shallower) low-resolution data. The *HST* WFC3/IR FOV (2.05 arcmin × 2.27 arcmin) is indicated by the green box in all panels.

Table 5. Summary of the number of fields, *N*, with each filter in the WISP photometric catalogue.

Filter	*u*	*g*	*F475X*	*F606W*	*r*	*F600LP*	*i*	*F814W*	*F110W*	*F140W*	*F160W*	I1	I2
$\lambda_{\rm eff}$ (μm)	0.3551	0.4686	0.4939	0.5893	0.6165	0.7444	0.7481	0.8060	1.1534	1.3923	1.5396	3.550	4.493
N	113	167	19	43	68	19	57	99	233	168	261	175	3

Notes. I1=IRAC1, I2 = IRAC2. The catalogue contains data for 439 out of 483 WISP fields. For details on individual fields, refer to Table 8.

for the WFC3/IR images. This is because, by design, every source detected in the WFC3/IR catalogue has an entry for each additional filter with data (either a detection or upper limit; i.e. 100 per cent 'complete'). Below we outline the method for estimating these values for the WFC3/IR images. We also provide a representative depth for each filter based on median 5σ magnitude (AB) for detected sources, which we also describe below.

For the *HST* imaging data, there are considerable differences in the depths and completeness achieved from field to field. This arises from the parallel nature of the program, where exposure times are set by the Primary observation. To estimate this, we use the code described in Prichard et al. (2022),[9] and briefly summarize here. First, the code determines the rms background value in 1000 randomly selected empty sky regions, where the segmentation map from SEXTRACTOR is used to exclude regions overlapping with sources. Next, it generates a histogram of the rms values and fits this distribution with a Gaussian profile to determine the sigma-clipped median. The median rms value is multiplied by the correlated pixel noise correction factor for that filter (for details, see Section 3.3 of the DRIZZLEPAC Handbook[10]). For the drizzle parameters of the WFC3/IR images, as described in Section 2.2.1, the correlated noise correction factor is 1.678. The corrected median rms is then converted into a limiting magnitude using the zero-point and a specified aperture size. We provide an estimate of depth for the WFC3/IR filters for a fixed-aperture size of 0.5 arcsec radius in Table 7.

To measure the completeness, we inject simulated sources of varying brightness into the images and estimate their recoverability with SEXTRACTOR, following methodology similar to Revalski et al. (2023). First, we inject point sources modelled as the PSF for each WFC3/IR filter. We adopt the PSFSTD models,[11] taking the average across the detector (3 × 3 array), and resampling it to match the 0.08 arcsec resolution of the drizzled images. We inject 40 sources into each image and require that these do not overlap with themselves or with real sources, as determined from the SEXTRACTOR segmentation map from the photometric pipeline. The magnitudes of the sources is scaled from [−3 mag, + 1 mag], in steps of 0.1 mag (40 steps), relative to the filter zero-point magnitude. This process is then repeated for a simulated galaxy, which is modelled as a Sérsic profile with $n = 2$, inclined at 45°, and an effective radii of 0.48 arcsec. This corresponds to the average size of a log $[M_\star/{\rm M}_\odot] = 10$ star-forming galaxy at $z = 1$ using the size–mass relation of van der Wel et al. (2014). The simulated galaxies are also convolved with each filter PSF. The resulting completeness measurements are shown in Fig. 6. The point-source completeness values show close agreement ($\lesssim 0.2$ mag) with the 5σ depth based on 0.5 arcsec apertures.

A full list of the median depth in each filter, based on median 5σ magnitude (AB; $0.15 < m_{\rm err} < 0.25$ mag) for detected sources, are provided in Table 8 and summarized in Fig. 7. We note that this depth value is the median of the entire sample of 5σ sources in each field, which span a range of sizes/surface brightness and hence is a not a uniform representation of depth. With regard to the WFC3/IR data, the depths reported here typically lie in-between the depth values for the point-source and simulated galaxies provided in Table 7. We provide some additional details on the variation in depth for the ground-based and *Spitzer* imaging data. For the ground-based data, there are considerable differences in the depths achieved from field to field. This difference is a result of the ground-based data being obtained through numerous programs, with different facilities/instruments and different science goals (see Table 3), and

[9] https://github.com/lprichard/hst_sky_rms
[10] https://hst-docs.stsci.edu/drizzpac
[11] https://www.stsci.edu/hst/instrumentation/wfc3/data-analysis/psf

Table 6. Description of WISP photometric catalogue.

Title	Description
PAR	WISP field ID number
OBJ	Object ID number
RA	Decimal RA from *HST* images (deg)
DEC	Decimal Dec. from *HST* images (deg)
A_WORLD	SExtractor profile RMS along major axis in world coordinates, measured on *HST* images (arcsec)
B_WORLD	SExtractor profile RMS along minor axis in world coordinates, measured on *HST* images (arcsec)
THETA_WORLD	Position angle measured counterclockwise from world *x*-axis, measured on *HST* images (deg)
X_IMAGE_040	Object *x* pixel coordinate in 0.04 arcsec pixel^{-1} scale image (pixels)
Y_IMAGE_040	Object *y* pixel coordinate in 0.04 arcsec pixel^{-1} scale image (pixels)
A_IMAGE_040	SExtractor profile RMS along major axis measured on 0.04 arcsec pixel^{-1} *HST* images (pixels)
B_IMAGE_040	SExtractor profile RMS along minor axis measured on 0.04 arcsec pixel^{-1} *HST* images (pixels)
THETA_IMAGE_040	Position angle counterclockwise from *x*-axis, measured on 0.04 arcsec pixel^{-1} *HST* images (deg)
X_IMAGE_080	Object *x* pixel coordinate in 0.08 arcsec pixel^{-1} scale image (pixels)
Y_IMAGE_080	Object *y* pixel coordinate in 0.08 arcsec pixel^{-1} scale image (pixels)
A_IMAGE_080	SExtractor profile RMS along major axis measured on 0.08 arcsec pixel^{-1} *HST* images (pixels)
B_IMAGE_080	SExtractor profile RMS along minor axis measured on 0.08 arcsec pixel^{-1} *HST* images (pixels)
THETA_IMAGE_080	Position angle counterclockwise from *x*-axis, measured on 0.08 arcsec pixel^{-1} *HST* images (deg)
FLUX_[APTYPE]_[HSTFILTER][a]	Sky-subtracted signal within aperture (μJy)
FLUXERR_[APTYPE]_[HSTFILTER][a]	Aperture signal uncertainty (μJy)
MAG_[APTYPE]_[HSTFILTER][a]	Magnitude for FLUX_[APTYPE]_[HSTFILTER]; −99 for undetected sources (AB mag)
MAGERR_[APTYPE]_[HSTFILTER][a]	Magnitude uncertainty; −99 for undetected sources (AB mag)
FLUX_SKY_[HSTFILTER][a]	Background used for sky subtraction in all apertures; use 10 arcsec rectangle with sources masked (μJy)
FLUXERR_SKY_[HSTFILTER][a]	Sky aperture signal uncertainty (μJy)
MAG_AUTO_[NIRFILTER][b]	SExtractor Kron-like elliptical aperture magnitude; limit reported for undetected sources (AB mag)
MAGERR_AUTO_[NIRFILTER][b]	SExtractor RMS error for AUTO magnitude; −99 for undetected sources (AB mag)
FLAG_[NIRFILTER][b]	SExtractor extraction flags (8 bit flags; see Appendix B)
AB_MAG_[FILTER][c]	Filter magnitude from TPHOT output; limiting magnitude for undetected sources (AB mag)
AB_MAGERR_[FILTER][c]	Filter magnitude uncertainty from TPHOT output; −99 for undetected sources (AB mag)
TPHOT_FLAG_[FILTER][c]	TPHOT extraction flags (3 bit flags; see Appendix B)
INSTR_FLAG_[GROUNDFILTER][d]	String indicating 'Telescope-instrument' for ground-based data (e.g. 'Magellan-Megacam')
APCOR_UVIS	Aperture correction to convert WFC3/UVIS isophotal magnitudes to total magnitude (AB mag)

Notes. [a] [ATYPE] is one of: ISO (isophotes based on the *F110W* and *F160W* segmentation map) or APER (circular apertures of radii $r = 0.2$, 0.5, 1.0, and 1.5 arcsec; array has dimensions of (n_{obj}, 4)). [HSTFILTER] is one of: *F475X*, *F475Xc*, *F606W*, *F606Wc*, *F600LP*, *F600LPc*, *F814W*, *F814Wc*, *F110W*, *F110Wc*, or *F160W*. 'c' refers to photometry on images convolved to match the PSF of *F160W*. Only convolved fluxes are available for the ISO apertures. [b] [NIRFILTER] is one of: *F110W*, *F140W*, or *F160W*. [c] [FILTER] is one of: *u*, *g*, *r*, *i*, I1, or I2. [d] [GROUNDFILTER] is one of: *u*, *g*, *r*, *i*.

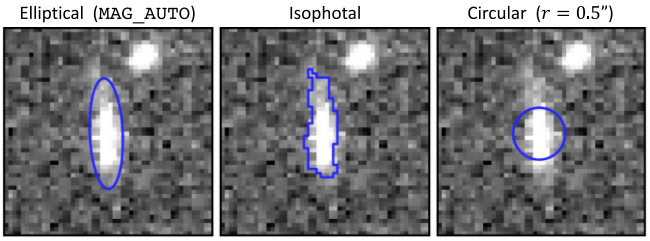


Figure 5. A demonstration of the three types of apertures available for *HST* imaging data in the photometric catalogue.

due to variations in the observing conditions. For *Spitzer*, the programs run over the different cycles (see Section 2.4) used slightly different exposures that were intended to achieve roughly uniform depth for each targeted field (see Section 2.4.1).

4.3 Comparison of ground-based catalogue photometry to SDSS

For the fraction of WISP fields that overlap with SDSS, we can make a comparison of the values in our photometric catalogue to values in the SDSS photometric catalogue. Due to SDSS being much shallower (5σ point-source depth for *u*, *g*, *r*, *i* is 22.15, 23.13, 22.70, and 22.20 mag, respectively) than our data (see Fig. 7), this comparison is quite limited (particularly for *u* band) and only possible for the brightest galaxies in each WISP field. We restrict our comparison to galaxies (SDSS PhotoType = 4) and extract PetroMag values from SDSS DR14 public catalogues to use as our reference values. The PetroMag values are the most suitable for representing total galaxy flux that will be similar to our TPHOT photometry (psfMag was used to compare stars for the zero-points in Section 3.1.2). However, the TPHOT photometry is based on priors from the deeper *HST* data such that the aperture sizes are likely to differ with respect to SDSS. A distance threshold of ∼0.5 arcsec is adopted, and we only compare sources with S/N ≥ 5 in both data sets. A comparison for the *u*, *g*, *r*, and *i* filters is shown in Fig. 8.

We find good general agreement in the photometry, with median offsets of ≲0.2 mag, although a slight bias (0.2 mag) may exist for the *r* and *i* bands. However, these offsets are within the variance for each band indicating they may not be significant. These results support or previous claim that the optical zero-points have a typical accuracy of ∼0.1 mag. Individual sources with large differences ($\Delta > 0.5$ mag) are likely due to source blending and/or incorrect cross-matching. We note that the sample available for this comparison is much smaller than the sample used in the zero-point determination for each image because the zero-point estimate made use of the entire instrument FOV (>15× the area of WFC3/IR) whereas the catalogue only includes the small region overlapping with the WFC3/IR data.

Table 7. Summary of WFC3/IR imaging depth and completeness for WISP fields in the photometric catalogue.

	F110W depth (AB mag)				*F140W* depth (AB mag)				*F160W* depth (AB mag)			
Par	*N*	Aperture	Point source	Galaxy	*N*	Aperture	Point source	Galaxy	*N*	Aperture	Point source	Galaxy
1	425	26.28	26.26	25.25	355	25.65	25.78	24.77	––	––	––	––
2	––	––	––	––	354	25.52	25.58	24.59	––	––	––	––
3	622	26.27	26.41	25.43	413	25.65	25.68	24.65	––	––	––	––
5	502	26.48	26.42	25.42	520	26.23	26.28	25.28	––	––	––	––
6	346	25.86	25.87	24.84	415	25.78	25.86	24.89	––	––	––	––
						...						

Notes. N is the number of detected objects in WISP photometric pipeline. 'Aperture' is the average 5σ depth within 0.5 arcsec sky regions. 'Point source' and 'Galaxy' are the 50 per cent completeness depths for injected simulated point sources and galaxies, respectively (see Section 4.2 and Fig. 6). A full ASCII version of this table is available online.

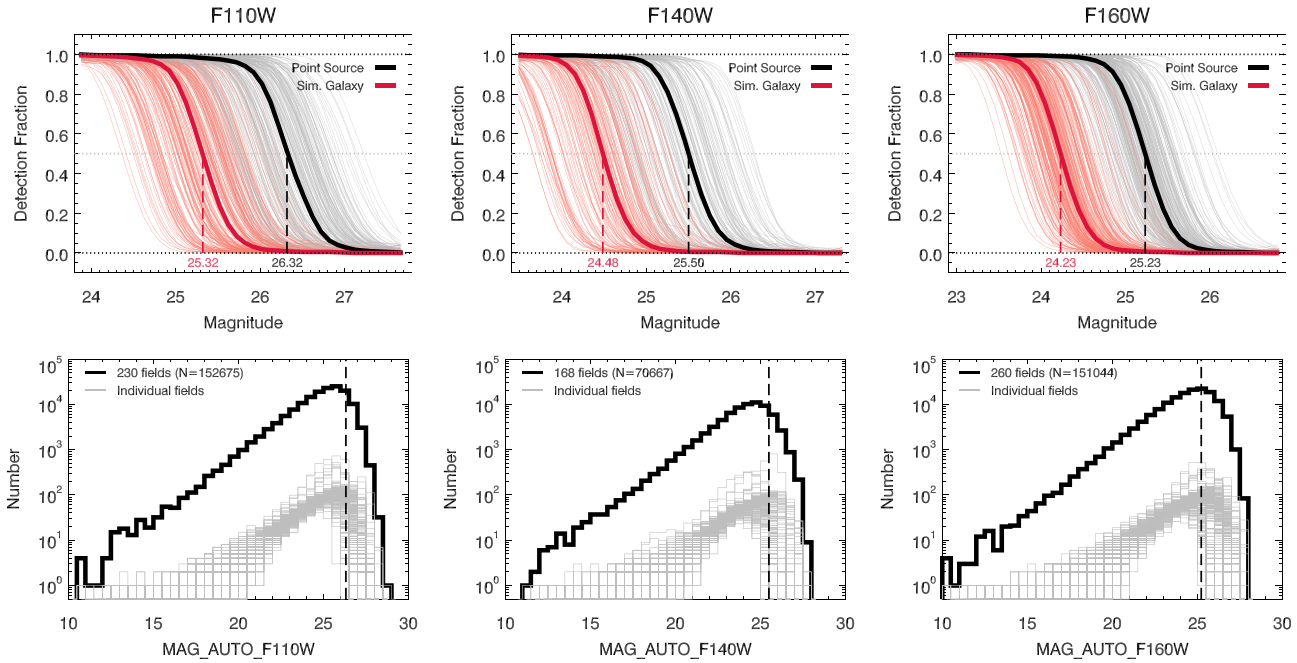


Figure 6. Top row: comparison of the fraction of artificial sources that are recovered as a function of magnitude for the WISP WFC3/IR images. The thin grey and red lines show the simulated point-source and galaxy completeness, respectively, for individual fields. The thick black and red lines are the medians of the point-source and galaxy completeness, respectively, across all fields. The completeness (adopted as the 50 per cent detection fraction) varies by up to 1.8 mag across the sample due to the varying exposure length of the visits. The completeness values for each individual field are provided in Table 7. The median 50 per cent completeness for each source type is indicated at the bottom of each panel. Bottom row: histograms of SExtractor MAG_AUTO values for the WFC3/IR images in the WISP catalogue. The black vertical dashed line denoted the median 50 per cent completeness of point sources for all fields (same as top panels).

Table 8. Summary of photometric availability and depth (AB mag) for WISP fields in this data release.

Par	*u*	*g*	*F475X*	*F606W*	*r*	*F600LP*	*i*	*F814W*	*F110W*	*F140W*	*F160W*	I1	I2
$\lambda_{\rm eff}$ (μm)	0.3551	0.4686	0.4939	0.5893	0.6165	0.7444	0.7481	0.8060	1.1534	1.3923	1.5396	3.550	4.493
1	25.06	26.29	––	––	25.82	––	26.35	––	26.02	25.59	––	23.88	––
2	––	27.16	––	––	––	––	26.28	––	––	24.94	––	––	––
3	––	26.68	––	––	––	––	26.14	––	25.91	25.42	––	23.75	––
5	24.48	26.90	––	––	24.64	––	26.23	––	26.36	25.97	––	23.53	––
6	––	26.91	––	––	––	––	26.06	––	25.68	25.38	––	23.85	––
							...						

Notes. The filters available for each field are indicated with the median 5σ magnitude (AB) for detected sources (dashes indicate no data). The *g* band and *F475X* cover roughly similar wavelengths and typically only one of these were obtained. This similarly occurs for *F606W*/*r* band, *F600LP*/*i* band/*F814W*, and *F140W*/*F160W*. I1=IRAC1 and I2 = IRAC2. A full ASCII and PDF version of this table is available online. In the full PDF version, fields with 'full' UV to NIR coverage (6 or more filters) are highlighted in orange.

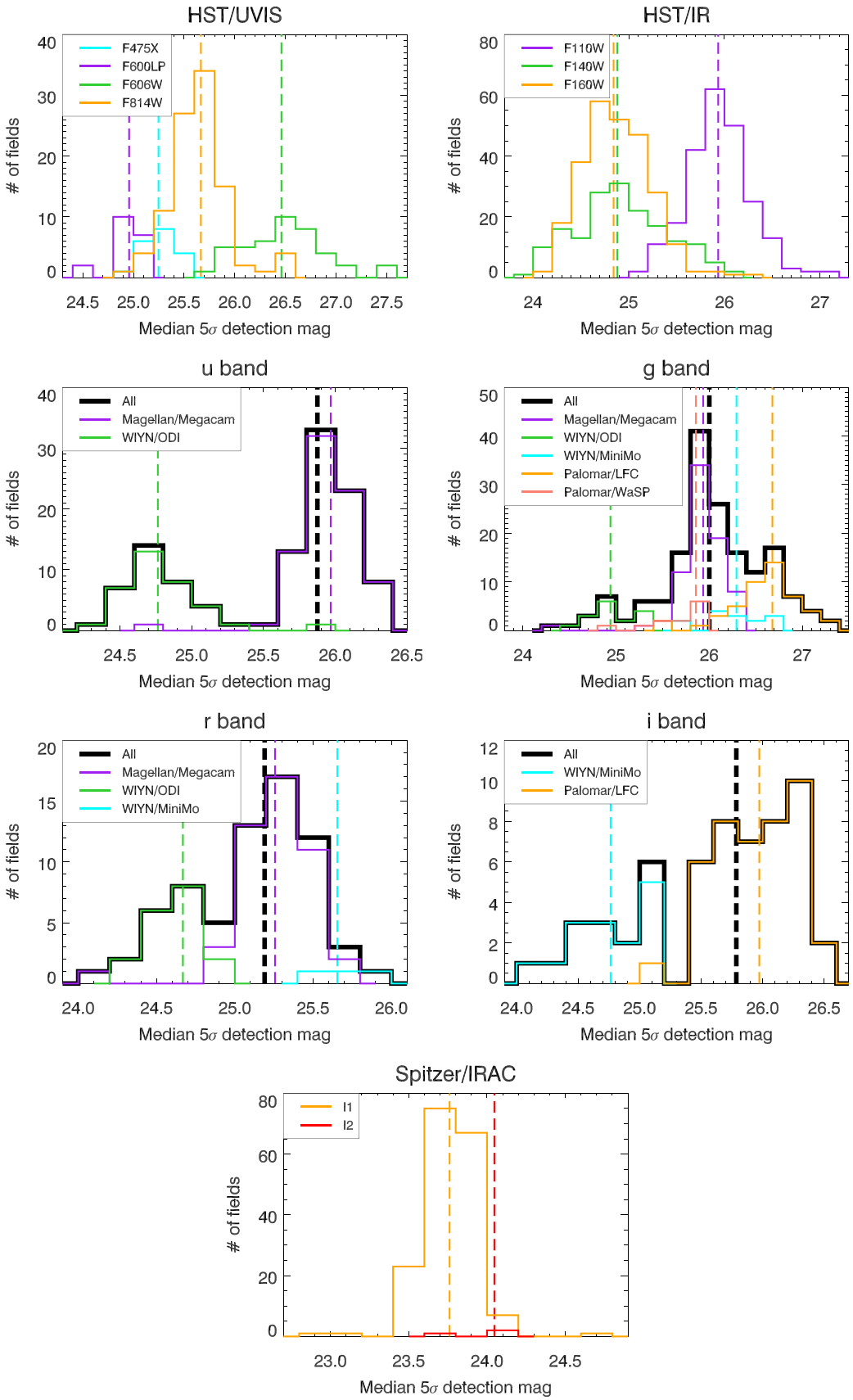


Figure 7. Comparison of the median magnitude for sources detected at $\sim 5\sigma$ significance (sources with $0.15 < m_{\rm err} < 0.25$ mag; S/N $\sim$ 5) for all WISP fields in this data release. Panels show (left to right and top to bottom) *HST* WFC3/UVIS, WFC3/IR, ground-based *u*, *g*, *r*, *i*, and *Spitzer*/IRAC. The vertical dashed lines denote the median magnitude value of 5σ-detected sources for a given filter (i.e. image depth). The range in depths are mainly due to the varying length of exposure times, observing conditions, and/or field crowding (see Section 4.2).

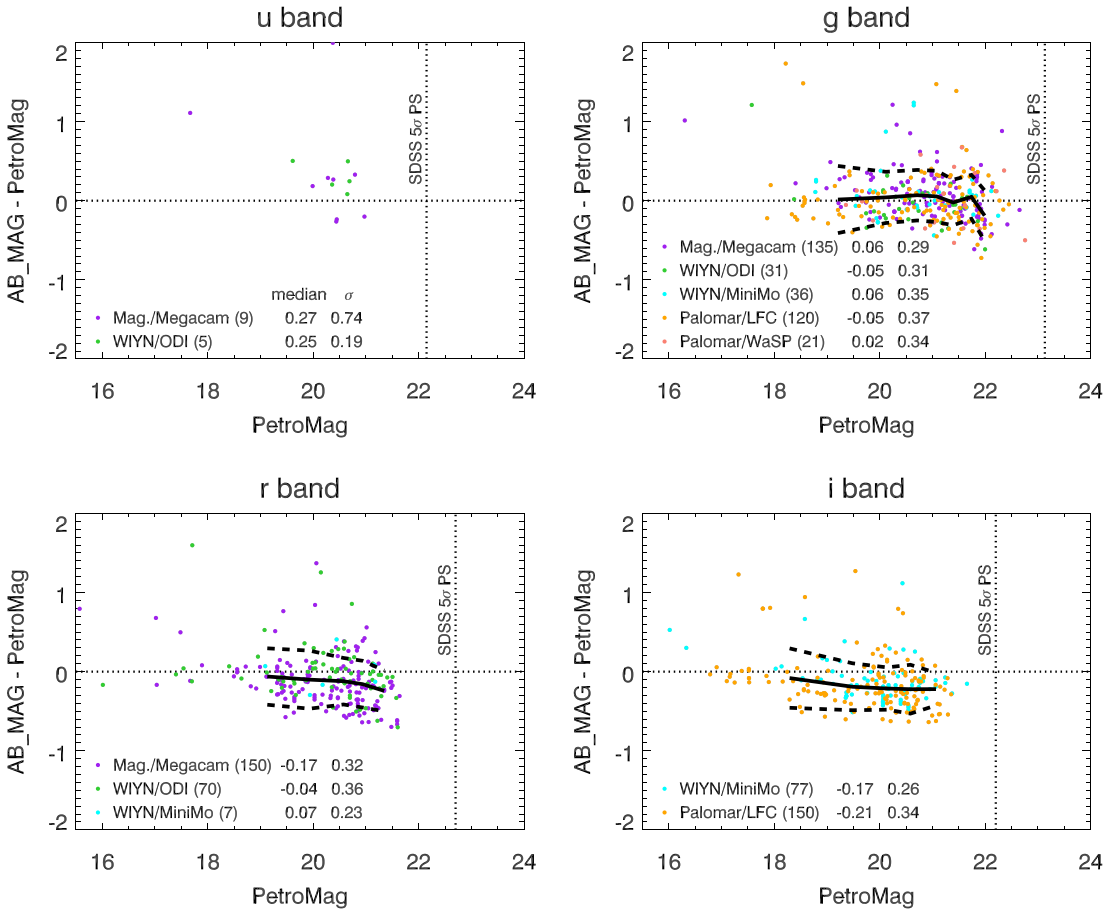


Figure 8. Comparison of the magnitude offset for galaxies detected at S/N ≥ 5 in optical bands from SDSS (`PetroMag`) and our WISP catalogue (`AB_Mag`). The vertical dashed lines denote the median 5σ point-source limit for SDSS. We note that the magnitude limit for galaxies is fainter than this (e.g. Fig. 6). The median offsets are typically $\lesssim$0.2 mag and agree within the variance of the data. There are very few *u*-band galaxies bright enough for reliable comparison. The solid and dashed lines shows the median and $\pm 1\sigma$ values combining all data in equal-number bins of 50 galaxies. For each instrument, we provide the median magnitude difference (i.e. *y*-axis) and its variance in the legend, with the number of available sources shown in parentheses.

5 WISP SPECTROSCOPIC PIPELINE

The WISP emission-line catalogue is created in three steps. First, an automatic detection algorithm examines every extracted spectrum to identify emission-line candidates. Next, each candidate is independently inspected by two reviewers both for confirmation and to measure the source redshift and emission-line properties. These first two steps are performed by the `wisp_analysis`[12] software package. Finally, the output from the two reviewers is combined to create one catalogue entry for each emission-line source. We describe each step in this process in the following sections.

5.1 Emission-line candidate identification

We identify emission-line candidates with a peak detection algorithm that uses a continuous wavelet transform (CWT) to select appropriately shaped peaks in 1-D spectra. A wavelet transform breaks a signal into its base components, each of which is a modified version of the same 'mother' wavelet function. This process is similar to Fourier analysis, but rather than sinusoidal components of varying frequencies, the base components identified by a wavelet analysis are scaled or shifted versions of the mother wavelet. The CWT is an improvement over the previous WISP line detection method, which identified emission lines as contiguous pixels above a signal-to-noise (S/N) threshold (Colbert et al. 2013). The previous amplitude-based peak finding process was strongly dependent on the fit to the continuum and the amount of smoothing applied to the spectrum. It also resulted in many spurious detections as noise spikes can be misidentified as faint emission lines. The CWT technique identifies emission-line features in the spectrum based on their shape as well as their amplitude, and therefore reduces the number of spurious emission-line candidates that reviewers need to inspect. A detailed description of the steps in applying our CWT algorithm and using the results to identify emission-line candidates is provided in Appendix C.

Emission-line identification using the CWT algorithm is performed on the 1D spectra extracted and calibrated by AXE (Kümmel et al. 2009). We first remove from each spectrum the flux introduced by the overlapping continua from nearby objects. For this purpose we use the contamination model created by AXE during spectral extraction, where all objects detected in the direct images are approximated as 2D Gaussians defined by the size and shape measured by SExtractor. This process is performed on the 1D

[12] https://github.com/HSTWISP/wisp_analysis

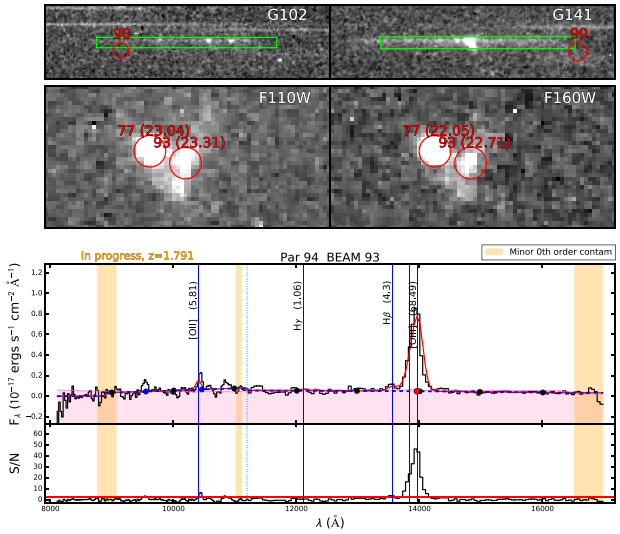


Figure 9. The display from the interactive portion of the emission-line finding process. Reviewers inspect all available information for a given object. Top row: the 2D spectral stamps for *G102* (left) and *G141* (right) are displayed with a green box indicating the trace of the spectra. Red circles identify the positions of zeroth orders of bright ($m < 23$) objects that may contaminate the spectra. Middle row: the direct images in all available filters, *F110W* (left) and *F160W* (right), with the positions of all objects and their corresponding magnitudes marked. The source of interest, here object 93, is centred in the direct images. Bottom panel: the 1D spectrum is displayed in both flux units (f_λ, top) and S/N (bottom). In the example shown here, the emission features identified in the spectrum of object 93 are not centred in the green box because they are coming from object 77. The emission-line offset is clearest in the *G102* spectrum. Object 93 was rejected as an emission-line candidate and these emission features were instead measured for object 77.

spectra of each grism individually, allowing for 500Å of overlap between the two grisms. We next require that at least two reviewers inspect each emission-line candidate, both to confirm the candidates and to measure the line properties.

5.2 Emission-line candidate inspection

The original, amplitude-based WISP line detection software resulted in a false detection rate of $\gtrsim 70$ per cent, depending on the depth of the fields. Although the improvements presented here – identifying emission lines using both amplitude and shape combined with the EW criterion – reduce this rate significantly to approximately 55 per cent, the visual inspection remains necessary for constructing a clean catalogue. This inspection is especially crucial for undithered parallel data, which cannot be properly cleaned of artefacts.

For each emission-line candidate, reviewers simultaneously inspect the direct images of the source, the 2D spectral stamps, the 1D extracted spectrum in units of f_λ, and a spectrum in units of S/N. An example of the full display from the interactive portion of the process is shown in Fig. 9.

We now briefly describe the inspection process and the series of checks reviewers perform for each candidate. The source displayed in Fig. 9, object 93 from WISP field Par94 (hereafter 94–93), is used as an example to illustrate the process. The two main questions reviewers must answer are whether the emission-line candidate identified by the detection algorithm is real, and whether it belongs to the object to which it is associated. In the first case, reviewers are validating the results of the detection software. In the second, they are considering and ruling out possible sources of contamination.

Candidate confirmation is required because spurious or false emission lines are occasionally identified by the detection program. False emission lines typically arise in two cases, when detector artefacts are identified as lines and when the continuum is improperly fit. The first case is a particular problem for a pure parallel survey such as WISP, since the telescope is not dithered between exposures. Combining multiple exposures therefore does not remove all cosmic rays, hot pixels, and other artefacts. Those that remain are sometimes selected as emission-line candidates. Reviewers can often reject these by comparing the shape of the emission-line candidate in the 2D spectral stamp (top row of Fig. 9) with the source shape in the direct image (middle row). Recall that an emission line in the spectral stamp is an image of the source at the given wavelength. The size, ellipticity, and position angle of the source are expected to be reflected to some degree in the emission line. This comparison is approximate, however, since the emission regions in a galaxy need not directly correlate with the broadband continuum emission detected in the imaging filters.

In the second case, false emission-line candidates are identified in objects with continua that are poorly fit by the automatic software, which uses a cubic spline to fit the spectrum at eight wavelengths nodes. A steep rise in the continuum of an object, often caused by contamination from the spectrum of a nearby object, that is not reflected in the model fit can be incorrectly selected as a spectral peak. Given the large range of object sizes, fluxes, and levels of spectral confusion, one set of parameters will not work perfectly for all objects. The software's continuum fit is therefore treated as a first pass. The dashed blue line in Fig. 9 is the continuum model for object 94–93, which in this case represents a good fit to the observed continuum. For objects with improperly fit continua, reviewers can tweak the model by adding, removing, or changing the wavelengths of the nodes used in the spline fit (black circles).

The reviewers must next determine whether the identified emission feature belongs to the source in question. With only a single roll angle, WISP spectra are often contaminated by overlapping spectra from other sources along the dispersion direction. This check generally involves four parts. First, the comparison between source and emission-line shape described above can help identify emission lines coming from another object. We do not expect a galaxy to be much larger or much brighter in an emission line than it is in imaging.

Second, the emission should be vertically centred in the trace of the spectral stamps, indicated by the green box in the top panels of Fig. 9. We can see that the line candidates in the spectrum of 94–93 are not centred, evidence that they likely belong to the nearby, brighter object 77. We note that it is of course possible that the emitting region of a source may not be centred on the continuum emission, and by rejecting emission lines that are off the centre of the trace we may also be rejecting real emission-line galaxies.

Third, if there are multiple emission lines visible in the spectrum, their relative wavelengths should match. The wavelength solution of the grism is determined by the source position in the direct image, and will therefore only be correct for the spectrum of that source. For example, in Fig. 9, given the assumed redshift for this object, the spectral peak around $\lambda \sim 10\,500$ Å should be [O II]λ3727. However, it does not line up exactly with the expected wavelength for [O II] at this redshift (indicated by the blue vertical line), further indication that these emission lines are contaminants from object 77.

Finally, the reviewers must consider the position of zeroth orders. We consider a portion of a spectrum to suffer from 'major' zeroth-order contamination if it directly overlaps with a zeroth order from a bright source ($m < 23$ mag). The position of these bright zeroth orders are indicated by red circles in the grism stamps in Fig. 9. The

Table 9. Model spectrum parameters.

Parameter	Description	Initial value	Limits
$n_{\rm nodes}$	Number of spline nodes used in continuum fit	8	Fixed
$\Delta\lambda_{\rm fitting}$	Size of region used for line fitting	1500 Å	Fixed
$z_{\rm init}$	Input redshift	$\left(\frac{\lambda_{\rm obs}}{\lambda_{\rm H\alpha}}-1\right)^a$	±0.02
$\Delta z_{\rm [O\,III]}$	Shift in redshift allowed for [O III] profile fit	0	±0.02
$\Delta z_{\rm [O\,II]}$	Shift in redshift allowed for [O II] profile fit	0	±0.02
$\Delta z_{\rm [S\,III,\,He\,I]}$	Shift in redshift allowed for [S III] and He I fits	0	±0.02
$\rm FWHM_{init}$	Input FWHM	$2a\,\Delta\lambda_{G141}$ Å^b	$-0.3\rm FWHM_{init}$ $+2.0\rm FWHM_{init}$
$A_{\rm line}$	Input amplitude for each emission line	…^{c}	$0,1^c$
$r_{\rm [S\,III]\lambda9532/\lambda9069}$	Ratio of [S III]λ9532 to [S III]λ9069 fluxes	2.48	Fixed
$\lambda_{\rm T}$	Grism transition wavelength	11 200 Å	Fixed
$\lambda_{\rm blue}$	Blue wavelength cut-off for the *G102* grism	8100 Å	Fixed
$\lambda_{\rm red}$	Red wavelength cut-off for the *G141* grism	17 000 Å	Fixed

Notes. [a]The input redshift is that which will give Hα for the current line, or is a redshift guess provided by the user.
[b]$\rm FWHM_{init}$ is taken as twice the source semimajor axis (A_IMAGE) multiplied by the dispersion in the red grism.
[c]Emission-line amplitudes are estimated as the flux value at line centre and are constrained to be positive.

chance alignment of a bright zeroth order, especially from a compact source, can appear as a very convincing emission line. Spectral peaks that suffer from this major contamination are automatically rejected by the automatic software. Meanwhile, 'minor' zeroth-order contamination could be caused by (1) bright zeroth orders that are close to but not directly overlapping the spectral trace or (2) from direct overlap with zeroth orders of fainter sources ($m > 23$). In some cases, again especially for the most compact objects, these fainter zeroth orders can be bright enough to masquerade as emission lines, and so the reviewers must remain vigilant for this possibility. The minor zeroth-order contamination regions for object 94–93 are shown as orange bands in the spectrum of Fig. 9.

5.3 Emission-line measurements

Once an emission-line candidate has been confirmed, the reviewers fit a model to the spectrum. The fitting is performed via Levenberg–Marquardt least-squares minimization. It is implemented with the software MPFIT, based on the MINPACK-1 FORTRAN package (Moré 1978) and translated to PYTHON by Mark Rivers.[13] With MPFIT, each parameter can be held fixed or can be constrained with upper and lower bounds. The full model includes over 20 parameters, which are described below and listed in Table 9.

The model spectrum fit to the data consists of a continuum with the following emission lines added: [O II]λλ3727 + 3729, Hγ, Hβ, [O III]λλ4959 + 5007, Hα, [S II]λλ6716 + 6730, [S III]λ9069, [S III]λ9531, and He Iλ10830. The continuum is modelled using a cubic spline fit to a series of $n_{\rm nodes}$ spectral nodes, and emission lines are modelled as Gaussians, where the line centre is determined by the redshift guess for the source, the input amplitude is estimated separately for each line within $\pm\Delta\lambda_{\rm fitting}$ of line centre, and the standard deviation depends on the source size and the dispersion of the grism in which the line appears. Emission lines are not fit individually. The entire spectrum – continuum plus all lines – is fit simultaneously, and all line profiles are constrained to have the same FWHM (full width at half-maximum) *in pixels*, not in Å. Hence, lines in the higher dispersion *G102* will have smaller FWHM by a factor of two. This approach is reasonable under the assumption that all emission lines are images of the same host source. The source redshift is determined by the centre of the profile fit to the Hα line at $z \lesssim 1.6$ and [O III] at $z \gtrsim 1.6$ (when Hα has redshifted out of the *G141* grism). The centres of each additional line are allowed to vary up to a maximum wavelength equivalent to $\Delta z_{\rm line} = 0.02$. In the absence of multiple emission lines, single lines are assumed to be Hα unless the clear asymmetry of the [O III]+Hβ line profile is visible.

As a consequence of the use of a full spectral model, all emission lines listed above are fit provided they fall within the grism wavelength coverage at the assigned redshift. Emission lines that were not identified by the detection algorithm will therefore be measured along with the identified lines. We refer to the lines strong enough to have been identified by the detection algorithm as 'primary lines', while the remaining lines are called 'secondary lines'. This distinction is an important one for the emission-line catalogue completeness, which is discussed in the next section (Section 5.4).

Lines with a flux S/N < 1 are set at 1σ and reported as upper limits in flux and therefore EW. These limits are calculated by summing in quadrature the error array within 2 ×FWHM of line centre. We find, however, that the error arrays calculated by AXE are underestimates of the spectral noise properties. The 1σ limits are all systematically lower than the sensitivity limits measured for the fields. We therefore apply a correction factor to the flux limits, correcting the amplitude of the limits while preserving the scatter in the measurements.

We note several emission lines – Hα and [N II], the [S II]λλ6716 + 6731 doublet, the [O II]λλ3727 + 3730 doublet, and the [O III]λ4959 + 5007 doublet – are blended at the resolution of the WFC3 grisms. The fluxes measured for Hα therefore include the contribution from [N II]λ6583 and [N II]λ6550, and those for [O II], [S II], and [O III] each include both doublet lines. The flux ratio [S III]λ9532/[S III]λ9069 is fixed to 2.48:1.

While inspecting each spectrum, reviewers can change multiple parameters in order to improve the fit to the spectrum. In addition to moving, adding, or subtracting nodes for the continuum fit, they can provide a new redshift guess for the source; modify the wavelength ranges of each grism to fit emission lines at the grism edges (changing the transition between grisms, $\lambda_{\rm T}$, and their wavelength cutoffs $\lambda_{\rm blue}$ and $\lambda_{\rm red}$); or provide a new guess for the $\rm FWHM_{init}$, usually decreasing the default guess for sources where the dispersion axis is along the minor axis. Finally, reviewers can mask regions of the spectrum that suffer from severe contamination from either zeroth orders or nearby continua, thereby making sure they do not affect the full spectral model. The emission-line candidate vetting, spectral

[13]University of Chicago, http://cars9.uchicago.edu/software/python/mpfit.html

fitting, and cataloguing are all performed as part of one streamlined process. The results from the two reviewers are then combined for each object, and a series of quality flags are assigned to the emission-line object determined by the level of redshift, flux and EW agreement between the reviewers' classifications (see Section 6 and Appendix E).

5.4 Completeness corrections

The completeness of a survey or catalogue is a description of how accurately the detected sample represents the true population in the universe. Understanding a survey's completeness is necessary before the results can be used to conclude anything about the true underlying distribution of sources and source properties. A survey such as WISP can suffer from incompleteness for a variety of reasons. Sources may be lost amidst the noise in images if their fluxes are close to the detection limit. Some sources may not be detected, or their emission lines missed in their spectra, because they overlap or are blended with nearby bright objects. The completeness of a survey depends on the specific selection function used to detect sources. In the case of the WISP emission-line catalogue, the selection function includes the detection of sources in the direct images, the identification of emission-line candidates via the detection algorithm, and the acceptance during the visual inspection. We also only look for emission lines of continuum-detected sources in the images. We must understand the fraction of sources and emission lines that are not detected through this full process as a function of their properties such as size, shape, and the strength of their emission.

In quantifying the completeness corrections that must be applied to a catalogue, we are determining the types and numbers of sources that are missed. To do so, we create a simulated catalogue of 10 000 sources and their spectra, 5000 each for the shallow and deep portions of the WISP Survey. The simulated sources have H-band magnitudes in the range $16.8 \leq m_{\rm H} \leq 26.2(27.6)$ and observed Hα fluxes in the range of $5 \times 10^{-17}(1 \times 10^{-17}) \leq f \leq 1 \times 10^{-15}$ erg s^{-1} cm^{-1} for the shallow (deep) fields. See Table D1 for the full list of input parameters and values. We insert these simulated sources into real WISP images, 25 sources at a time, and run them through the full WISP pipeline and emission-line detection software. The creation of the simulated data is described in Appendix D. In order to save time and the effort required during the visual inspection stage, the reviewers only inspect the spectra of simulated sources that were identified by the line finding algorithm. Yet not all of the emission-line candidates were real. Some were noise spikes, contamination, or the result of poorly fit continua. We note, however, that because of this choice, we cannot use the simulations to measure the rates of contamination or redshift mis-identification in the catalogue.

Of the 10 000 input sources, 7721 were recovered by the WISP reduction pipeline, with an equal number recovered in the shallow and deep fields. This 77 per cent recovery fraction represents the overall imaging completeness given the imaging depths and our set of source detection parameters. The majority of the sources that are not recovered in the imaging catalogue, and which therefore have no extracted spectra, are faint and/or extended. In Fig. 10, the input semimajor axis sizes (before convolution with the PSF) are shown as a function of magnitude for the simulated sources that were input and recovered. The distributions of real sources are shown for reference. The sources that are not recovered in imaging mainly have a semimajor axis of $a \geq 0.7$ arcsec and are fainter than 24.5 mag in the H band.

For the simulated sources recovered in imaging, we calculate the fraction of these that are recovered by the full line finding process.

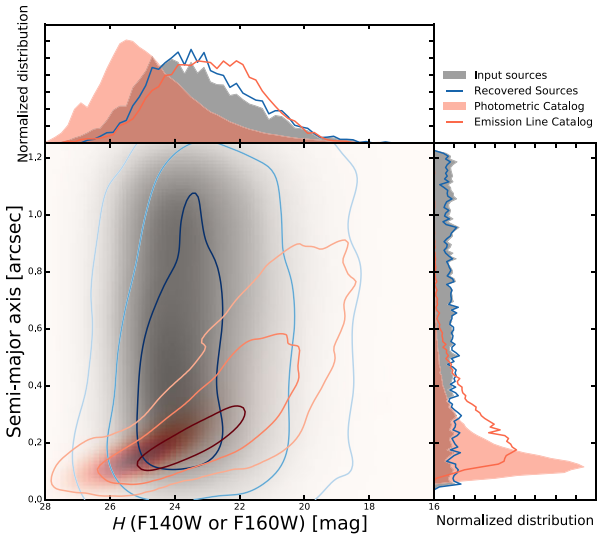


Figure 10. The semimajor axis as a function of the H-band magnitude for all input sources (grey) and those recovered in imaging (blue contours). The real sources from the photometric and emission-line catalogues are shown in red. The imaging completeness is a function of the magnitude and size of the sources. The recovery fraction drops for large ($a \geq 0.7$ arcsec) and faint ($H > 24.5$) sources.

We find that the completeness depends on source size and shape and emission-line EW and S/N. The S/N dependence is essentially a dependence on line flux, but includes the effects of the variable depths of the WISP fields (see Fig. 3). While we can measure the S/N of observed emission lines, there is no analogous definition of the input S/N for the simulated lines. The input template spectra do not include noise, and while we have added shot noise to the simulated grism data based on the integration times of the exposures into which they are added, this is not the only source of noise that will affect the flux measurements. The depths reached in each field depend also on the level of the zodiacal background for each pointing. We therefore instead characterize the completeness as a function of emission line 'scaled flux', or the emission-line flux (input or recovered) divided by the sensitivity limit of the field at the wavelength of the line.

As discussed in Section 5.3, a source will enter the catalogue because of the detection of the primary lines. We consider only one line per spectrum – both for the input simulated and the output measured spectra – taking the line with the brightest scaled flux as the source's primary line. As we are not attempting to quantify the rates of redshift misidentification, we consider a detected emission line recovered regardless of whether the reviewers have properly identified it (i.e. regardless of what redshift is assigned to the object). We recover 868 of the 5000 simulated sources added to deep WISP fields covered by both grisms and 1541 of the 5000 added to shallow, *G141*-only fields. This recovery reflects the completeness due to both the imaging and the spectroscopic selection functions, and is heavily influenced by source size and shape as discussed below.

The object size and shape will strongly affect the completeness, as large, low surface brightness emission lines may be missed by the peak finder. However, the large sources that suffer from the highest levels of incompleteness, those with $a \geq 0.7$ arcsec, constitute less than 1 per cent of the total catalogue. We simulate sources with a uniform distribution of sizes, but then weight the input sources by the distribution of observed sizes in the emission-line catalogue. This step both reflects the observed distribution and allows us to consider a two-parameter completeness correction, maintaining

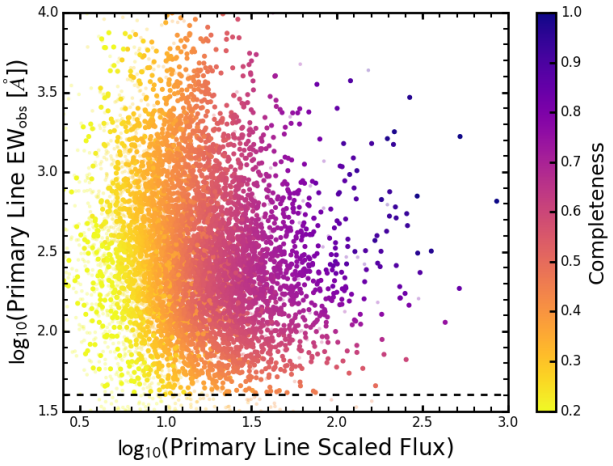


Figure 11. The completeness of the emission-line catalogue as a function of the scaled flux and observed EW of the strongest line for each source. The dashed horizontal line indicates the EW completeness limit 40 Å. The small, transparent points indicate the lines with EW < 40 Å and/or S/N < 5.

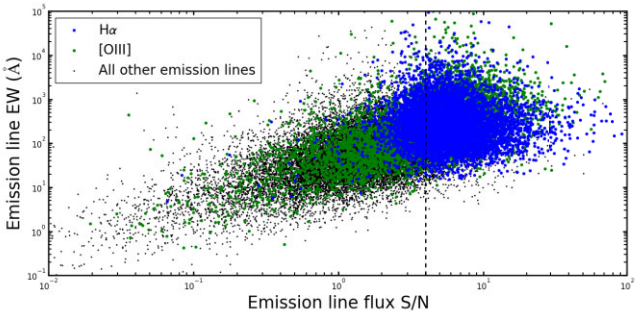


Figure 12. The emission-line S/N and EW of all emission lines in the catalogue. Hα is the two most commonly identified 'primary line' and therefore the majority of Hα in the catalogue has an S/N > 4. The majority of the lower-S/N lines were fit as a consequence of the redshift assigned by identification of a primary line.

sufficient number counts for the completeness analysis without requiring reviewers to visually inspect tens of thousands of sources.

The completeness is calculated in four bins of scaled flux and five bins of EW. The bin edges are determined by the distribution of sources in the real WISP emission-line catalogue such that there are an approximately equal number of real sources in each bin. The one exception is the bin of lowest EW, which we add in order to probe an area of the parameter space with low completeness (EW_{obs} < 40 Å, see Colbert et al. 2013). We use a radial basis function to approximate the 3D surface formed by the bin centres and recovery fractions calculated in each bin. The resulting completeness corrections calculated for each source in the WISP emission-line catalogue are shown in Fig. 11. They are applied according to the scaled flux and observed EW of the strongest line in the spectrum, which is most often Hα or [O III], but is occasionally [O II] when Hα or [O III] are masked out due to overlap with a bright zeroth order or other major contaminant. These completeness corrections are applicable at the source level or for the primary lines for each source. They are not appropriate for secondary lines. As expected, very few low-EW emission lines were recovered in the simulations, making the completeness corrections calculated in bins with EW_{obs} < 40 Å very uncertain. We therefore find that the EW criterion presented in Colbert et al. (2013) applies to the new version of the line finding process as well (see Bagley et al. 2020 for a comparison of these completeness corrections with those calculated by Colbert et al. 2013). We also recommend applying an emission-line flux S/N cut at S/N>5 when using this catalogue.

6 WISP EMISSION-LINE CATALOGUE

6.1 Emission-line catalogue description

The WISP emission-line catalogue contains 8192 sources in 419 WISP fields. These sources include only galaxies and AGN (spanning $0.137 < z_{grism} < 2.785$; no stars included). Fields that are absent, relative to the photometric catalogue, fall into one of a few categories: they were too crowded (a few WISP fields captured portions of dwarf galaxies), heavily contaminated by bright stars, or suffered from poor and uneven background subtraction in the grism data. We note that line availability is dependent on the grism coverage and galaxy redshift and that sources without data in a given emission line have values of '−1' for their entries. The S/N and EW of all emission lines in the catalogue are shown in Fig. 12. The vast majority of the Hα emission lines (blue circles) have a S/N greater than 4 (vertical dashed line), indicating that Hα is most often the primary line in a spectrum. Emission-

line fluxes to the left of the dashed line were likely fit as secondary lines and are therefore at a lower S/N. The emission lines in this catalogue are *not* extinction-corrected (neither foreground MW nor internal extinction applied).

We present an explanation of the WISP emission-line catalogue entries in Table 10. These include RA, Dec., WFC3/IR total magnitudes, sizes (based on *HST*), spectroscopic redshift, and the FWHM used for all emission lines. There are also emission-line flux, flux error, EW, and observed wavelength, for nine lines ([O II], Hγ, Hβ, [O III], Hα + [N II], [S II], [S III]λ9069, [S III]λ9532, and He Iλ10830).

There are also nine quality flags in the emission-line catalogue. A description of these flags is provided in Appendix E. These flags should be carefully considered when selecting sources from the catalogue. For the most robust selection of emission-line sources, we recommend only using cases with `REDSHIFT_FLAG` = 0 (5054 sources) to avoid cases with ambiguous redshift determinations. Cases with `REDSHIFT_FLAG` > 0 indicate some type of redshift disagreement between the reviewers or redshift uncertainty for single line emitters. The majority of sources with `REDSHIFT_FLAG` > 0 are those that were identified by only a single reviewer (`REDSHIFT_FLAG` ≥ 16; 2061 sources). These cases are likely marginal detections where one reviewer did not consider the identified emission feature to be real above the noise or rejected the emission feature as some type of artefact or contamination. Sources with `REDSHIFT_FLAG` ≥ 16 should therefore be considered with caution.

6.2 Grism redshift accuracy and precision

The redshift accuracy of the grism data are primarily driven by the number of available lines, with a greater number of lines generally providing more reliable redshifts. For example, Baronchelli et al. (2020) show that the default choice of assuming WISP single-line emitters are Hα is incorrect for ∼30 per cent of cases, where most of these are likely to be the [O III]λ5007 emission line. Therefore, in the absence of other independent information that can inform how to break single-emitter degeneracies, such as photometric redshifts (photo-z) and/or machine learning (e.g. Baronchelli et al.

Table 10. Description of WISP emission-line catalogue.

Title	Description
PAR	WISP field ID number
OBJ	Object ID number
RA	Decimal RA from *HST* images (deg)
DEC	Decimal Dec. from *HST* images (deg)
JMAG	*J*-band magnitude; SExtractor MAG_AUTO_F110W from WISP photometric catalogue (mag)
HMAG	*H*-band magnitude; SExtractor MAG_AUTO_[F140W or F160W] from WISP photometric catalogue (mag)
A_IMAGE	SExtractor profile RMS along major axis measured on 0.08 arcsec pixel^{-1} *HST* images (pixels)
B_IMAGE	SExtractor profile RMS along minor axis measured on 0.08 arcsec pixel^{-1} *HST* images (pixels)
FILTER_FLAG	Flag identifying filter coverage for field (10 bit flags; see Appendix E)
GRISM_FLAG	Integer flag identifying grism coverage for field; (1=*G102*, 2=*G141*, and 3 = *G102* + *G141*)
[GRISM_FILTER]_FLAG[a]	Flags identifying any problems with the grism data for this field (5 bit flags; see Appendix E)
EDGE_FLAG	Flag identifying objects close to edges of direct image (5 bit flags; see Appendix E)
NNEIGHBORS	Number of sources in WISP photometric catalogue within 1 arcsec of object
NLINES	Number of lines detected at $>2\sigma$ for object
REDSHIFT	Redshift from simultaneous fit to all lines in spectrum
REDSHIFT_ERR	Redshift 1σ uncertainty
dz_OIII	Shift in redshift of the [O III] profile fit compared with reported REDSHIFT for object
dz_OII	Shift in redshift of the [O II] profile fit compared with reported REDSHIFT for object
dz_SIII_HE1	Shift in redshift of the [S II] and He I profile fit compared with reported REDSHIFT for object
DELTA_REDSHIFT	Difference in redshift fits of the two reviewers
REDSHIFT_FLAG	Flag identifying quality of redshift fit (6 bit flags; see Appendix E)
FWHM_OBS	FWHM used for all emission-line profile fits; observed frame using the 46.5 Å pixel^{-1} dispersion in *G141* grism (Å)
FWHM_OBS_ERR	FWHM 1σ uncertainty (Å)
FWHM_FLAG	Flag identifying quality of reported FWHM (5 bit flags; see Appendix E)
COMPLETENESS	Source completeness reflecting the selection function of the full line-finding procedure (see Section 5.4)
[LINE]_NREVS[b]	Number of reviewers who measured emission line
[LINE]_FLUX[b]	Emission-line flux from profile fit (erg s^{-1} cm^{-2})
[LINE]_FLUX_ERR[b]	Emission-line flux 1σ uncertainty (erg s^{-1} cm^{-2})
[LINE]_DELTA_FLUX[b]	Difference in emission-line fluxes of the two reviewers (erg s^{-1} cm^{-2})
[LINE]_EW_OBS[b]	Emission-line EW reported in observed frame; -1 for undetected lines (Å)
[LINE]_DELTA_EW[b]	Difference in emission-line EWs of the two reviewers (Å)
[LINE]_FLAG[b]	Flag identifying quality of emission-line measurements (6 bit flags; see Appendix E)
[LINE]_CONTAM[b]	Flag identifying the contamination noted by each reviewer ('a' and 'b'); String has form 'a.b' (4 bit flags; see Appendix E)
[LINE]_WAVEOBS[b]	Observed wavelength of emission line (Å)
[LINE]_EDGE_FLAG[b]	Flag identifying emission lines close to edges of grism where sensitivity decreases (4 bit flags; see Appendix E)

Notes.[a] [GRISM_FILTER] is one of: *G102* or *G141*. [b] [LINE] is one of: O II, Hg, Hb, O III, HaN II, S II, S III_9069, S III_9532, or He1_10830.

2020, 2021), we urge caution in using sources with only a single emission line (NLINES ≤ 1; 3350 sources in catalogue). It is also important to consider the redshift quality flags (REDSHIFT_FLAG; see Appendix E) that indicate potential disagreements between emission-line reviewers.

For sources with multiple detected emission lines (i.e. NLINES > 1; 4842 sources in catalogue), the reliability of the redshifts is expected to higher, however this is difficult to quantify without independent metrics or follow-up higher resolution spectroscopy. We refer readers to Baronchelli et al. (2020, 2021) for a detailed discussion on the issues pertaining to contamination and purity in a grism spectroscopic sample. Masters et al. (2014) presented a subset of 26 emission-line galaxies (from 23 WISP fields) where follow-up observations with Magellan/FIRE ($R \sim 6000$; 0.8–2.5 μm) were made on sources with S/N $\gtrsim 10$ (from grism) in [O III] and/or Hα. They found very good agreement in line identification (close to 100 per cent) for these sources, although these cases have higher S/N than most of the WISP sources. A larger sample with follow-up observations from VLT/FORS2 (Very Large Telescope/FOcal Reducer and low dispersion Spectrograph; $R \sim 1200$; 0.51–0.85 μm) will be presented in Boyett et al. (in preparation), which we briefly highlight here. This program observed 85 emission-line galaxies out of 138 potential emission sources in the four targeted WISP fields. Emission lines were detected in the FORS2 data for 38/85 cases, with agreement in the line identification between the grism and FORS2 data seen for 36 of those galaxies (suggesting 95 per cent accuracy of line identification). Of the 38 galaxies with detected lines, 17 are cases where the grism redshift is based on a single line. All cases where lines were not detected with FORS2 are either cases where no lines were predicted to lie in the FORS2 window (15/85) or the grism redshift is based on a single line (32/85). In summary, these findings add further support that selecting sources with multiple emission lines is necessary to ensure robust redshift estimates.

The precision of the redshifts (relating to REDSHIFT_ERR) are a separate metric from accuracy, with the former mainly driven by the spectral resolution of the grism data. For reference, the WFC3/IR detector using the *G141* grism has a sampling of 46.5 Å pixel^{-1} and a FWHM ~ 110 Å. We characterize the precision as $\sigma_z/(1+z)$ (i.e. REDSHIFT_ERR /(1 + REDSHIFT), which has a median value of 0.00088 ($\sim$0.09 per cent) and the 16th and 84th percentiles are 0.00045 and 0.00154, respectively. Bagley et al. (2020) performed an independent test of the precision using the 36 WISP fields that overlap to some degree with each other. These fields result in $\sim$140 sources that were observed multiple times, often with very different exposure times (field depths) and roll angles. This comparison

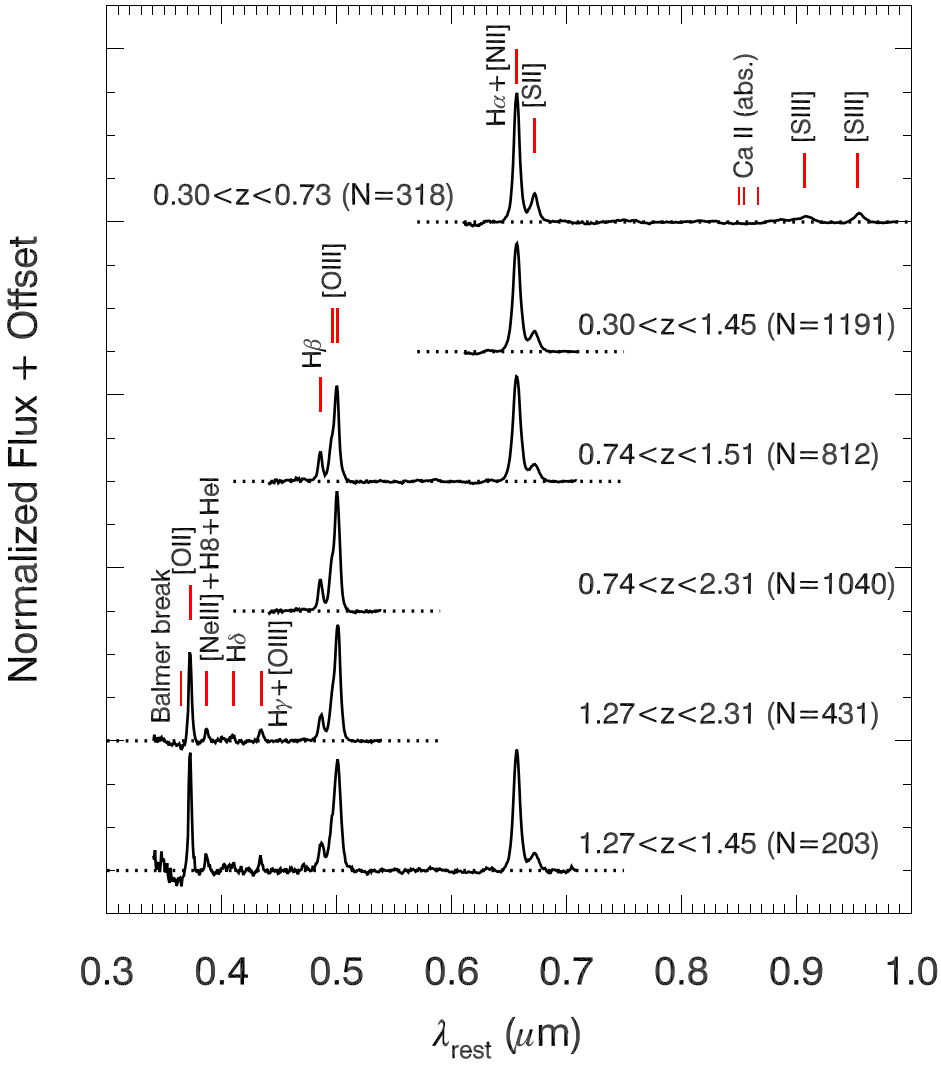


Figure 13. Demonstration of the wavelength coverage of WISP stacked spectra for different redshift windows. Multiple emission lines are simultaneously available for several windows in the redshift range from $z = 0.3$ to 2.3. The sample size indicated corresponds to sources that satisfy the criteria for our science results (criterion 1–8; see Section 7.1).

showed an empirical precision of $\sigma_z/(1 + z) = 0.00136$, which is in close agreement with the median value in the catalogue when accounting for the fact that the test combines the uncertainty of two line measurements.

7 RESULTS

We demonstrate the utility of combining the WISP photometric and emission-line catalogues to study galaxy evolution by dividing the sample into four grism windows (redshift regimes) where different sets of emission lines are available in *G102* and/or *G141* (i.e. do not always require both) to constrain properties of the interstellar medium (ISM):

(i) grism coverage of Hα + [N II], [S II], and [S III] doublet ($0.30 < z < 0.73$),
(ii) grism coverage of Hα + [N II] and [S II] ($0.30 < z < 1.45$),
(iii) grism coverage of Hβ and [O III] ($0.74 < z < 2.31$),
(iv) grism coverage from [O II] to [S II] ($1.27 < z < 1.45$; maximum λ-coverage).

We do not examine the window where Hβ, [O III], and Hα + [N II] are simultaneously covered to constrain the Balmer decrement (Hα/Hβ; $0.74 < z < 1.51$), which is a subset of sources in the third window above, because this was the focus of Battisti et al. (2022). Similarly, we do not examine the window where [O II], Hβ, and [O III] are covered ($1.27 < z < 2.31$; wider redshift than fourth window above) to constrain the metallicity via the $R_{23} \equiv$ ([O II]$\lambda\lambda$3726, 3729 + [O III]$\lambda\lambda$4959, 5007)/Hβ diagnostic because this was the focus of Henry et al. (2021). Both of those studies also supplement their samples with other *HST* grism surveys such that they have larger sample sizes than are available from WISP alone. A visual demonstration of the spectral coverage from WISP for these windows is shown in Fig. 13, together with the sample size satisfying the criteria for our science results (Section 7.1).

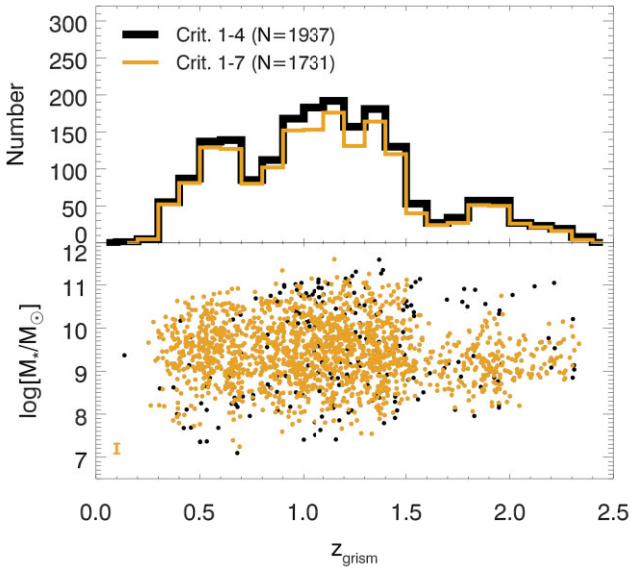


Figure 14. Top panel: grism redshift, $z_{\rm grism}$, distribution for galaxies in the WISP sample that have robust photometric and spectrocopic data (criterion 1–4; black) and the subset that also have good SED fits, accurate stellar masses, and are not AGN candidates (criterion 1–7; orange), as described in Section 7.1. Bottom panel: stellar mass versus redshift. A representative median 1σ error bar for the criterion (1)–(7) sample (orange) is indicated in the lower left, with values of $\langle\sigma(z_{\rm grism})\rangle = 0.0017$ (i.e. negligible) and $\langle\sigma(\log[M_\star/{\rm M}_\odot]) = 0.12$ dex. No significant selection effects on stellar mass are apparent with redshift.

7.1 Sample selection criteria

Combining the overlapping sources in both the WISP photometric and emission-line catalogues provides us with a parent sample of 1937 galaxies that satisfy the following spectroscopic and photometric selection criteria:

(i) One emission line with S/N $\geq$ 3 and one additional line with S/N $\geq$ 2.
(ii) At least three bands of photometry with S/N $>$ 3.
(iii) Independent redshift agreement between the two reviewers (`REDSHIFT_FLAG = 0`).
(iv) Emission-line FWHM $<$ 600Å (`FWHM_OBS`$<$600).

Criteria (1) ensures an accurate $z_{\rm grism}$, which is important for optimally aligning the spectra, as well as reducing false identifications. Criteria (2) ensures that we have adequate SED coverage for characterizing stellar masses, $M_\star$. Criteria (3) removes ambiguous sources where the WISP pipeline reviewers have either assigned different redshift solutions or the spectrum was rejected by one of the reviewers (no confidence in redshift solution). We note that criteria (3) removes $\sim$1/3 of the potential sample, which indicates the difficulty in determining reliable redshifts from low S/N spectra. Criteria (4) removes very broad emission-line sources ($\sim$1 per cent of sample). These broad profiles may be due to an AGN, which we want to exclude for our analysis. The redshift and stellar mass distribution of sources satisfying criterion (1)–(4) are shown in Fig. 14.

For creating stacked spectra according to stellar mass, we also impose additional cuts on SED goodness of fit and uncertainty on stellar mass:

(v) $\chi^2_{\rm red} < 3$,
(vi) $\sigma(\log M_\star) < 0.3$ dex,

which is detailed in Section 7.2. Together, criterion (5) and (6) remove 147 sources (6 fail both criteria) or 7.6 per cent of the sample. For reference, width of the stellar mass bins in our analysis are roughly 0.3 dex or larger to improve the reliability of bin assignment for each galaxy.

For the subset sources with grism coverage and detection of [O III] and Hβ, we also exclude AGN candidates based on the mass–excitation diagram (MEx, log[[O III]/Hβ] versus log[$M_\star$/M$_\odot$]; Juneau et al. 2014):

(vii) non-AGN in MEx diagram

which is detailed in Section 7.4.1. Criteria (7) removes an additional 59 sources. This leave 1731 sources that satisfy criterion (1)–(7), and this sample is also shown in Fig. 14.

Finally, a criteria limiting the contamination of the main emission lines if that line is the focus of the stacking analysis:

(viii) neither reviewer marked line contamination (`[LINE]_CONTAM<4` from both reviewers).

For example, we require Hβ and [O III] are uncontaminated for the stacks in the Hβ and [O III] window (Section 7.4.1). Depending on the window considered, this criteria typically removes an additional 15 per cent–20 per cent of sources in that window. This highlights the high rate of contamination that can occur in single-orientation grism data.

7.2 SED modelling for stellar masses

To estimate stellar masses, we perform SED fitting on our galaxy sample using the MAGPHYS (high-z) spectral modelling code (da Cunha et al. 2015; Battisti et al. 2020), adopting the grism redshift as the input redshift (i.e. fixed-z). Prior to fitting, the photometry is corrected for foreground MW extinction using the Galactic dust extinction maps from Schlafly & Finkbeiner (2011) via the NASA/IPAC Infrared Science Archive.[14] MAGPHYS uses spectral population synthesis models of Bruzual & Charlot (2003), and we refer readers to that paper for details on the stellar tracks, templates, and isochones adopted in that model, also noting that it does include a prescription for thermally pulsing asymptotic giant branch stars. MAGPHYS adopts a Chabrier (2003) initial mass function (IMF) and has 14 free model parameters (high-z version), and for full details we refer readers to the documentation on the MAGPHYS website.[15] In brief, they include: a uniform prior in metallicity from 0.2 to 2 times solar (one parameter); a parametric star formation history (SFH; three parameters), which rises linearly at early ages and then declines exponentially (delayed-tau model) with additional instantaneous bursts of star formation; the dust model of Charlot & Fall (2000, four parameters) for which the interstellar dust is distributed into two components, one associated with star-forming regions (migration time of 10 Myr) and the other with the diffuse ISM, with the addition of the 2175 Å absorption feature (Battisti et al. 2020); the dust emission models of da Cunha, Charlot & Elbaz (2008), which uses templates based on four components (five parameters); and a normalization that sets the stellar mass and star formation rate (SFR) from the SFH (one parameter). The adopted SFH parametrization may introduce systematic biases to the stellar mass estimates (e.g. Leja et al. 2019). For reference, when comparing overlapping WISP galaxies in Henry et al. (2021), who use similar photometry but adopt a non-parametric SFH for their SED modelling, we find their stellar

[14] https://irsa.ipac.caltech.edu/applications/DUST/
[15] http://www.iap.fr/magphys/index.html

masses are systematically larger than the MAGPHYS-derived values by 0.2 dex.

MAGPHYS does not include templates for emission-line fluxes and therefore we perform emission-line subtraction prior to SED fitting, when available. This is especially important for this study because we are using an emission-line-selected sample. If one assumes a roughly flat continuum (in F_λ), the average flux density measured in the photometry can be approximated as (e.g. Whitaker et al. 2014)

$$F_\lambda \simeq F_{\lambda,\mathrm{cont}} + F_{\mathrm{line}}/\Delta\lambda\,, \quad (1)$$

where $F_{\lambda,\,\mathrm{cont}}$ is the continuum-only flux density and $\Delta\lambda$ is the width of the filter, which we take to be its FWHM. We subtract emission-line fluxes for all lines with S/N $>$ 2 from the photometric data. The impact of emission lines on photometry are the largest for galaxies with fainter continuum emission (typically lower $M_\star$) and higher equivalent widths (EW $= F_{\mathrm{line}}/F_{\lambda,\,\mathrm{cont}}$). An example of a MAGPHYS fit for a WISP galaxy is shown in Fig. 15.

For each MAGPHYS fit, a goodness of fit is determined based on the best-fitting model using a reduced χ^2 metric, $\chi^2_{\mathrm{red}} = \chi^2/N_{\mathrm{bands}}$, where N_{bands} is the number of bands observed with non-zero flux. We exclude cases of poor-quality fits by removing galaxies with $\chi^2_{\mathrm{red}} > 3$ from our analysis (criteria 5), which removes 140 of the criterion (1)–(4) sources. Cases of poor fits may be associated with poor/inconsistent photometric data and/or AGN contamination (AGN models are not included in the SED fitting). We also require accurate stellar masses because we will bin our data according to stellar masses, and exclude sources with $\sigma(\log M_\star) > 0.3$ dex (based on the 16th and 84th percentiles of the posterior PDF (probability distribution function); $\sigma(\log M_\star) = (\log M_{\star,\,\mathrm{p84}} - \log M_{\star,\,\mathrm{p16}})/2$; criteria 6). This occurs for 13 of the criterion (1)–(4) sources. The median values of the remaining WISP sample is $\bar{\chi}^2_{\mathrm{red}} = 0.32$. The low χ^2 values are due to the fact that the models tend to 'overfit' the data (more free parameters than data points).

Table 11 provides the $\log M_\star$ and log SFR(SED) percentiles from MAGPHYS for our sample of 1937 galaxies that satisfy criterion (1)–(4) in Section 7.1. SFR(SED) represents the average SFR over the last 100 Myr of the SFH. For reference, the median 1σ uncertainty on $\log M_\star$ and log SFR(SED) (taken as (84th–16th percentile)/2) is 0.12 and 0.25 dex, respectively. We note that our uncertainties may be underestimated due to the SFH parametrization used in MAGPHYS. For overlapping WISP galaxies in Henry et al. (2021), which uses a non-parametric SFH for SED modelling, their median 1σ uncertainties on $\log M_\star$ and log SFR(SED) are larger by 0.03 and 0.04 dex, respectively. SED-derived SFRs have very large uncertainty when relying only on UV through NIR data. For this reason, we avoid using SED-derived SFRs for our analysis and instead use Hα-based estimates whenever possible. We note that the median number of bands available for the subsample is 5, with 85 per cent of the sample having *coverage* in IRAC (i.e. detection or an upper limit). The median uncertainty on $\log M_\star$ for sources with 3, 4, 5, and 6 + bands available is 0.19, 0.13, 0.11, and 0.10 dex, respectively. We find that fields without IRAC coverage have only marginally higher stellar mass uncertainty (0.01 dex) and no significant bias. The difference in median value for the two samples is $\log M_\star$(w/ IRAC) $-$ $\log M_\star$(no IRAC) = 0.045 dex and below the typical uncertainty and our binning size. We stress that these are formal uncertainties and that the true mass uncertainty is higher when accounting for systematic uncertainties arising from model assumptions (e.g. SFH). We include the number bands available in the SED modelling for each source as a column in Table 11.

7.3 Spectral stacking and emission-line fitting of stacked spectra

All galaxy spectra are stacked and fit following similar methods to those described in Dai et al. (2021), Henry et al. (2021), and Battisti et al. (2022). In brief, we use the continuum-subtracted spectra (using a cubic spline; see Section 5.3) and normalize them by the 'typical' brightest line in the spectral window considered, which for the windows we consider is either Hα + [N II] (low-z windows) or [O III] (high-z windows). The spectra are de-redshifted using a linear interpolation to shift them onto a common grid of rest wavelengths and we take the median of the normalized fluxes at each wavelength.

The procedure to fit the stacked spectra differs slightly from the method used for individual sources for the emission-line catalogue (Section 5.3) and this is due to the fact that the stacks reach greater depth such that more parameters are generally required for good fits (e.g. narrow + broad components). To fit the stacked spectra, we fit a set of Gaussian profiles to the emission lines in the region of interest. We adopt two Gaussian components for each line, one narrow and one broad component. Multiple components can arise due to kinematic differences among ionizing sources (e.g. H II versus AGN), but can also occur in grism spectra due to line profiles having a dependence on the spatial distribution of the emitting sources. The FWHM of the broad component is fixed to be the same for all of the lines and also required to be between 1–4$\times$ the FWHM of the narrow components. The amplitudes of the broad components for each line are allowed to vary independently (among positive values).

The emission lines are simultaneously fit with the following assumptions/restrictions: (1) the ratio of [O III]λ5007/[O III]λ4959 is fixed to 2.98:1 (Storey & Zeippen 2000) and [S III]λ9532/[S III]λ9069 is fixed to 2.47:1 (Berg et al. 2021), (2) single profiles are used for the closely spaced blends of [O II]$\lambda\lambda$3727, 3729, Hα+[N II]$\lambda\lambda$6548, 6583, and [S II]$\lambda\lambda$6716, 6731, (3) we require the narrow component FWHM of close pairs to match (e.g. Hβ and [O III], Hα + [N II], and [S II]). We do not match all components in order to account for the effect that the spectral resolution difference between the *G102* and *G141* grisms can have on the profiles, (4) we require the FWHM of the narrow components to be within a factor of two with each other, (5) we allow a $\pm$10 Å shift (rest frame) of the emission-line centroids to accommodate systematic uncertainties in the grism wavelength solution, and (6) we account for any (small) residual continuum offsets due to imperfect continuum subtraction by including free parameters for the spectra amplitudes (i.e. constant offsets) in the regions near emission-line groups (e.g. 4400 Å $< \lambda_{\mathrm{rest}} <$ 5500 Å for Hβ and [O III]; and 6000 Å $< \lambda_{\mathrm{rest}} <$ 7100 Å for Hα + [N II] and [S II]).

Line flux measurements for the stacks are based on scaling the average flux of the normalized line in each bin (i.e. reversing the normalization). Line flux measurement uncertainties on the stacked spectra are obtained by bootstrapping with replacement. In brief, for each sample of N galaxies that are stacked, we draw N random galaxies from that sample, allowing individual objects to be selected more than once. We create a new stack from these objects and measure the lines and repeat this procedure 1000 times and calculate the standard deviation on the line fluxes from these realizations and adopt this as the uncertainty.

7.4 Stellar mass stack results

For all stellar mass bins, we require that they contain $N \gtrsim 100$ galaxies to ensure reliable recovery of faint emission lines and ensure stacks are not sensitive to any individual outlier galaxies in the stack. This also ensures that corrections, which are based on averages, are

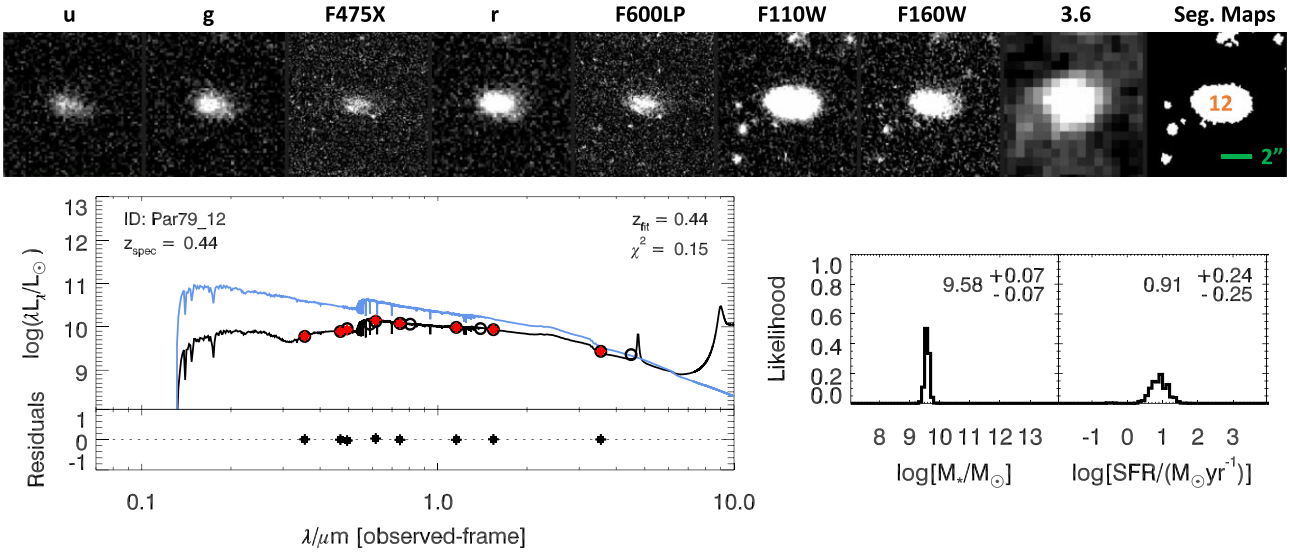


Figure 15. Top row: native resolution 10 arcsec × 10 arcsec image thumbnails (linear scale) of the available photometry for Par79_12 (i.e. parallel field 79, object 12). The last panel shows the SExtractor segmentation map, based on *F110W* and *F160W*, which is an input for TPHOT, with object 12 indicated. Bottom left: the MAGPHYS best-fitting SED (black line) for Par79_12. The red squares are the observed photometry and the black circles are the corresponding model values. The blue line shows the predicted intrinsic stellar population SED (without attenuation). Bottom right: the posterior PDFs for stellar mass and SFR are also shown. The available photometric coverage is sufficient to accurately constrain stellar masses (∼0.1 dex), however SFRs have poorer constraints due to the lack of rest-frame IR data (i.e. age/dust reddening degeneracy; ∼0.3 dex).

Table 11. Stellar mass and SFR percentiles from MAGPHYS for the 1937 WISP sources with 'robust' spectroscopy and photometry (criterion 1–4 in Section 7.1).

Par	Obj	$z_{\rm grism}$	$z_{\rm grism,\,err}$	AGN_flag[a]	log $[M_\star/{\rm M}_\odot]$					log [SFR(SED)/(M$_\odot$yr^{-1})]					$\chi^2_{\rm red}$	$N_{\rm bands}$
					2.5	16	50	84	97.5	2.5	16	50	84	97.5		
1	10	0.5084	0.0013	0	9.92	10.03	10.15	10.26	10.40	0.01	0.73	1.16	1.58	1.90	0.011	5
1	13	0.5309	0.0011	0	9.80	9.94	10.02	10.13	10.21	−0.60	0.26	0.92	1.28	1.74	0.055	7
1	15	0.6699	0.0011	0	9.86	9.96	10.04	10.13	10.22	0.68	0.95	1.32	1.73	2.08	0.464	7
1	28	1.3443	0.0028	0	9.41	9.50	9.61	9.68	9.80	1.07	1.16	1.32	1.48	1.58	0.297	7
1	41	1.3065	0.0018	0	9.67	9.81	9.89	10.00	10.13	0.91	1.21	1.53	1.77	1.94	0.326	7
								...								

Notes. A full ASCII version of this table is available online. Percentiles provided are 2.5 per cent, 16 per cent, 50 per cent, 84 per cent, and 97.5 per cent. SFR(SED) corresponds to the average SFR over the last 100 Myr of the SFH. $\chi^2_{\rm red} = \chi^2/N_{\rm bands}$ is the reduced χ^2 of the best-fitting model and $N_{\rm bands}$ is the number of bands observed with non-zero flux.
[a]AGN_flag = 1 indicates sources that lie in the AGN region of the MEx diagram (Section 7.4.1).

reasonable for the sample (e.g. Section 7.4.2). Below we show our results using stellar mass stacks for four redshift windows.

7.4.1 Hβ and [O III] window (0.74 < z < 2.31)

We start with galaxies with both Hβ and [O III] grism coverage ($0.74 < z_{\rm grism} < 2.31$) because it is one of the largest groups we consider ($N = 1040$ with criterion 1–8) and the ratio [O III]/Hβ as a function of stellar mass, known as a the MEx diagram, has been established as a reliable tool for distinguishing star forming galaxies (SFGs) and AGN at both low (Juneau et al. 2014) and high redshifts (Coil et al. 2015; Kashino et al. 2019). The results of Kashino et al. (2019) and Coil et al. (2015) suggest that the demarcation line to distinguish SFGs and AGN should shift to higher stellar masses with increasing redshift, requiring a 0.5 and 0.75 dex shift at $z \sim 1.6$ and ~ 2.3, respectively. We use the demarcation from Kashino et al. (2019) as our reference because we are examining similar redshifts ($z \sim 1.0$ and ~ 1.6). The position on the MEx diagram of individual galaxies in our subsample is shown in Fig. 16 (bottom). We find that 10.4 per cent of the WISP sources in this subsample are detected in both Hβ and [O III] (S/N > 3), 61.1 per cent detected only in [O III], and the rest are undetected in both lines (S/N < 3; 28.5 per cent). For Hβ undetected cases, we use the 3σ Hβ error and treat the ratio as a lower limit. We find 59 individual sources above the AGN lower boundary line, which are excluded from the stacks.

We subdivide galaxies in this window into two redshift ranges, $0.74 < z < 1.22$ and $1.22 \leq z < 2.31$ ($z = 1.22$ is the median for this window), and each of those into five equal-number bins in stellar mass (10 bins total). The spectra of these stellar mass stacks and their position on the MEx diagram are shown in Fig. 16. The average galaxy properties and emission-line values for these stacks are listed in Table 12. As expected, most of the stacks lie in the SFG region of the diagram. However, despite excluding all individual galaxies with emission-line ratios or upper limits that lie in the AGN region of the MEx diagram prior to stacking (purple squares), the position of the highest mass stacks for each redshift remains on/above the lower AGN boundary. We attribute this result to the 28.5 per cent of individual sources that are unconstrained in the MEx diagram (i.e. those with S/N < 3 for both [O III] and Hβ), together with the fact

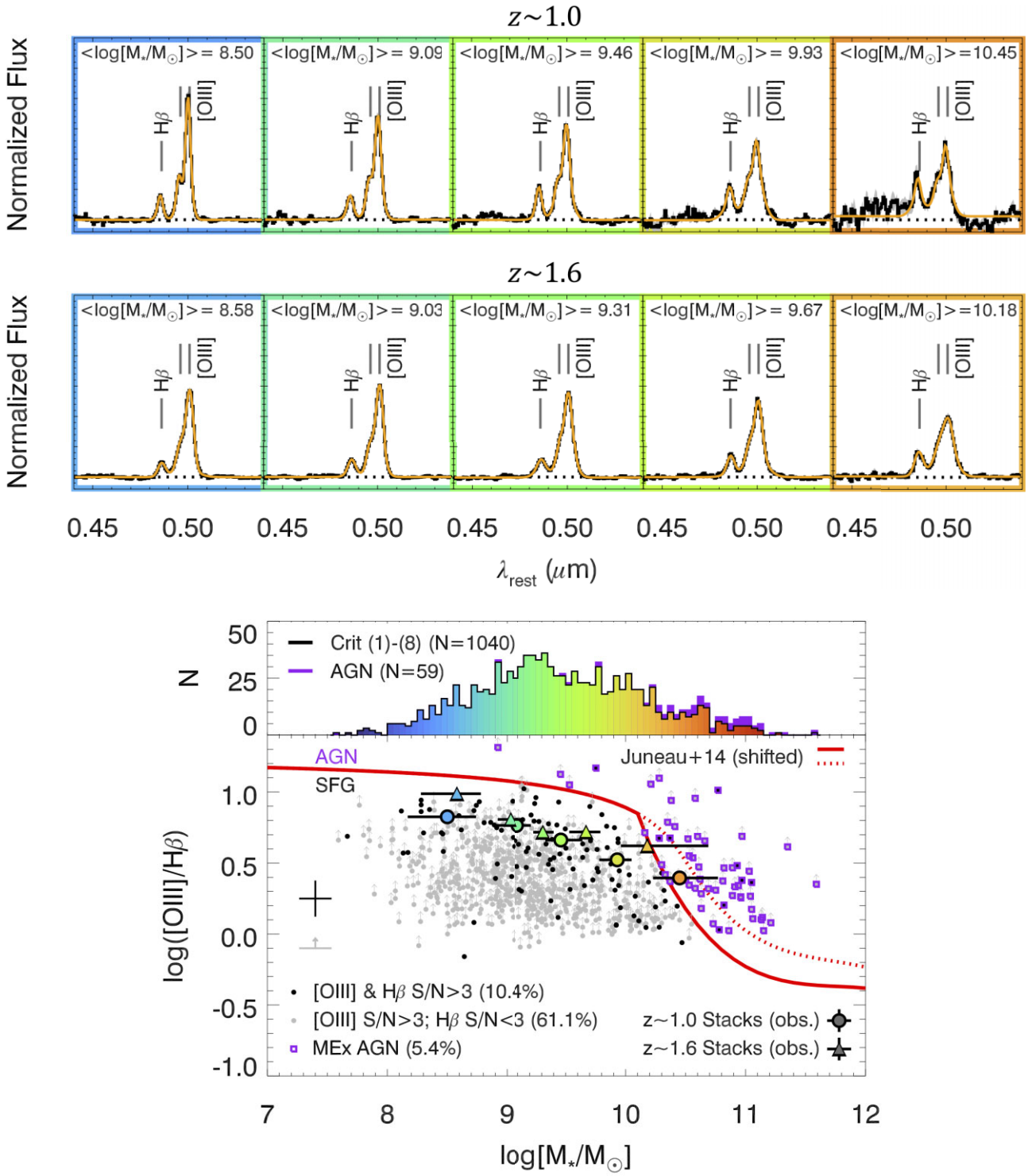


Figure 16. Top: stacked spectra in the Hβ and [O III] window normalized to [O III] total flux in bins of stellar mass for our subsample at $0.74 < z < 2.31$ ($N = 1040$ with criteria 1–8). The median $\log M_\star$ is indicated in each panel. The spectral fits are shown by the orange lines. Bottom: MEx diagram for our sample. The solid (dashed) red lines corresponds to the lower (upper) boundaries of Juneau et al. (2014) shifted to higher masses by 0.5 dex (i.e. to the right), corresponding to the relation found at $z \sim 1.6$ by Kashino et al. (2019), and up by 0.13 dex to account for using total [O III] instead of only [O III]λ5007. Small black and grey points denote individual galaxies where MEx positions are constrained (71.5 per cent of sample), with median error bars given on the middle-left, and AGN candidates are indicated by purple squares. Hβ undetected cases are treated as lower limits (see Section 7.4.1). The large circles/triangles (colour based on median $\log M_\star$), correspond to the *observed* values from the stacked spectra shown at the top (i.e. Hβ is not corrected for stellar absorption). Stack x-axis error bars denote the 1σ mass range spanned by each bin (Table 12 lists full mass range). The y-axis error bars denote the line ratio error on the stacked spectra, which are smaller than the symbol sizes in this case (<0.05 dex; see Table 12). Despite excluding AGN candidates in the stacking, the position of our highest mass stacks remain on/above the lower AGN boundary. We attribute this to the 28.5 per cent of our individual sources being unconstrained (S/N < 3 in both [O III] and Hβ), with larger $\log M_\star$ galaxies preferentially lacking detection in these lines (see Section 7.4.1).

Table 12. Average properties, emission-line luminosities, and line ratios for $N = 1040$ stacked spectra in the Hβ and [O III] window ($0.74 < z < 2.31$) that satisfy criterion (1)−(8).

N	$\log M_\star$ range	$\langle \log M_\star \rangle$	Hβ	[O III]	$\frac{\text{[O III]}}{\text{H}\beta}$
			(10^{41} erg s^{-1})		
		$0.74 < z_{\rm grism} < 1.22$ ($z \sim 1.0$)			
104	[7.59, 8.82]	8.50	1.09 ± 0.06	7.27 ± 0.11	6.69 ± 0.37
104	[8.83, 9.27]	9.09	1.24 ± 0.06	7.24 ± 0.11	5.82 ± 0.31
104	[9.29, 9.71]	9.46	1.09 ± 0.06	5.01 ± 0.09	4.60 ± 0.28
104	[9.72,10.12]	9.93	1.42 ± 0.09	4.74 ± 0.11	3.34 ± 0.23
104	[10.12,11.60]	10.45	1.17 ± 0.11	2.92 ± 0.15	2.49 ± 0.27
		$1.22 \leq z_{\rm grism} < 2.31$ ($z \sim 1.6$)			
104	[7.77, 8.85]	8.58	2.11 ± 0.10	20.56 ± 0.25	9.72 ± 0.49
104	[8.86, 9.19]	9.03	3.60 ± 0.15	23.11 ± 0.18	6.42 ± 0.27
104	[9.19, 9.47]	9.31	4.70 ± 0.16	24.50 ± 0.18	5.22 ± 0.19
104	[9.47, 9.89]	9.67	3.82 ± 0.25	20.00 ± 0.31	5.24 ± 0.35
104	[9.89,11.26]	10.18	2.60 ± 0.29	10.89 ± 0.30	4.18 ± 0.48

Note. Hβ is not corrected for stellar absorption. [O III] is the sum of [O III]$\lambda\lambda$4959, 5007.

that higher stellar mass galaxies have preferentially weaker [O III] and Hβ relative to Hα than lower stellar mass galaxies (e.g. fig. 5 of Battisti et al. 2022). As a reference, the lowest stellar mass bins for each redshift range have ∼90 per cent of individual sources detected in [O III] or Hβ, whereas the highest stellar mass bins have 30 per cent–40 per cent of individual sources detected in [O III] or Hβ (for these cases, usually Hα and [S II] are detected). Thus, a majority of sources in the higher mass bins might be AGN that we are unable to identify individually using the MEx diagram. This indicates that there may be significant AGN contamination for WISP emission-line sources at high stellar masses.

These WISP results are consistent with Förster Schreiber et al. (2019), who examined AGN occurrence rates, $f_{\rm AGN}$, in a sample of 600 galaxies at $0.6 < z < 2.7$ in KMOS$^{\rm 3D}$ (K-band Multi Object Spectrograph). They found that galaxies with $\log(M_\star/{\rm M}_\odot) < 10.2$ have an AGN occurrence rate of $f_{\rm AGN} \lesssim 10$ per cent, with $f_{\rm AGN}$ increasing dramatically with increasing mass (e.g. ∼60 per cent at $\log(M_\star/{\rm M}_\odot) = 11$; see their fig. 6). These results are also similar to findings in Henry et al. (2021), which included both the WISP and CLEAR + 3D-HST surveys. For the subsequent analysis, we do not exclude galaxies based on a stellar mass threshold (e.g. Battisti et al. 2022), but caution that AGN may contaminate bins above $\log(M_\star/{\rm M}_\odot) \gtrsim 10.2$.

7.4.2 *Hα + [N II] and [S II] window ($0.30 < z < 1.45$)*

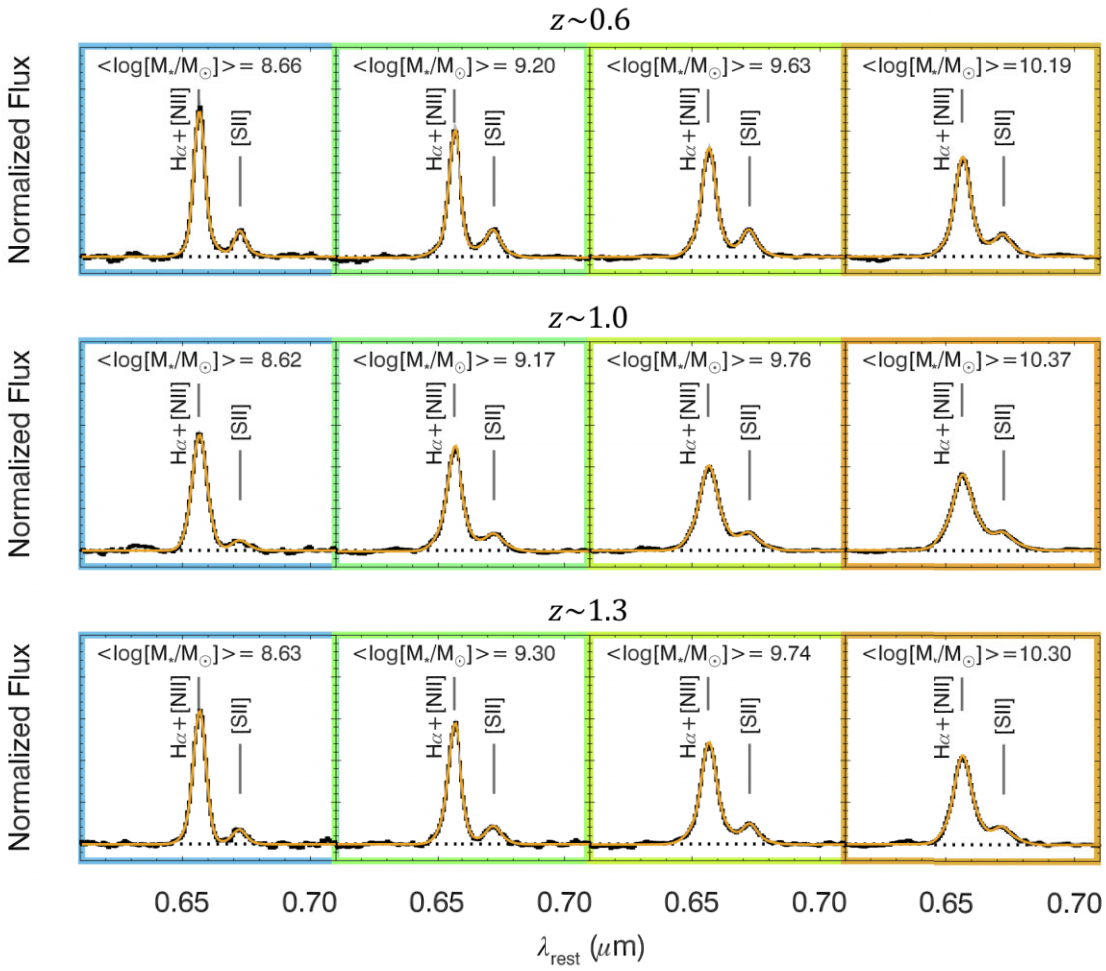
Next, we consider galaxies with grism coverage of Hα + [N II] and [S II] ($0.30 < z_{\rm grism} < 1.45$; $N = 1191$ with criterion 1–8), which is our largest group size. This group can be used to examine how representative our sample is of typical SFGs at these redshifts. To do this, we characterize WISP galaxies relative to the star-forming galaxy MS (log SFR versus log$M_\star$; e.g. Brinchmann et al. 2004; Speagle et al. 2014; Leslie et al. 2020), using SFRs based on Hα (described below).

We divide galaxies in this window into three redshift ranges, $0.30 < z \leq 0.83$, $0.83 < z \leq 1.15$, and $1.15 \leq z < 1.45$ (equal number in each), and each of those into four equal-number bins in stellar mass (12 bins total). The spectra of these stellar mass stacks are shown in Fig. 17. The average galaxy properties and emission-line values for these stacks are listed in Table 13.

In order to estimate SFRs from Hα, we need to apply several corrections. We note that we only apply correction factors on stacked spectra based on groups of $N > 100$ galaxies to minimize the impact of intrinsic variation of the correction factors. First, we deblend the Hα + [N II] line using the stellar mass- and redshift-dependent functional relation from Faisst et al. (2018), which used ∼190 000 SDSS galaxies combined with the observed BPT locus evolution of SFGs from $0 < z < 2.5$. These corrections were derived over the range of $0 < z < 2.7$ and $8.5 < \log(M_\star/{\rm M}_\odot) < 11.0$ and have an intrinsic scatter of ∼0.2 dex. A small fraction of our stellar masses are $\log(M_\star/{\rm M}_\odot) < 8.5$ but extrapolating the Faisst et al. (2018) relation to lower masses should have minimal impact as the expected [N II] contribution is $\lesssim 5$ per cent at all redshifts in this mass regime. Based on this relation, the contribution of [N II] range from ∼6 per cent for the lowest mass bins, up to ∼35 per cent for the most massive bins.

Second, we correct Hα for stellar absorption, which cannot be directly fit in the low-resolution grism spectra. To correct for stellar absorption, we use the stellar mass- and SFR-dependent functional relation from Kashino & Inoue (2019), which is based on trends observed for ∼190 000 SDSS galaxies. The fractional corrections were determined over the range of $7.2 < \log(M_\star/{\rm M}_\odot) < 11.4$, with an intrinsic scatter of ∼10 per cent–20 per cent. For each stack, we use the median value of stellar masses and SFRs from MAGPHYS. However, if the inferred SFR from MAGPHYS is lower than SFR(H$\alpha_{\rm obs}$) (corrected for [N II] blending), then we adopt the latter as the input to determine the fraction of stellar absorption because SFR(H$\alpha_{\rm obs}$) can be considered a lower limit. Due to the relatively high specific-SFR (SFR/$M_\star$) of WISP emission-line galaxies, the Hα absorption correction factor is negligible for most of the sample (∼ 1 per cent).

Lastly, we correct for dust extinction. Due to the fact that Hβ is not available for most galaxies in this sample (due to coverage and/or depth), we cannot use the Balmer decrement for dust corrections. Instead, we use the τ_B^l–log $M_\star$ relation of SDSS galaxies ($z \sim 0$; Battisti et al. 2022), where τ_B^l is the Balmer optical depth, and assume a 0.15 dex uncertainty on τ_B^l for a given log $M_\star$. This relation appears to show minimal evolution with redshift (e.g. Battisti et al. 2022; Shapley et al. 2022). SFRs based on these corrections were found to be roughly consistent with SED-based SFRs in Battisti et al. (2022) for a similar sample. We adopt the MW extinction curve from Fitzpatrick et al. (2019), but note that most extinction curves have shapes that are very similar in the optical/NIR regime such that this choice has a smaller impact relative to the uncertainty on τ_B^l.



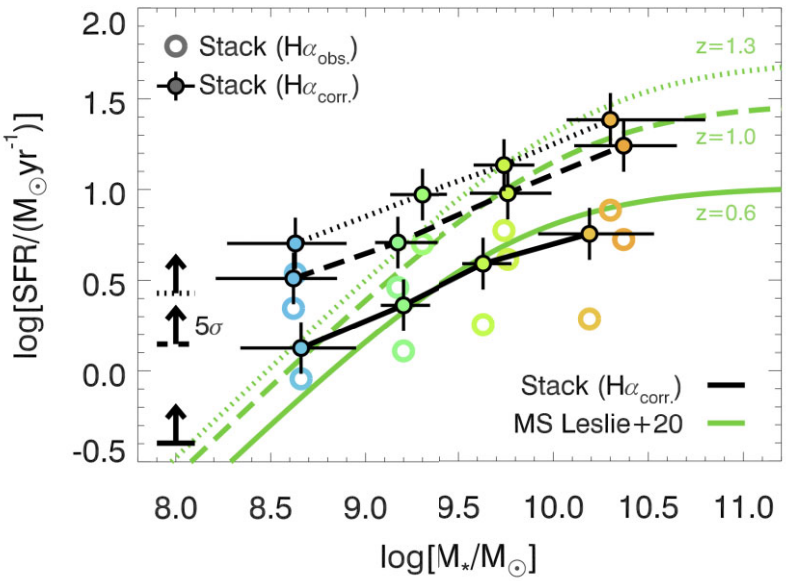


Figure 17. Top: similar to Fig. 16, but now for the Hα + [N II] and [S II] window and the stacked spectra are normalized to Hα + [N II] total flux. Bottom: the galaxy MS (log SFR versus log$M_\star$), for our sample binned by stellar mass (large coloured symbols) across three redshift ranges, which are distinguished by the line style connecting the filled data points (not a fit; $z \sim 0.6$, 1.0, and 1.3 are solid, dashed, and dotted, respectively). Bins are based on stacked Hα without dust corrections (SFR(H$\alpha_{\rm obs}$)) and with dust corrections (SFR(H$\alpha_{\rm corr}$)), where the latter is based on the log$M_\star$–τ_B^l relation (Battisti et al. 2022). The green lines are the galaxy MS from Leslie et al. (2020) at the median redshift of the groups (line styles match as above). Our sample mostly coincides with the MS except for a bias at low masses due to sensitivity limits. The median Hα line detection threshold (5σ) for each redshift range is indicated (bottom left), but we note this has large variation due to differing opportunity lengths.

Table 13. Average properties, emission-line luminosities, and line ratios for $N = 1191$ stacked spectra in the Hα + [N II] and [S II] window ($0.30 < z < 1.45$) that satisfy criterion (1)−(8).

N	log $M_\star$ range	$\langle \log M_\star \rangle$	$\langle \log \mathrm{SFR} \rangle$ H$\alpha_{\rm obs}$	$\langle \log \mathrm{SFR} \rangle$ H$\alpha_{\rm corr}$	Hα + [N II]	[S II]	H$\alpha_{\rm obs}$ (10^{41} erg s^{-1})	H$\alpha_{\rm corr}$
			$0.30 < z_{\rm grism} \leq 0.83$ ($z \sim 0.6$)					
100	[7.25, 9.01]	8.66	−0.04 ± 0.01	0.13 ± 0.14	1.81 ± 0.02	0.35 ± 0.02	1.68 ± 0.04	2.49 ± 0.82
99	[9.01, 9.44]	9.20	0.11 ± 0.01	0.36 ± 0.14	2.73 ± 0.03	0.79 ± 0.02	2.40 ± 0.06	4.29 ± 1.41
99	[9.45, 9.84]	9.63	0.25 ± 0.02	0.59 ± 0.14	4.23 ± 0.05	1.13 ± 0.03	3.34 ± 0.13	7.25 ± 2.39
99	[9.84,10.92]	10.19	0.29 ± 0.02	0.75 ± 0.14	5.28 ± 0.06	1.28 ± 0.04	3.59 ± 0.16	10.58 ± 3.49
			$0.83 < z_{\rm grism} \leq 1.15$ ($z \sim 1.0$)					
100	[7.59, 8.93]	8.62	0.35 ± 0.01	0.51 ± 0.14	4.41 ± 0.04	0.43 ± 0.03	4.13 ± 0.08	6.03 ± 1.98
99	[8.94, 9.44]	9.17	0.46 ± 0.01	0.71 ± 0.14	6.01 ± 0.06	1.16 ± 0.05	5.38 ± 0.11	9.51 ± 3.12
99	[9.45,10.04]	9.76	0.61 ± 0.03	0.98 ± 0.14	9.60 ± 0.08	2.08 ± 0.06	7.65 ± 0.48	17.73 ± 5.90
99	[10.04,11.34]	10.37	0.72 ± 0.02	1.24 ± 0.14	15.04 ± 0.14	3.31 ± 0.09	9.84 ± 0.47	32.40 ± 10.71
			$1.15 \leq z_{\rm grism} < 1.45$ ($z \sim 1.3$)					
100	[7.54, 9.02]	8.63	0.54 ± 0.01	0.70 ± 0.14	6.83 ± 0.07	0.84 ± 0.07	6.42 ± 0.12	9.41 ± 3.08
99	[9.04, 9.50]	9.30	0.70 ± 0.01	0.97 ± 0.14	10.49 ± 0.10	1.80 ± 0.09	9.36 ± 0.17	17.47 ± 5.72
99	[9.50, 9.96]	9.74	0.77 ± 0.02	1.13 ± 0.14	13.35 ± 0.14	2.65 ± 0.11	11.04 ± 0.39	25.33 ± 8.33
99	[9.97,11.60]	10.30	0.89 ± 0.04	1.38 ± 0.15	21.22 ± 0.17	4.29 ± 0.12	14.29 ± 1.32	45.02 ± 15.30

Notes. [S II] is the sum of [S II]λλ6716, 6731. H$\alpha_{\rm obs}$ are values after correcting for [N II] blending, where the uncertainty is the 1σ dispersion in correction values for individual galaxies and the line measurement uncertainties added in quadrature. H$\alpha_{\rm corr}$ are values after also correcting for stellar absorption and dust extinction, with the latter introducing significant uncertainty (see Section 7.4.2).

A comparison of our sample, both before and after dust corrections, to the galaxy MS at these redshifts from Leslie et al. (2020) is shown in Fig. 17. The Leslie et al. (2020) relations are based on stacked radio data from ∼200 000 galaxies in the COSMOS (COSMOlogical evolution Survey) field (we use their 'All' sample). For our SFRs, we adopt the conversion from Kennicutt & Evans (2012):

$$\log\left[\frac{\mathrm{SFR(H}\alpha)}{\mathrm{M}_\odot\ \mathrm{yr}^{-1}}\right] = \log\left[\frac{L(\mathrm{H}\alpha)}{\mathrm{erg\ s}^{-1}}\right] - 41.27 \quad (2)$$

which assumes the IMF of Kroupa & Weidner (2003, this is comparable to IMF used in MAGPHYS). All of the dust-corrected SFR values with $\log(M_\star/\mathrm{M}_\odot) \gtrsim 9$ appear roughly consistent with the MS, indicating they are fully representative of 'normal' star-forming galaxies at these redshifts. Even though our WISP galaxies are selected by emission-line strength, this provides a sample similar to those selected by traditional broad-band continuum photometry. This reflects the fact that all star-forming galaxies at $z \gtrsim 0.5$ have strong emission lines which WISP can detect.

At the lowest stellar masses ($\log(M_\star/\mathrm{M}_\odot) \lesssim 9$), however, WISP galaxies tend to reside above the MS, which we attribute to the line detection threshold of WISP ($\sim 5 \times 10^{-17}$ erg s^{-1} cm^{-2}; Atek et al. 2010). In other words, WISP can only detect emission lines with high EWs at low stellar masses. We show the lower limit on SFR(H$\alpha_{\rm obs}$) for this line sensitivity at $z = 0.6$, 1.0, and 1.3 (median redshifts of our bins) in the lower left of Fig. 17.

7.4.3 Hα + [N II], [S II], and [S III] doublet window (0.30 < z < 0.73)

Next, we consider galaxies with grism coverage of Hα + [N II], [S II], and [S III] doublet ($0.30 < z_{\rm grism} < 0.73$; $N = 318$ with criterion 1–8). The $S_{32} \equiv$ [S III]λλ9069,9531/[S II]λλ6716,6731 ratio is a proxy for the ionization parameter of a galaxy and is relatively insensitive to the gas-phase metallicity and ISM pressure (Kewley, Nicholls & Sutherland 2019). However, there is considerable uncertainty regarding the exact conversion of S_{32} to ionization parameter because [S II] lines are typically underestimated by photoionization models (Kewley et al. 2019). For this reason, we simply report the S_{32} line ratios instead of ionization parameter. For reference, 16.0 per cent of the WISP sources in our subsample are detected in both [S II] and [S III] (S/N > 3), 26.5 per cent detected only in [S II], and the rest are undetected in both lines (S/N < 3; 57.5 per cent).

We divide galaxies in this window into three equal-number bins in stellar mass over the full redshift range and the spectra of these stacks are shown in Fig. 18 (top). We label the region where the Ca II triplet absorption features occur, which are not accounted for in the continuum fitting (simple polynomial spline), and it appears that this may affect the stacked spectrum in the region near the [S III]λ9069 line, particularly for the largest stellar mass bin. We enforce a fixed [S III]λ9532/[S III]λ9069 ratio such that the impact should be mitigated, however the WISP values should be treated with caution. Attempting to account for this is beyond the scope of our results. The average galaxy properties and emission-line values for these stacks are listed in Table 14.

We show a comparison of the S_{32} ratio versus $\log M_\star$ for WISP relative to the results for the MaNGA (Mapping Nearby Galaxies at Apache Point Observatory; $z \sim 0$) and MOSDEF (Multi-Object Spectrometer for infra-red exploration Deep Evolution Field; $z \sim 1.5$) samples (Sanders et al. 2020) in Fig. 18 (bottom). We show both the observed and extinction corrected line ratios. For MaNGA and MOSDEF, these extinction corrections are based on direct measurements of Balmer decrements (Hα/Hβ) and assuming an MW extinction curve (Cardelli, Clayton & Mathis 1989). For WISP, we do not have Hβ in our redshift window and, similar to the previous section, we adopt the τ_B–$\log M_\star$ relation from Battisti et al. (2022, with 0.15 dex uncertainty) and the MW extinction curve from Fitzpatrick et al. (2019). We again note that different extinction curves give a similar outcome because of the similarity in their shape in the NIR (due to using a line ratio). The WISP sample shows similar values to the MaNGA sample relation prior to making corrections for diffuse-ionized gas (DIG). If the WISP galaxies ($z \sim 0.55$) have a similar component of DIG as MaNGA ($z \sim 0$), then this would imply they have comparable ionization parameter values as the MaNGA and MOSDEF samples. We also find a decreasing value S_{32} with increasing $\log M_\star$, similar to findings of Sanders et al. (2020) at both lower and higher redshifts, indicating

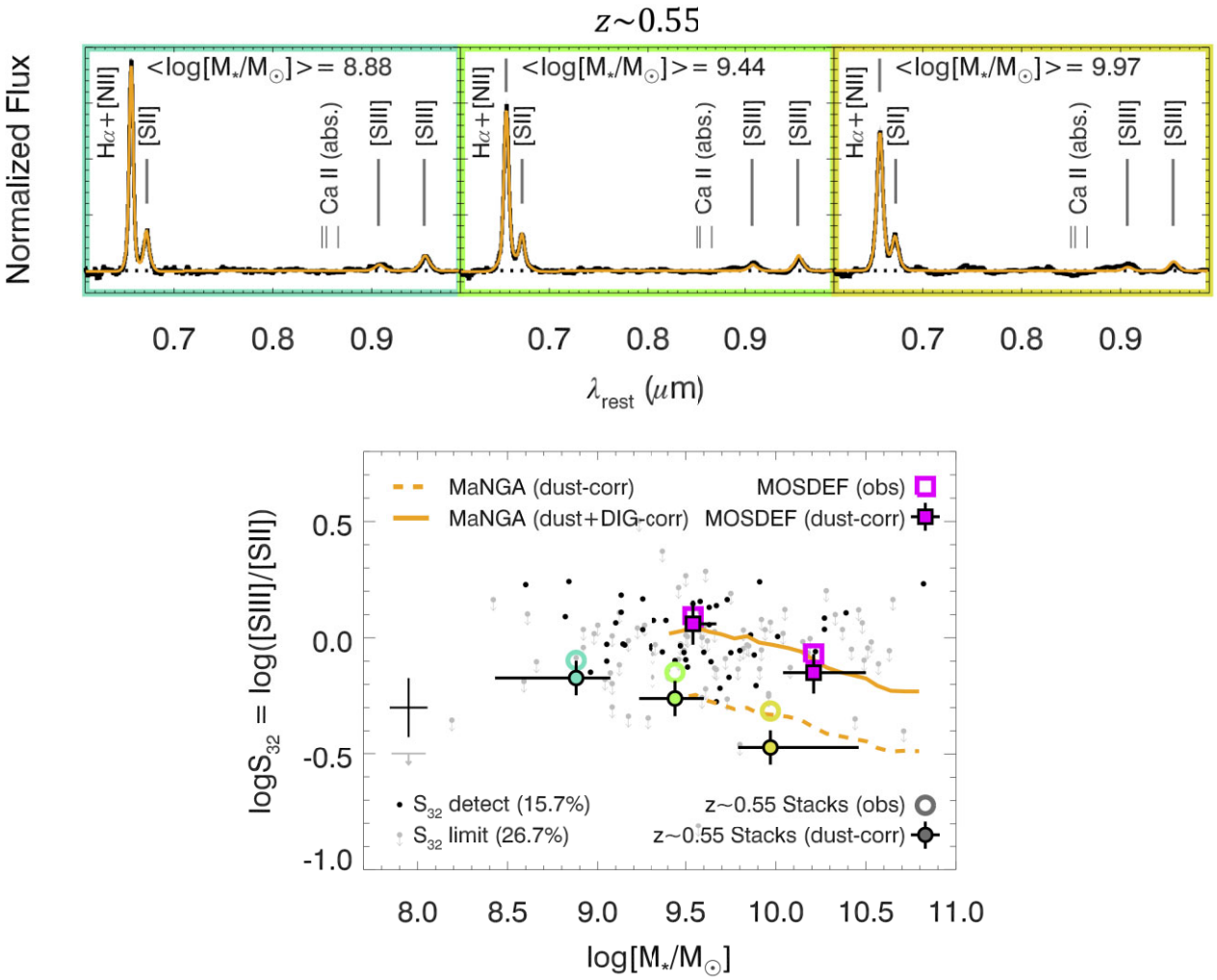


Figure 18. Top: similar to Fig. 16, but now for the Hα + [N II], [S II], and [S III] window ($0.30 < z < 0.73$) and the stacked spectra are normalized to Hα + [N II] total flux. The Ca II triplet absorption features are also indicated and appear to affect the stacked spectrum in the region near the [S III]λ9069 line, particularly for the largest stellar mass bin. Bottom: we compare the S_{32} ratio versus stellar mass from WISP ($z \sim 0.55$) to the results in Sanders et al. (2020) for the MaNGA ($z \sim 0$) and MOSDEF ($z \sim 1.5$) samples. Small black and grey points denote individual galaxies where S_{32} positions are constrained (42.4 per cent of sample), with median error bars given on the middle left. Coloured open and closed symbols correspond to the stellar mass stacks before and after dust extinction correction, labelled 'obs' and 'dust-corr', respectively (see Section 7.4.3). The WISP stacks appear most similar to the MaNGA relation before correcting for DIG, perhaps indicating a non-negligible DIG contribution in the WISP sample. All samples show a mildly decreasing S_{32} with increasing $\log M_\star$.

Table 14. Average properties, line luminosities, and line ratios for $N = 318$ stacked spectra in the Hα + [N II], [S II], and [S III] doublet window ($0.30 < z < 0.73$) that satisfy criterion (1)−(8).

N	$\log M_\star$ range	$\langle \log M_\star \rangle$	Hα + [N II]	[S II]	[S III]	$H\alpha_{obs}$ (10^{41} erg s^{-1})	$H\alpha_{corr}$	$[S\,II]_{corr}$	$[S\,III]_{corr}$	$\frac{[S\,III]}{[S\,II]}$	$\frac{[S\,III]_{corr}}{[S\,II]_{corr}}$
106	[7.25, 9.13]	8.88	1.72 ± 0.02	0.44 ± 0.01	0.35 ± 0.01	1.57 ± 0.04	2.49 ± 0.82	0.69 ± 0.22	0.47 ± 0.09	0.80 ± 0.04	0.67 ± 0.11
106	[9.13, 9.66]	9.44	3.19 ± 0.05	0.84 ± 0.04	0.60 ± 0.02	2.67 ± 0.11	5.29 ± 1.75	1.63 ± 0.52	0.89 ± 0.17	0.71 ± 0.04	0.55 ± 0.10
106	[9.66,10.83]	9.97	4.31 ± 0.05	1.15 ± 0.03	0.55 ± 0.02	3.06 ± 0.17	7.93 ± 2.63	2.87 ± 0.91	0.97 ± 0.19	0.48 ± 0.02	0.34 ± 0.06

Notes. [S II] is the sum of [S II]λλ6716, 6731. [S III] is the sum of [S III]λλ9069, 9532. $H\alpha_{obs}$ and $H\alpha_{corr}$ have same meaning as in Table 13. $[S\,II]_{corr}$ and $[S\,III]_{corr}$ are values after correcting for dust extinction (see Section 7.4.2).

a lower ionization parameter with increasing stellar mass. We highlight that S_{32} is relatively insensitive to the shape of the ionizing spectrum (e.g. Sanders et al. 2020), and that studies favour a redshift evolution in the hardness of the ionizing spectrum (e.g. Steidel et al. 2016).

7.4.4 [O II] to [S II] window (1.27 < z < 1.45)

Finally, we consider galaxies with grism coverage from [O II] all the way to [S II], which occurs for only a narrow redshift window ($1.27 < z_{grism} < 1.45$; $N = 203$ with criterion 1–8). This window is unique in that it provides coverage across most

strong optical emission lines simultaneously, which is beneficial for characterizing various properties of the ISM, including: $O_{32} \equiv$ [O III]$\lambda\lambda$4959,5007/[O II]$\lambda\lambda$3726,3729, a proxy for the ionization parameter (Kewley et al. 2019); $R_{23} \equiv$ ([O II]$\lambda\lambda$3726, 3729 + [O III]$\lambda\lambda$4959, 5007)/Hβ, a proxy for gas-phase metallicity (e.g. Curti et al. 2017); and Hα/Hβ, a proxy for dust attenuation (Calzetti 2001). We note that unlike S_{32}, the relation between O_{32} and ionization parameter is more sensitive to gas-phase metallicity and ISM pressure (Kewley et al. 2019).

We divide galaxies in this window into two equal-number stellar mass bins. The spectra of these stellar mass stacks are shown in Fig. 19. The average galaxy properties and emission-line values for these stacks are listed in Table 15. Unlike previous stacks, we can perform dust corrections based on the measured Balmer decrement. This is based on the methods detailed in Battisti et al. (2022), which require corrections for Hα + [N II] blending and Balmer absorption based on empirical relations from Faisst et al. (2018) and Kashino & Inoue (2019), respectively. We note that the average dust attenuation based on SED modelling is systematically lower than that inferred from the Balmer decrement, with log SFR(SED) being ∼0.2 dex lower than log SFR(H$\alpha_{\rm corr}$). This is consistent with the findings in Battisti et al. (2022) and is expected. In the absence of IR data, the age–dust degeneracy (older populations can produce redder colours) will result in older (lower SFR) templates in MAGPHYS being able to reproduce the data.

We show a comparison of the O_{32} and R_{23} line ratios versus log $M_\star$ for WISP relative to results from SDSS ($z \sim 0$) and the CLEAR survey ($1.1 < z < 2.3$; Papovich et al. 2022), which was an *HST* grism program in GOODS-N and GOODS-S (Great Observatories Origins Deep Survey-North and South; Simons et al. 2023). We show both the observed and extinction corrected line ratios for WISP (only extinction corrected for CLEAR). For CLEAR, the extinction corrections are based on the attenuation measurements from the SED fitting that assume the Calzetti et al. (2000) attenuation curve and assuming the nebular and continuum reddening are the same (see Papovich et al. 2022, for details). For WISP, we use Balmer decrements and the MW extinction curve from Fitzpatrick et al. (2019). The WISP sample has similar values and behaviour to the CLEAR stacks for both line ratios, which are offset above local galaxies based on SDSS. The SDSS contours are based on ∼15 6000 galaxies at $z < 0.2$ (for selection criteria, see Battisti et al. 2022), noting that we use the stellar mass enclosed in the SDSS fibre and not the 'total' stellar mass. For O_{32}, the decreasing values in line ratios with increasing log $M_\star$ suggests a lower ionization parameter and/or softer ionizing spectrum with increasing stellar mass. For R_{23}, the decreasing values in line ratios with increasing log $M_\star$ reflects an increase in metallicity with increasing stellar mass. However, we note that our WISP R_{23} values are close to the log $R_{23} \sim 1$ turnover in the R_{23}–metallicity relation (e.g. Curti et al. 2017), which introduces an ambiguity in relating R_{23} to a metallicity. The CLEAR sample is representative of galaxies on the star-forming galaxy MS (see fig. 2 of Papovich et al. 2022) and our agreement with their trends further supports the argument that the WISP sample also reflects MS galaxies. For a more detailed analysis of the metallicity of WISP galaxies, we refer readers to Henry et al. (2021).

8 CONCLUSION

Slitless spectroscopic surveys are an efficient method to perform large spectroscopic surveys of galaxies across a wide range of cosmic time. We present the public data release of reduced ancillary photometric images obtained for the WISP survey, a large pure-parallel *HST* program, as well as a consistent photometric catalogue containing ∼230 000 sources. This catalogue is based on the SExtractor and TPHOT codes, the latter of which uses the high spatial resolution *HST* data to perform forced photometry on the low spatial resolution ground-based and *Spitzer* data. We also present the WISP emission-line catalogue containing ∼8000 sources. This catalogue is based on a novel combination of an automated line detection algorithm and visual inspection. These data can be used to study a broad range of topics in galaxy evolution over ∼60 per cent of cosmic time ($0.3 \lesssim z \lesssim 3$) and will serve as a useful reference sample for future slitless surveys with *JWST*, *Euclid*, and *Roman*.

We combine the WISP photometric and spectroscopic catalogues to examine the properties of WISP galaxies using stacked spectra in bins of stellar mass over four grism windows (redshift regimes) where specific emission-line ratios are available to study their ISM properties (Section 7.4):

(i) For Hβ and [O III] coverage ($0.74 < z_{\rm grism} < 2.31$), we bin 1040 galaxies and examine their position on the MEx diagram ([O III]/Hβ versus $M_\star$; Juneau et al. 2014). We find that our stacks at log ($M_\star$/M$_\odot$) < 10.2 lie in the star-forming region of the diagram and log ($M_\star$/M$_\odot$) $\gtrsim 10.2$ lie on/above the AGN region of the diagram. This suggests that there may be a non-negligible fraction of sources with AGN at higher masses and is qualitatively consistent with findings from other high-z spectroscopic studies (e.g. Förster Schreiber et al. 2019).

(ii) For Hα + [N II] and [S II] coverage ($0.30 < z_{\rm grism} < 1.45$), we bin 1191 galaxies and examine their position on the galaxy MS (SFR versus $M_\star$; Leslie et al. 2020). Bins with log ($M_\star$/M$_\odot$) $\gtrsim 9$ appear roughly consistent with the 'star-forming MS', indicating they are representative of 'normal' star-forming galaxies. At log ($M_\star$/M$_\odot$) $\lesssim$ 9, our bins reside above the MS.

(iii) For Hα + [N II], [S II], and [S III] coverage ($0.30 < z_{\rm grism} < 0.73$), we bin 318 galaxies to examine the [S III]/[S II] ratio (ionization parameter proxy). Our stacks are consistent with ratios found in MaNGA ($z \sim 0$) and MOSDEF ($z \sim 1.5$; Sanders et al. 2020) if WISP galaxies ($z \sim 0.5$) contain a similar [S II] contribution from DIG as MaNGA galaxies.

(iv) For [O II] to [S II] coverage ($1.27 < z_{\rm grism} < 1.45$), we bin 203 galaxies to examine the [O III]/[O II] ratio (ionization parameter/spectral slope proxy) and ([O II] + [O III])/Hβ ratio (metallicity proxy). Our stacks are consistent with line ratios found in CLEAR ($1.1 < z < 2.3$; Papovich et al. 2022). In particular, the relative strength of [O III] emission is substantially higher than in local star-forming galaxies observed by SDSS.

These results indicate that the majority of WISP galaxies are representative of typical MS galaxies.

Finally, we note that several current large-area photometric surveys, such as the Dark Energy Survey (*grizY* bands; Dark Energy Survey Collaboration 2016), the DECam Legacy Survey of the SDSS Equatorial Sky (*grz* bands; Dey et al. 2019), the Beijing–Arizona Sky Survey (*gr* bands; Zou et al. 2017), and the Mayall z-band Legacy Survey (MzLS, *z* band; Dey et al. 2019), provide shallow coverage (∼22–24 AB mag) for some of the WISP fields that do not have existing coverage in this data release. In the future, the Legacy Survey of Space and Time (*ugrizy* bands; Ivezić et al. 2019) with the Rubin Observatory will provide more uniform and deeper coverage (final depth ∼25–27 AB mag) that will supersede the depth of most existing optical data for WISP fields in the Southern Hemisphere (Dec. < 0). Incorporating these data sets may be the subject of a future data release.

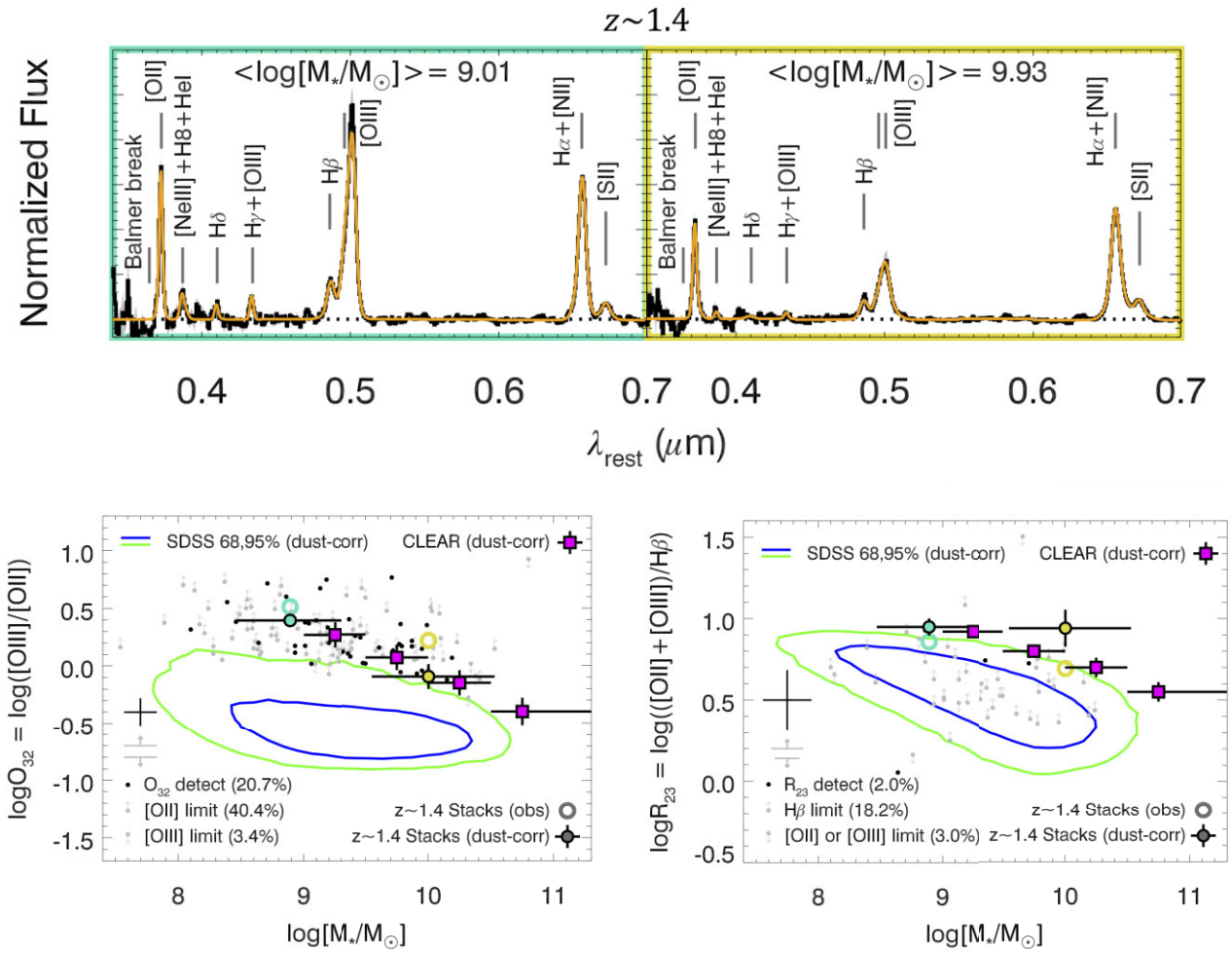


Figure 19. Top: similar to Fig. 16, but now for the window covering [O II] to [S II] and the stacked spectra are normalized to Hα + [N II] total flux. The numerous optical emission lines in this window allows several ISM diagnostics to be simultaneously available. The region around the Balmer break (0.3645 μm) shows a discontinuity, which may result from the simplistic treatment of the stellar continuum (see Section 5.3). Bottom: we compare the O_{32} and R_{23} ratios versus stellar mass from WISP ($z \sim 1.4$) to values from CLEAR ($1.1 < z < 2.3$; Papovich et al. 2022) and also show contours for SDSS ($z < 0.2$; coloured lines). Coloured open and closed symbols correspond to the stellar mass stacks before and after dust extinction correction, respectively, based on the measured Balmer decrement (see Section 7.4.4). The WISP stacks are in rough agreement with CLEAR, lying above the SDSS sample, and showing a decreasing O_{32} and R_{23} with increasing $\log M_\star$. Larger O_{32} ratios indicate of higher ionization potential and/or harder ionizing spectrum. Larger R_{23} indicate lower metallicities (for upper branch, $12 + \log(\mathrm{O/H}) > 8.1$), however the R_{23}–metallicity relation has a turnover at $\log R_{23} \sim 1$ (e.g. Curti et al. 2017).

Table 15. Average properties, emission-line luminosities, and line ratios for stacked spectra in the [O II] to [S II] window ($1.27 < z < 1.45$).

N	$\log M_\star$ range	$\langle \log M_\star \rangle$	[O II]	Hβ	[O III]	Hα + [N II]	[S II]	H$\alpha_{\rm obs}$	τ_B^l	log O_{32}	log R_{23}
					(10^{41} erg s^{-1})						
102	[7.54, 9.39]	9.01	5.03 ± 0.24	2.46 ± 0.14	16.41 ± 0.31	10.11 ± 0.09	1.45 ± 0.09	9.28 ± 0.26	0.27 ± 0.07	0.40 ± 0.06	0.95 ± 0.05
101	[9.40,11.26]	9.93	6.56 ± 0.35	2.39 ± 0.21	10.83 ± 0.32	18.29 ± 0.16	3.29 ± 0.11	14.50 ± 1.72	0.71 ± 0.19	-0.10 ± 0.11	0.94 ± 0.11

Notes. Hβ shown is not corrected for stellar absorption. [O III] is the sum of [O III]$\lambda\lambda$4959, 5007. [S II] is the sum of [S II]$\lambda\lambda$6716, 6731. The Balmer optical depth, $\tau_B^l = \ln((F(\mathrm{H}\alpha)/F(\mathrm{H}\beta)/2.86)$ is based on absorption-corrected lines (see Section 7.4.4). Values listed for O_{32} and R_{23} are after extinction correction.

ACKNOWLEDGEMENTS

The authors thank the anonymous referee, whose suggestions helped to clarify and improve the content of this work. AJB thanks the staff at the WIYN and Magellan facilities for their assistance with technical aspects of the observations. AJB is also thankful for attending ASTRO 3D writing retreats that provided a helpful environment to complete portions of this manuscript. Parts of this research were supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project no. CE170100013. We acknowledge the invaluable labour of the maintenance and clerical staff at our institutions, whose contributions make our scientific discoveries a reality. This research was conducted on Ngunnawal Indigenous land, as well as within the traditional homelands of the Dakota people. This work is sponsored by the National Key R&D Program of China for grant no. 2022YFA1605300, the National Nature Science Foundation of China (NSFC) grants nos 12273051 and 11933003. HA is supported by CNES (Centre National d'Etudes Spatiales). AJ Bunker acknowledges funding from

the 'FirstGalaxies' Advanced Grant from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 789056). Support for WISP (*HST* programs GO-11696, 12283, 12568, 12902, 13517, 13352, and 14178) was provided by NASA through grants from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. This work uses data from WIYN. The WIYN Observatory is a joint facility of the NSF's National Optical-Infrared Astronomy Research Laboratory, Indiana University, the University of Wisconsin-Madison, Pennsylvania State University, the University of Missouri, the University of California-Irvine, and Purdue University. This paper includes data gathered with the 6.5 meter Magellan Telescopes located at Las Campanas Observatory, Chile. This work uses data from Palomar. The Palomar Observatory is a joint facility of the Caltech/Caltech Optical Observatories, Jet Propulsion Laboratory, Yale University, and National Astronomical Observatories of China. This work uses data from SDSS. Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University. This work uses data from PAN-STARRS. The Pan-STARRS1 Surveys (PS1) and the PS1 public science archive have been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, the Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under grant no. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation grant no. AST–1238877, the University of Maryland, Eotvos Lorand University (ELTE), the Los Alamos National Laboratory, and the Gordon and Betty Moore Foundation.

DATA AVAILABILITY

All of the WISP photometric and spectroscopic data used in this paper are publicly available through data releases on the WISP website.[16] on *MAST*. This release adds the following data products to the WISP website: (1) fully reduced 5 arcmin × 5 arcmin cutouts (centred on WISP field) of ground-based observations, (2) fully reduced *Spitzer* images, (3) the photometric catalogue (FITS binary table; described in Table 6), (4) the emission-line catalogue (FITS binary table; described in Table 10), (5) the full versions of Table 2 (grism depths for each field), Table 7 (WFC/IR photometric depths/completeness for each field), Table 8 (photometric depths for each field), and Table 11 (MAGPHYS properties for subsample). Other data products can be made available upon reasonable request to the first author.

[16] https://archive.stsci.edu/prepds/wisp/

REFERENCES

Anders P., Fritze-v. Alvensleben U., 2003, A&A, 401, 1063
Astropy Collaboration, 2013, A&A, 558, A33
Astropy Collaboration, 2018, AJ, 156, 123
Atek H. et al., 2010, ApJ, 723, 104
Atek H. et al., 2014, ApJ, 789, 96
Bagley M. B. et al., 2017, ApJ, 837, 11
Bagley M. B. et al., 2020, ApJ, 897, 98
Bagley M. B. et al., 2024, ApJ, 961, 209
Baronchelli I. et al., 2020, ApJS, 249, 12
Baronchelli I. et al., 2021, ApJS, 257, 67
Battisti A. J., Cunha E. d., Shivaei I., Calzetti D., 2020, ApJ, 888, 108
Battisti A. J. et al., 2022, MNRAS, 513, 4431
Bedregal A. G. et al., 2013, ApJ, 778, 126
Berg D. A., Chisholm J., Erb D. K., Skillman E. D., Pogge R. W., Olivier G. M., 2021, ApJ, 922, 170
Bertin E., 2010, Astrophysics Source Code Library, record ascl:1010.068
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bertin E., Mellier Y., Radovich M., Missonnier G., Didelon P., Morin B., 2002, in Bohlender D. A., Durand D., Handley T. H., eds, ASP Conf. Ser. Vol. 281, Astronomical Data Analysis Software and Systems XI. Astron. Soc. Pac., San Francisco, p. 228
Bourque M., Baggett S., 2016, WFC3/UVIS Dark Calibration: Monitoring Results and Improvements to Dark Reference Files, Instrument Science Report WFC3 2016–08, The Association of Universities for Research in Astronomy, Inc.
Brammer G. B. et al., 2012, ApJS, 200, 13
Brammer G., Ryan R., Pirzkal N., 2015, Source-dependent Master Sky Images for the WFC3/IR Grisms, Instrument Science Report WFC3 2015–17, The Association of Universities for Research in Astronomy, Inc.
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Calzetti D., 2001, PASP, 113, 1449
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
Chabrier G., 2003, PASP, 115, 763
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Coil A. L. et al., 2015, ApJ, 801, 35
Colbert J. W. et al., 2013, ApJ, 779, 34
da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595
da Cunha E. et al., 2015, ApJ, 806, 110
Curti M., Cresci G., Mannucci F., Marconi A., Maiolino R., Esposito S., 2017, MNRAS, 465, 1384
Dai Y. S. et al., 2021, ApJ, 923, 156
Dark Energy Survey Collaboration, 2016, MNRAS, 460, 1270
Dey A. et al., 2019, AJ, 157, 168
Dickinson H. et al., 2018, Res. Notes Am. Astron. Soc., 2, 120
Domínguez A. et al., 2013, ApJ, 763, 145
Du P., Kibbe W. A., Lin S. M., 2006, Bioinformatics, 22, 2059

Faisst A. L., Masters D., Wang Y., Merson A., Capak P., Malhotra S., Rhoads J. E., 2018, ApJ, 855, 132
Fitzpatrick E. L., Massa D., Gordon K. D., Bohlin R., Clayton G. C., 2019, ApJ, 886, 108
Förster Schreiber N. M. et al., 2019, ApJ, 875, 21
Gonzaga S., Hack W., Fruchter A., Mack J., 2012, The DrizzlePac Handbook, The Association of Universities for Research in Astronomy, Inc.
Henry A. et al., 2013, ApJ, 776, L27
Henry A. et al., 2021, ApJ, 919, 143
Hoffmann S. L., Mack J., Avila R., Martlin C., Cohen Y., Bajaj V., 2021, American Astronomical Society Meeting Abstracts. p. 216.02
Ivezić Ž. et al., 2019, ApJ, 873, 111
Juneau S. et al., 2014, ApJ, 788, 88
Kashino D., Inoue A. K., 2019, MNRAS, 486, 1053
Kashino D. et al., 2019, ApJS, 241, 10
Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531
Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, ARA&A, 57, 511
Kroupa P., Weidner C., 2003, ApJ, 598, 1076
Kümmel M., Kuntschner H., Walsh J., 2007, Space Telesc. Eur. Coord. Facil. Newsl., 43, 8
Kümmel M., Walsh J. R., Pirzkal N., Kuntschner H., Pasquali A., 2009, PASP, 121, 59
Leja J., Carnall A. C., Johnson B. D., Conroy C., Speagle J. S., 2019, ApJ, 876, 3
Leslie S. K. et al., 2020, ApJ, 899, 58
Lucy L. B., 1974, AJ, 79, 745
Malkan M. A. et al., 2021, PASSAGE-Parallel Application of Slitless Spectroscopy to Analyze Galaxy Evolution, JWST Proposal. Cycle 1, STSCI
Masters D. et al., 2012, ApJ, 752, L14
Masters D. et al., 2014, ApJ, 785, 153
Mehta V. et al., 2015, ApJ, 811, 141
Merlin E. et al., 2015, A&A, 582, A15
Merlin E. et al., 2016, A&A, 595, A97
Moré J. J., 1978, The Levenberg-Merquardt Algorithm: Implementation and Theory. Springer-Verlag, Berlin
Oke J. B., Gunn J. E., 1983, ApJ, 266, 713
Osterbrock D. E., 1989, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei. University Science Books, Mill Valley, CA
Papovich C. et al., 2022, ApJ, 937, 22
Pasquali A., Pirzkal N., Walsh J. R., Rhoads J. E., Malhotra S., Tsvetanov Z., Apples Team, 2003, in Shaver P. A., Dilella L., Giménez A., eds, Astronomy, Cosmology and Fundamental Physics. Springer-Verlag, Berlin, p. 471
Pirzkal N. et al., 2004, ApJS, 154, 501
Pirzkal N. et al., 2013, ApJ, 772, 48
Pirzkal N. et al., 2017, ApJ, 846, 84
Pirzkal N. et al., 2018, ApJ, 868, 61
Prichard L. J. et al., 2022, ApJ, 924, 14
Rafelski M. et al., 2015, AJ, 150, 31
Revalski M. et al., 2023, ApJS, 265, 40
Richardson W. H., 1972, J. Opt. Soc. Am. (1917–1983), 62, 55
Sanders R. L. et al., 2020, ApJ, 888, L11
Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103
Shapley A. E. et al., 2022, ApJ, 926, 145
Simons R. C. et al., 2023, ApJS, 266, 13
Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJS, 214, 15
Steidel C. C., Strom A. L., Pettini M., Rudie G. C., Reddy N. A., Trainor R. F., 2016, ApJ, 826, 159
Storey P. J., Zeippen C. J., 2000, MNRAS, 312, 813
Tonry J. L. et al., 2012, ApJ, 750, 99
Treu T. et al., 2015, ApJ, 812, 114
van der Wel A. et al., 2014, ApJ, 788, 28
Wang X. et al., 2022, ApJ, 926, 70
Whitaker K. E. et al., 2014, ApJ, 795, 104
Zanella A. et al., 2016, ApJ, 824, 68
Zou H. et al., 2017, PASP, 129, 064101

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

suppl_data

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: TPHOT INPUT AND CONFIGURATION DETAILS

In order for TPHOT to obtain reliable measurements, the geometric centres of the sources detected in different bands must be perfectly aligned. To obtain this result, the coordinates of the sources originally extracted from the low-resolution images are refined by comparing the positions of the sources in the images at different wavelengths. This coordinate recentring is performed using an iterative procedure. Initially, source positions in the low-resolution data are determined by running SExtractor on the original images. The most likely *HST* counterparts are identified using an initial search radius of 2 arcsec. At this stage, each source is associated to only one potential counterpart. Then, an average shift correction for the coordinates is computed by comparing the (RA, Dec.) positions of the counterparts in the two bands and the corresponding new set of WCS is applied to the low-resolution images. After this initial iteration, the entire process is repeated using, this time, a shorter searching radius (0.75 arcsec). Therefore, the WCS correction is refined by performed the second iteration using only sources with a more secure counterpart identification. At the end of the process described, the coordinates in the low-resolution images are precisely recentred to the *HST* reference frame.

TPHOT also requires input images with identical (or integer multiple) pixel-scale and identical pixel orientation. We use the SWARP software to obtain low-resolution images consistent with the *HST* reference images. We note that these SWARP intermediate products (i.e. images with identical pixel scales and orientations) are not provided in the current data release. TPHOT also requires both a catalogue of sources extracted in the high-resolution image (*HST*), and a corresponding segmentation map. In particular, the IDs reported in the input catalogue must be identical to the pixel values of the corresponding sources in the segmentation map. Our catalogues were created by merging catalogues of sources detected in the *J* and/or in the *H* band (plus a catalogue obtained from a sum of the $J + H$ images). As a consequence, the IDs in our catalogues had to be re-assigned (and ordered on a brightness basis). Finally, not all the sources extracted (and included in the segmentation maps) were included in our catalogues. For example, we removed all the sources located in the image borders, characterized by bad photometric measurements.

Due to the fact that not all sources uniquely correspond to a counterpart across the different *HST* segmentation maps, we computed an appropriate unique segmentation image, from the original segmentation maps obtained by SExtractor. In the new segmentation map, every source of the catalogue is uniquely related with a corresponding source (with identical ID) in the map. Moreover, all the sources removed from the catalogues are also removed from the final segmentation map.

In addition to the IDs, the TPHOT input catalogue must include, for each source, its (*x*,*y*) position in the *HST* reference image, the values $x_{\rm min}$, $x_{\rm max}$, $y_{\rm min}$, and $y_{\rm max}$ defining the borders of the source in

the segmentation map, a local background value (we set this value to 0 as the background was already subtracted from the images before running TPHOT), and the reference flux for each source in the high-resolution band.

One of the most important inputs required by TPHOT is the image of the kernel required to perform the inverse of the following convolution operation:

$$\mathrm{PSF(LRI)} = \mathrm{Kernel} * \mathrm{PSF(HRI)}, \tag{A1}$$

where LRI and HRI stand for low- and high-resolution images, respectively. We obtain the convolution kernel image by deconvolving the low-resolution PSF using the *HST* PSF as a deconvolution kernel. For this, we adopt the Richardson–Lucy approach (Richardson 1972; Lucy 1974). The average low-resolution and *HST* PSFs are obtained averaging the PSFs of a selection of point-like, non-saturated sources detected in each image. In accordance with this choice, we set the TPHOT configuration parameter `usereal = True` (and `usemodels` and `useunresolved = False`).

The TPHOT algorithm is organized in different 'stages', each of which performs a specific task. The list and the order of the various stages in each iteration can be set by the user. As described in the user manual, the best results are obtained when the various stages are run in two separate iterations ('passes'). We proceed in the default manner, setting the keyword `order` to `standard` and to `standard2` in the two iterations, respectively.

For each source, TPHOT computes (x,y) shifts during the first pass ('dance' stage). Using these corrections, in the second pass, local kernels are registered to each single source improving the accuracy of the outputs. This can be obtained by setting the keyword `multikernel = true` for the second iteration. We set the size of the region in which the PSF shift is computed (keyword `dzonesize`) to 100 pixels and the maximum shift allowed (keyword `maxshift`) to 20 pixels. We reduce these two parameters to 50 and 1, respectively, for the second iteration. The shifts computed are smoothed over 100 and 50 neighbours in the two passes, respectively (keyword `nneighinterp`). Among the three different methods available for solving the linear systems during the fitting stage (i.e. LU, Cholesky and the iterative biconjugate gradient), we selected the matrix inversion method of LU (default). Additionally, we select the option to clip out large negative fluxes before obtaining the final fit. We set all the remaining TPHOT parameters to the default options.

APPENDIX B: WISP PHOTOMETRIC CATALOGUE FLAG DESCRIPTION

There are two entries for quality flags in the photometric catalogue, each consisting of a sum of bit flags (i.e. sum of powers of 2). One is for the default SExtractor 'internal' flags[17] (`FLAG_[NIRFILTER]`) and the other is for the default TPHOT flags (Merlin et al. 2015, `TPHOT_FLAG_[FILTER]`). These are described below.

`FLAG_[NIRFILTER]` contains a sum of 8 flag bits (i.e. sum of powers of 2):

(i) 1 = photometry likely to be biased by neighbouring sources or bad pixels;
(ii) 2 = object has been deblended;
(iii) 4 = at least one object pixel is saturated;
(iv) 8 = object is close to image boundary;

[17] https://astromatic.github.io/sextractor/Flagging.html

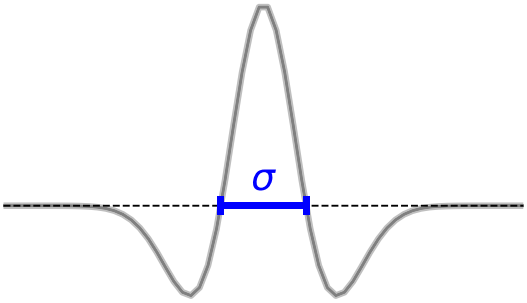


Figure C1. The Ricker wavelet used for emission-line detection in the CWT. It is proportional to the second derivative of a Gaussian function and is defined by a width parameter, σ.

(v) 16 = at least one photometric aperture is incomplete or corrupted;
(vi) 32 = the isophotal footprint is incomplete or corrupted;
(vii) 64 = a memory overflow occurred during deblending;
(viii) 128 = a memory overflow occurred during extraction.

For example, a saturated detection close to an image boundary will have `FLAG_[NIRFILTER]` $= 4 + 8 = 12$.

`TPHOT_FLAG_[FILTER]` contains a sum of three flag bits:

(i) 1 = the prior has saturated or negative flux;
(ii) 2 = the prior is blended;
(iii) 4 = the source is at the border of the image.

APPENDIX C: WISP CONTINUOUS WAVELET TRANSFORM ALGORITHM

Starting from the 1D spectra extracted and calibrated by AXE, we perform a CWT on the spectrum using a Ricker wavelet, which is proportional to the second derivative of a Gaussian function. The Ricker wavelet models the function

$$f = \frac{2}{\sqrt{3\sigma}\pi^{\frac{1}{4}}}\left(1 - \frac{x^2}{\sigma^2}\right) e^{-\frac{x^2}{2\sigma^2}}, \tag{C1}$$

and is illustrated in Fig. C1.

The transform is performed using $N_\sigma = 10$ kernels or wavelets of varying widths, with the minimum width set to 1.5 pixels, corresponding to ∼36 Å (∼70 Å) in *G102* (*G141*), and the maximum width equal to twice the FWHM estimate for the object. Spectra obtained in slitless mode are essentially images of the source at each wavelength, and so the width of emission lines will be correlated with the source size. We approximate the FWHM as twice the dispersion measured along the semimajor axis (`A_IMAGE` reported by SExtractor), and use this $\mathrm{FWHM_{est}}$ to define upper bounds for both the wavelet widths and the FWHM of the emission-line profile fits described in Section 5.3.

The CWT transform compares the wavelet with the 1D spectrum, shifting the wavelet to cover all wavelengths and scaling or stretching to cover all input widths. The resulting array of CWT coefficients is a matrix of dimension $N_\sigma \times N_\lambda$ representing the correlation of the spectrum and wavelet at each scale and wavelength. Large CWT coefficients indicate regions of the spectrum with a strong correlation with the wavelet, and so the largest coefficients will occur where both the position and width of the wavelet best match a spectral feature. An example coefficient matrix is displayed in the bottom panel of Fig. C2.

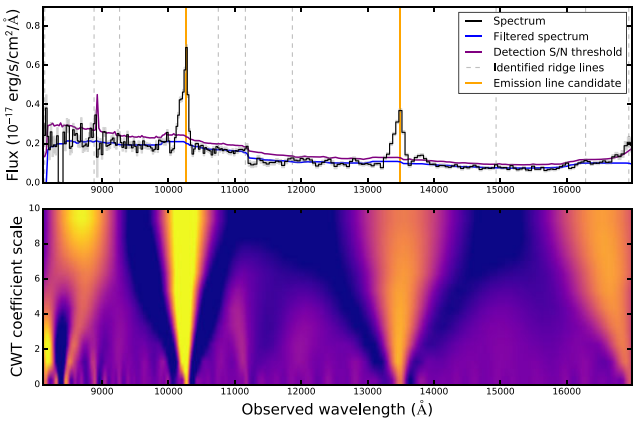


Figure C2. An example of the emission-line detection algorithm. Top panel: The input spectrum is plotted in black. The filtered spectrum (blue curve) is used as an estimate of the continuum. The purple curve is the detection threshold, defined as a minimum S/N ≥ 2.31 above the continuum. Bottom panel: the 2D CWT coefficient matrix represents the correlation of the spectrum and wavelet at each scale and wavelength. Larger coefficients (yellow colours) identify regions of higher correlation. Emission lines are correlated with the wavelet at many scales, creating connected ridge lines in the CWT matrix. The dashed grey lines in the top panel show the positions of detected ridge lines. The yellow solid lines show the ridges that passed our additional selection criteria to be identified as true emission-line candidates.

As can be seen in the bottom panel of Fig. C2, strong emission peaks in the spectrum are strongly correlated with the Ricker wavelet at many scales. The resulting peaks in the 2D CWT coefficient matrix extend to multiple scales and can be visualized as mountain ridges. Emission-line features in the spectrum can now be identified using the ridges in the CWT coefficient matrix. For this step, we use the SCIPY program `find_peaks_cwt`, which is an implementation of the procedure presented in Du, Kibbe & Lin (2006). We briefly summarize the process here and refer the reader to Du et al. (2006) for a full description of the algorithm.

The first step in the peak identification process involves identifying ridge lines in the coefficient matrix. At each scale, a local maximum is matched with the nearest maximum calculated for the adjacent scale. The two local maxima constitute a ridge provided they are within a set distance, $\mathrm{dist}_{\mathrm{max}}$, which we define as one pixel larger than the width of the wavelet at the given scale. A ridge line continues through the matrix as long as the local maximum at each scale is within d_{max} of the previously identified maximum. The ridge line is terminated if a local maximum is not matched at more than g_{thresh} consecutive kernels. We conservatively use $g_{\mathrm{thresh}} = 1$ kernel.

The second step is to identify peaks from the detected ridge lines. We only consider ridge lines that cover at least $l_{\mathrm{min}} = 3$ scales and have a ridge signal-to-noise of $\mathrm{S/N}_{\mathrm{min,\,CWT}} \geq 1$. The ridge S/N is estimated at each scale, and is taken to be the maximum CWT coefficient value on the ridge divided by the overall noise floor. The noise floor is calculated as the 10th percentile of the coefficients around the ridge line at the smallest scale (σ_{min}). These identified peaks are represented as dashed grey lines in the top panel of Fig. C2.

Following the peak finder, we additionally screen the selected peaks to remove sources that are likely to be spurious. The spectrum is filtered to remove noise using a median filter with a window size of 31 pixels, corresponding to ∼760 Å in *G102* and ∼1440 Å in *G141*. The purpose of this filtering is to obtain a crude measure of the continuum, which is used in evaluating the strength of any identified peaks. The large window is chosen so that emission lines will not significantly affect the continuum measurement. From this continuum estimate, the contrast between the peak and the continuum is calculated $C = (f_{\lambda,\,\mathrm{peak}} - f_{\lambda,\,\mathrm{continuum}})/f_{\lambda,\,\mathrm{continuum}}$, and any peak with $C < C_{\mathrm{min}} = 0.15$ is rejected. This criterion is approximately equivalent to a cut on very low EW emission lines and is implemented in order to remove noise spikes. Next, all peaks that overlap with zeroth orders of nearby bright ($m < 23.5$ in the corresponding direct image) sources are rejected. Recall that we also do not consider any spectra on the right edge of the detector where we cannot determine the position of zeroth orders. We additionally remove peaks that are within $d_{\mathrm{edge}} = 5$ pixels of the edge of the spectrum to avoid convolution edge effects. Finally, we require that emission-line candidates have an overall signal-to-noise for the emission line of $\mathrm{S/N}_{\mathrm{line}} \geq 4$. For a line consisting of three contiguous pixels, this corresponds to $\mathrm{S/N}_{\mathrm{min,\,pix}} \geq 2.31$, where here S/N is defined traditionally as the flux in the continuum divided by the error. This resulting noise threshold is displayed as the purple curve in Fig. C2. A summary of the emission line detection parameters is provided in Table C1.

Table C1. Emission-line detection parameters.

Parameter	Value	Description
	Continuous wavelet transform	
N_σ	10	Number of CWT kernels (widths) used in transform
σ_{min}	1.5 [pixels]	Minimum CWT kernel width
σ_{max}	2.0	Maximum CWT kernel width, fraction of $\mathrm{FWHM}_{\mathrm{est}}$
	Ridge and peak definitions	
d_{max}	$\sigma + 1$ [pixels]	Maximum acceptable separation distance between local maxima at each scale on the same ridge line
g_{max}	1 [kernels]	Maximum acceptable gap between connected ridges
l_{min}	3 [kernels]	Minimum acceptable no. of kernels in which peak is significant
$\mathrm{S/R}_{\mathrm{min,\,CWT}}$	1	Minimum acceptable ridge S/N at each scale
	Requirements for real lines	
C_{min}	0.15	Minimum acceptable contrast between peak and continuum fluxes
d_{edge}	5 [pixels]	Minimum acceptable distance to edge of spectrum
n_{min}	3 [pixels]	Minimum number of pixels above the noise threshold
$\mathrm{S/N}_{\mathrm{min,\,pix}}$	2.31	Minimum S/N per pixel

APPENDIX D: WISP SURVEY COMPLETENESS SIMULATIONS

The selection function in slitless spectroscopic data is complex, depending on line S/N, EW, galaxy size and concentration, and observed wavelength. Extensive simulations exploring the full parameter space are required. Moreover, the completeness strongly depends on the observing strategy and depth of the data sets. We therefore perform two sets of simulations, one each for the shallow and deeper WISP fields. We add synthetic sources to a selection of real WISP fields and reprocess the fields through the WISP pipeline and line finding procedures presented in Section 5. We describe the full process below.

We create 10 000 simulated galaxies and divide them equally between the shallow and deep fields. All sources are assigned a spectral template from the models of Bruzual & Charlot (2003) created with a Chabrier (2003) IMF, a constant SFH observed 100 Myr after the beginning of star formation, and one of three metallicities: $Z/Z_{\odot}$ =0.02, 0.2, or 1. The template spectra are redshifted to the observed frame such that all emission-line and continuum fluxes are normalized to the desired observed values. We add the following emission lines to the spectra: [O II]λ3727, [O III]λ4363, Hβ, [O III]λ4959, [O III]λ5007, Hα, [S II]λ6716, [S II]λ6730, [S III]λ9069, and [S III]λ9531. Each emission line is modelled as a Gaussian with $\sigma = 3$ Å.

The basic question addressed by completeness corrections is whether an input source or emission line is recovered by the reduction and processing performed on real data. As it is not a measure of the rates of source misclassification or redshift misidentification, the inputs need not represent the physical distributions observed in the universe. Unless otherwise noted, we therefore uniformly populate the input parameter space so we can determine the ranges most affected by incompleteness.

For each parameter, we choose input ranges that bracket the observed values. Source redshifts are pulled from a uniform distribution ranging from the redshift at which Hα enters the wavelength coverage to that at which [O III]λ5007 leaves the wavelength coverage.[18] We adopt conservative wavelength cutoffs for each grism to avoid the wavelengths where the sensitivity drops rapidly: $8500 \leq \lambda_{G102} \leq 11200$ Å and $11200 \leq \lambda_{G141} \leq 16500$ Å for the *G102* and *G141* grisms, respectively. Real emission lines that lie outside of these wavelength ranges are flagged in the catalogue (see Appendix E for details). These cut-offs correspond to redshift ranges of $0.3 \leq z \leq 2.3$ for the deep fields with coverage in both grisms and $0.7 \leq z \leq 2.3$ for the shallow fields. Input Hα fluxes for sources in shallow fields are pulled uniformly from $5 \times 10^{-17} \leq f_{\mathrm{H}\alpha} \leq 1 \times 10^{-15}$ erg s^{-1} cm^{-2}, with a lower limit of 1×10^{-17} erg s^{-1} cm^{-2} in the deep fields. We increase the number of faint line fluxes in the deep fields by separating the synthetic sources into two groups: half with fluxes pulled from a uniform distribution with a maximum at 1×10^{-16} erg s^{-1} cm^{-2} and half with fluxes extending up to 1×10^{-15}. Using two upper flux limits allows us to populate the bright end where sources should be easily detected in the deep fields, while ensuring we have an adequate number of faint objects even if the recovered fraction is small. The input observed EW distribution is uniform across the range $20 \leq \mathrm{EW}_{\mathrm{H}\alpha,\,\mathrm{obs}} \leq 700$ Å. The flux density in the continuum at the observed wavelength of Hα – that is, the ratio of input Hα flux and EW – is used to normalize the spectral template to the desired observed units and brightness.

There is a range of observed emission-line ratios in the WISP catalogue, which we account for in the simulations by varying the input Hα/[O III]λ5007 ratios in the synthetic spectra. Although we are not attempting to quantify the redshift misidentification in the catalogue, simulating emission lines with a variety of flux ratios is necessary to include any biases related to single versus multiline emitters. For example, reviewers are more likely to identify low S/N lines if there are additional emission lines visible in the spectrum to confirm the source's redshift. Input $\log_{10}$(Hα/[O III]λ5007) ratios are drawn from a Gaussian distribution centred at $\mu = 0$ with $\sigma = 0.2$, matching the observed distribution in the catalogue but with a slightly larger FWHM. The intrinsic Hα/Hβ ratio for case B recombination is adopted for all sources, 2.86 (Osterbrock 1989). For the remaining emission lines, we adopt the ratios from Anders & Fritze-v. Alvensleben (2003) with respect to [O III]λ5007 for each metallicity assuming an electron density of $n_e = 100$ cm^{-3} and electron temperature of $T_e = 10\,000$ K. The input *H*-band magnitudes are not assigned to the sources, but instead are the result of the flux normalization of the template spectral continua and depend on the Hα fluxes and EWs of each source. We do not add the effects of dust to the spectra, but instead rely on the range of line ratios to cover observed values. We note that the adopted case B Balmer decrement, Hα/Hβ, is the only line ratio that remains unchanged and therefore always exhibits the value expected for dust-free galaxies. However, we remind the reader that as we are not trying to replicate reality, this choice will not affect the resulting completeness calculations.

The completeness of sources in the WISP catalogue depends strongly on object size and shape. Object size first affects the completeness in imaging, where the low surface brightness of faint, extended objects may fall below the adopted SExtractor detection thresholds, while the higher surface brightnesses of more compact sources are detected. In addition, the pipeline removes the most extremely elongated detected sources from the catalogue in an attempt to remove artefacts such as diffraction spikes and persistence from bright first orders. Since emission lines observed via slitless spectroscopy are essentially images of the sources at the given wavelengths, the source shape and size will also affect the completeness of the line finding procedure. Extended sources with a low EW are missed by the peak finder, and reviewers are more consistent with their treatment of compact, high-S/N emission lines. We therefore assign each object a profile RMS along the major (minor) axis pulled from a uniform distribution in the range $0.05 \text{ arcsec} \leq a \leq 1.2 \text{ arcsec}$ ($0.05 \text{ arcsec} \leq b \leq a$), again matching the range but not the shape of the observed distribution in the catalogue. The input parameter distributions are summarized in Table D1.

We add 25 simulated sources at random locations to the raw images for a set of WISP fields, using the AXESIM (Kümmel et al. 2007) software package to create the synthetic direct and grism images of each source. These fields are then fully processed as real data, including the visual inspection by two reviewers. As mentioned in Section 5.4, in order to save on the time and effort required for this step, the reviewers only inspect the spectra of simulated sources that were identified by the line finding algorithm. This is not to say that all emission-line candidates were real. Some were noise spikes, contamination, or the result of poorly fit continua. However, it does mean we cannot use the simulations to measure the rates of contamination or redshift misidentification in the catalogue.

[18] We do not add any emission lines to the spectrum in the two narrow wavelength ranges that AXESIM uses for spectral normalization: 10 400–10 600 (*G102*) and 15 400–15 600 Å (*G141*). Emission lines in these ranges would artificially boost the normalization factor thereby significantly reducing the simulated continuum of the spectrum. For more information, see the AXESIM (Kümmel, Kuntschner & Walsh 2007) manual available at axe.stsci.edu/axesim/. There are therefore several narrow redshift ranges that we do not populate with synthetic sources.

Table D1. Input parameters for simulated sources.

Parameter	Deep fields	Shallow fields
Redshift	$0.3 \leq z \leq 2.3$	$0.7 \leq z \leq 2.3$
Semimajor axis[a]	0.05 arcsec $\leq a \leq$ 1.2 arcsec	0.05 arcsec $\leq a \leq$ 1.2 arcsec
Elongation[b]	$1 \leq a/b \leq 24$	$1 \leq a/b \leq 24$
Observed Hα flux (erg s^{-1} cm^{-2})	$1 \times 10^{-17} \leq f \leq 1 \times 10^{-15}$	$5 \times 10^{-17} \leq f \leq 1 \times 10^{-15}$
Observed Hα EW (Å)	$20 \leq \mathrm{EW_{obs}} \leq 700$	$20 \leq \mathrm{EW_{obs}} \leq 700$
Hα/[O III]λ5007 ratio	Gaussian: $\mu = 0, \sigma = 0.2$	Gaussian: $\mu = 0, \sigma = 0.2$
H-band magnitude[c]	$16.8 \leq m_H \leq 27.6$	$16.8 \leq m_H \leq 26.2$

Notes. Parameters indicated with a range are uniformly populated between that range.
[a]The semimajor axis a is treated as the profile RMS along the major axis.
[b]The semiminor axis b values are pulled from the range 0.05 arcsec $\leq b \leq a$, and so these elongations a/b list the minimum and maximum possible values.
[c]The *H*-band magnitude is computed from the input template spectra normalized according to the Hα flux and EW. The m_H here list the minimum and maximum possible values.

The WISP completeness calculations, described in Section 5.4, involve determining the fraction of sources that have been recovered in bins of source size and emission-line flux and EW. After the simulated sources are fully processed through the WISP software, we compare all properties of the recovered sources with their input values. This step is necessary to confirm that the recovered sources are counted in the proper bins, that is, that the parameter values are not systematically different due to the simulation and reduction processes. The input and output fluxes for Hα and [O III] show a clear correlation down to $f \sim 7 \times 10^{-17}$ erg s^{-1} cm^{-2}. At fainter fluxes, there is a slight trend toward brighter output fluxes, which may be caused by spectral contamination. Overlapping spectra can boost a source's measured emission-line flux, and we expect the severity of this contamination to increase with decreasing line flux. Yet there are far too few recovered Hα or [O III] lines with fluxes $<7 \times 10^{-17}$ erg s^{-1} cm^{-2} to properly evaluate the trend. The other emission lines are almost always fit as secondary lines and will therefore be fainter and at a lower S/N than the primary lines. There is a similarly good agreement between input and output Hα and [O III] EWs.

A comparison of the source sizes, however, shows that the inputs are systematically larger than the outputs. This effect is not surprising, since the flux in the wings of the simulated Gaussian sources can fall below the SExtractor detection limit. The extracted 'footprints' of the sources are then smaller than what was simulated. We must understand the relationship between the input and output sizes in order to properly determine the number of sources that are recovered as a function of size. We model this relationship with a combination of a fourth-order polynomial for semimajor and minor axes $\lesssim 0.6$ arcsec and a linear fit for larger sizes. The WISP survey completeness is applied to sources according to their observed fluxes, EWs, and sizes, as their intrinsic values are unknown. The completeness is therefore calculated as a function of the output values measured for the simulated sources rather than the input values. We use the models to scale the input a and b to their measured values. This step is necessary to ensure that input sources that are not recovered are counted in the correct bins.

APPENDIX E: WISP EMISSION-LINE CATALOGUE FLAG DESCRIPTION

There are nine entries for quality flags in the emission-line catalogue, all consisting of a sum of bit flags (i.e. sum of powers of 2). These are described below.

`FILTER_FLAG` indicates the filter coverage as a sum of 10 flag bits:

(i) 1 = F110 coverage;
(ii) 2 = F140 coverage;
(iii) 4 = F160 coverage;
(iv) 8 = UVIS1 coverage (*F475X* or *F606W*);
(v) 16 = UVIS2 coverage (*F600LP* or *F814W*);
(vi) 32 = IRAC coverage (Ch1 and/or Ch2);
(vii) 64 = *u* coverage;
(viii) 128 = *g* coverage;
(ix) 256 = *r* coverage;
(x) 512 = *i* coverage.

`GRISM_FLAG` contains a sum of two flag bits:

(i) 1 = *G102* coverage;
(ii) 2 = *G141* coverage.

`[GRISM_FILTER]_FLAG` contains a sum of five flag bits:

(i) 1 = Artefact, a satellite trail, strange features, significant persistence, etc.
(ii) 2 = Sky subtraction problem, residual sky, or structure remaining in the background.
(iii) 4 = One or more bright sources present that significantly contaminate the field.
(iv) 8 = A very crowded field, usually indicating some type of star cluster.
(v) 16 = Scattered light, leading to sensitivity depths that vary significantly depending on source position.

`EDGE_FLAG` contains a sum of five flag bits:

(i) 0 = Object is not near an image edge in direct image.
(ii) 1 = Object is within ∼20 pixels of bottom edge of direct image.
(iii) 2 = Object is within ∼20 pixels of top edge of direct image.
(iv) 4 = Object is within ∼20 pixels of left edge of direct image.
(v) 8 = Object is potentially within the region along the right edge in which the position of zeroth orders is unknown. The exact x-position is different for both grisms and the wavelength of the emission line must also be considered. The line-finding algorithm takes these details into account.

`REDSHIFT_FLAG` contains a sum of eight flag bits:

(i) 0 = Redshift agreement within 1σ errors.
(ii) 1 = Redshift disagreement within 1σ errors; adopted redshift taken as the case with three or more lines with S/N > 3.
(iii) 2 = Redshift disagreement within 1σ errors; adopted redshift taken as the case with line identified as Hα. If both reviewers

measured Hα (i.e. they identified two different lines as Hα), the one with the best χ^2 is taken.

(iv) 4 = Redshift disagreement within 1σ errors and neither reviewer identified multiple high S/N lines nor Hα; adopted redshift taken as the case with the best χ^2.

(v) 8 = Redshift based on a single line.

(vi) 16 = Only one reviewer identified object.

In rare cases where both redshift errors are zero, the redshifts are considered to be in agreement if the per cent difference in redshift is <1 per cent.

`FWHM_FLAG` contains a sum of five flag bits:

(i) 0 = FWHM agreement within 1σ errors.

(ii) 1 = Reported FWHM is larger than 2×`A_IMAGE`.

(iii) 2 = One or more measurements had $\sigma_{\rm FWHM} = 0$. This occurs when the FWHM bumps up against either the upper or lower bounding constraints set on the model fit.

(iv) 4 = FWHM disagreement within 1σ errors; reported FWHM is from best χ^2 fit to full spectrum.

(v) 8 = Reported FWHM is based on measurement from only one reviewer. This either indicates there was redshift disagreement, in which case the reported FWHM is from the best χ^2 fit to full spectrum, or only one reviewer accepted the object.

`[LINE]_FLAG` contains a sum of six flag bits:

(i) 0 = Line flux and EW agreement within 1σ errors.

(ii) 1 = EW disagreement within 1σ errors; reported EW is from best χ^2 fit to full spectrum.

(iii) 2 = Flux disagreement within 1σ errors; reported flux is from best χ^2 fit to full spectrum.

(iv) 4 = Only one reviewer measured flux. The other reviewer has line masked or outside coverage. Alternatively, due to redshift disagreement, all line measurements are from one reviewer.

(v) 8 = Flux limits reported for line.

(vi) 16 = No measurement for line; masked or outside of wavelength coverage

`[LINE]_CONTAM` contains a sum of four flag bits:

(i) 0 = Uncontaminated.

(ii) 1 = Reviewer marked general contamination.

(iii) 2 = Reviewer marked continuum contaminated. May affect emission-line fit and therefore flux measurement.

(iv) 4 = Reviewer marked emission line contaminated.

`[LINE]_EDGE_FLAG` contains a sum of four flag bits:

(i) 0 = Emission line is away from grism edges.

(ii) 1 = Emission line is at $\lambda_{\rm obs} \leq 8500$ Å.

(iii) 2 = Emission line is at $\lambda_{\rm obs} \geq 16\,750$ Å.

(iv) 4 = Emission line is in grism overlap region, $10\,900 \leq \lambda_{\rm obs} \leq 11\,500$ Å.

[1]*Research School of Astronomy and Astrophysics, Australian National University, Cotter Road, Weston Creek, ACT 2611, Australia*
[2]*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia*
[3]*Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA*
[4]*Space Telescope Science Institute, 3700 San Martin Dr, Baltimore, MD 21218, USA*
[5]*Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*
[6]*INAF – Istituto di Radioastronomia, Via Gobetti 101, I-40129 Bologna, Italy*
[7]*Italian ALMA Regional Centre, Via Gobetti 101, I-40129 Bologna, Italy*
[8]*Chinese Academy of Sciences South America Center for Astronomy (CASSACA), National Astronomical Observatories of China (NAOC), 20A Datun Road, Beijing, 100012, China*
[9]*Institut d'Astrophysique de Paris, CNRS, Sorbonne Université, 98bis Boulevard Arago, F-75014, Paris, France*
[10]*IPAC, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA*
[11]*Department of Physics and Astronomy University of California, Los Angeles, Los Angeles, CA 90095-1547, USA*
[12]*National Optical-Infrared Astronomy Research Laboratory, Tucson, AZ 85719, USA*
[13]*Minnesota Institute for Astrophysics, University of Minnesota, Minneapolis, MN 55455, USA*
[14]*Department of Physics and Astronomy, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA*
[15]*School of Physics, University of Melbourne, Parkville 3010, VIC, Australia*
[16]*Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX13RH, UK*
[17]*Astrophysics Science Division, NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA*
[18]*Minnesota State University-Mankato, Telescope Science Institute, TN141, Mankato, MN 56001, USA*
[19]*School of Astronomy and Space Science, University of Chinese Academy of Sciences (UCAS), Beijing 100049, China*
[20]*National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China*
[21]*Institute for Frontiers in Astronomy and Astrophysics, Beijing Normal University, Beijing 102206, China*

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