The 3D-HST Survey: *Hubble Space Telescope* WFC3/G141 Grism Spectra, Redshifts, and Emission Line Measurements for 100,000 Galaxies

Journal Item

The Astrophysical Journal Supplement Series, 225:27 (35pp), 2016 August doi:10.3847/0067-0049/225/2/27

THE 3D-HST SURVEY: *HUBBLE SPACE TELESCOPE* WFC3/G141 GRISM SPECTRA, REDSHIFTS, AND EMISSION LINE MEASUREMENTS FOR ~100,000 GALAXIES

Ivelina G. Momcheva[1], Gabriel B. Brammer[2], Pieter G. van Dokkum[1], and Rosalind E. Skelton[1,3], Katherine E. Whitaker[4,17], Erica J. Nelson[1], Mattia Fumagalli[5], Michael V. Maseda[6], Joel Leja[1], Marijn Franx[5], Hans-Walter Rix[6], Rachel Bezanson[7,17], Elisabete Da Cunha[8], Claire Dickey[1], Natascha M. Förster Schreiber[9], Garth Illingworth[10], Mariska Kriek[11], Ivo Labbé[5], Johannes Ulf Lange[1], Britt F. Lundgren[12], Daniel Magee[10], Danilo Marchesini[13], Pascal Oesch[1], Camilla Pacifici[2], Shannon G. Patel[14], Sedona Price[11], Tomer Tal[10], David A. Wake[12,15], Arjen van der Wel[6], and Stijn Wuyts[16]
[1] Department of Astronomy, Yale University, 260 Whitney Avenue, New Haven, CT 06511, USA; ivelina.momcheva@yale.edu
[2] Space Telescope Science Institute, Baltimore, MD 21218, USA
[3] South African Astronomical Observatory, Cape Town, 7935, South Africa
[4] Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA
[5] Leiden Observatory, Leiden University, Leiden, The Netherlands
[6] Max Planck Institute for Astronomy, D-69117, Heidelberg, Germany
[7] Steward Observatory, University of Arizona, Tucson, AZ 85721, USA
[8] Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia
[9] Max-Planck-Institut für extraterrestrische Physik, Garching, Germany
[10] Department of Astronomy & Astrophysics, University of California, Santa Cruz, CA, USA
[11] Astronomy Department, University of California, Berkeley, CA 94720, USA
[12] Department of Astronomy, University of Wisconsin-Madison, Madison, WI 53706, USA
[13] Department of Physics and Astronomy, Tufts University, Medford, MA 02155, USA
[14] Carnegie Observatories, Pasadena, CA 91101, USA
[15] Department of Physical Sciences, The Open University, Milton Keynes, MK7 6AA, UK
[16] Department of Physics, University of Bath, Claverton Down, Bath BA2 7AY, UK
Received 2015 October 6; revised 2016 May 24; accepted 2016 May 24; published 2016 August 11

ABSTRACT

We present reduced data and data products from the 3D-HST survey, a 248-orbit *HST* Treasury program. The survey obtained WFC3 G141 grism spectroscopy in four of the five CANDELS fields: AEGIS, COSMOS, GOODS-S, and UDS, along with WFC3 H_{140} imaging, parallel ACS G800L spectroscopy, and parallel I_{814} imaging. In a previous paper, we presented photometric catalogs in these four fields and in GOODS-N, the fifth CANDELS field. Here we describe and present the WFC3 G141 spectroscopic data, again augmented with data from GO-1600 in GOODS-N (PI: B. Weiner). We developed software to automatically and optimally extract interlaced two-dimensional (2D) and one-dimensional (1D) spectra for all objects in the Skelton et al. (2014) photometric catalogs. The 2D spectra and the multi-band photometry were fit simultaneously to determine redshifts and emission line strengths, taking the morphology of the galaxies explicitly into account. The resulting catalog has redshifts and line strengths (where available) for 22,548 unique objects down to $JH_{\rm IR} \leqslant 24$ (79,609 unique objects down to $JH_{\rm IR} \leqslant 26$). Of these, 5459 galaxies are at $z > 1.5$ and 9621 are at $0.7 < z < 1.5$, where Hα falls in the G141 wavelength coverage. The typical redshift error for $JH_{\rm IR} \leqslant 24$ galaxies is $\sigma_z \approx 0.003 \times (1+z)$, i.e., one native WFC3 pixel. The 3σ limit for emission line fluxes of point sources is 2.1×10^{-17} erg s^{-1} cm^{-2}. All 2D and 1D spectra, as well as redshifts, line fluxes, and other derived parameters, are publicly available.[18]

Key words: catalogs – galaxies: evolution – methods: data analysis – techniques: spectroscopic

1. INTRODUCTION

Since its deployment in 1990, the *Hubble Space Telescope* (*HST*) has not only been used as an imager but also as a spectrograph. Space-based spectroscopy offers the same advantages as space-based imaging: spatial resolution that is difficult or impossible to achieve from Earth and access to wavelength regimes that are blocked by the atmosphere. Whereas dedicated *HST* spectrographs such as STIS and COS use a slit or a small aperture to isolate the light of an individual object, several of the imaging instruments on *HST* employ, or employed, a different technique. The NICMOS, ACS, and WFC3 cameras were all equipped with dispersing grisms that can be placed in the light path in lieu of a filter. This technique is very efficient because it provides spectra of all objects in the imaging field simultaneously (Pirzkal et al. 2004; Malhotra et al. 2005; Straughn et al. 2008; van Dokkum & Brammer 2010). Slitless spectroscopy has limited appeal in ground-based astronomy, as the brightness of the sky greatly reduces the signal-to-noise ratio (S/N) compared to slit spectroscopy. However, the much fainter background from space makes slitless *HST* spectroscopy competitive with, and in several respects superior to, ground-based slit spectroscopy.

While the NICMOS grisms have left little mark on the field of galaxy formation, the ACS grisms were successfully used to obtain deep optical spectroscopy in several fields (e.g., the PEARS and GRAPES surveys; Pirzkal et al. 2004; Malhotra et al. 2005; Straughn et al. 2008). Among other successes, the ACS G800L data of GRAPES led to the spectroscopic identification of passively evolving galaxies at $z \sim 2$ in the

[17] Hubble Fellow.
[18] http://3dhst.research.yale.edu

Hubble Ultra Deep Field (Daddi et al. 2005). This discovery was aided by a particular aspect of *HST* grism observations. As galaxies are spatially resolved at *HST* resolution, their spectra are spread over multiple pixels. Therefore, the S/N is strongly dependent on morphology, and is higher for galaxies that are more compact. As it turns out, the passively evolving galaxies in the Ultra Deep Field have extremely small sizes, yielding relatively high S/N spectra. The high spatial resolution of *HST* also enables the study of the spatial variation in spectral features; as discussed in detail in Nelson et al. (2015), this opens up the possibility to study the spatial distribution of line emission at scales of $\sim 0\farcs1$.

The grism mode of the WFC3 camera's near-IR channel is realizing the full potential of space-based slitless spectroscopy. Although the sky background from space is lower than that from Earth at all wavelengths, the differences are more pronounced in the near-IR than in the optical: *HST*'s near-IR background is similar to that of a 30 m class telescope on Earth. As a result, the per-object sensitivity of WFC3 grism spectroscopy *without* slits is similar to that of ground-based spectrographs on 10 m telescopes *with* slits (as will be quantified later in this paper). With the added benefits of superb spatial resolution and highly efficient multi-plexing, the WFC3 camera is an excellent spectroscopic survey instrument at near-IR wavelengths. It is complementary to ground-based multi-object spectrographs such as MOSFIRE (McLean et al. 2012) and KMOS (Sharples et al. 2013, 2014): these ground-based spectrographs have much higher spectral resolution ($R \sim 3500$ for MOSFIRE versus $R \sim 100$ for the WFC3/G141 grism) but cannot match the continuum sensitivity or observing efficiency of WFC3.

The 3D-HST Treasury program (van Dokkum et al. 2011; Brammer et al. 2012b; Skelton et al. 2014) has obtained 2-orbit depth WFC3/G141 grism observations over four large sky areas, comprising a total of 124 pointings. The G141 grism has a wavelength coverage of 1.1–1.7 μm, approximately corresponding to ground-based J and H (including the region in between these bands, which is inaccessible from the ground due to H_2O absorption). The main aim of the survey is to obtain a large, representative spectroscopic sample of galaxies at $0.7 < z < 3$, the epoch when most of the stars in the present-day universe were formed. As we show below, a typical single 2-orbit WFC3/G141 grism observation provides redshifts of $\sim$130 galaxies at $0.7 < z < 3$ down to $H_{160} \leqslant 24$ in a 4.6 arcmin^2 area. The survey also obtained ACS/G800L grism observations in parallel, covering 0.5–0.9 μm, as well as short direct imaging exposures in the WFC3 JH_{140} and ACS I_{814} bands.

The survey fields were chosen to coincide with those of the CANDELS Multi-Cycle Treasury project (Grogin et al. 2011; Koekemoer et al. 2011), which has obtained WFC3 and ACS imaging of five fields, comprising a total area of $\sim$0.25 degree^2 (see Table 1). These fields have a wealth of complementary imaging data at other wavelengths from ground- and space-based observatories (see Grogin et al. 2011; Brammer et al. 2012b; Skelton et al. 2014, and references therein), and have quickly become the "standard" deep, moderately wide areas for studies of the distant universe. The four fields observed by the 3D-HST Treasury program are AEGIS, COSMOS, GOODS-S, and UDS. The GOODS-N field had already been observed in a Cycle 17 program (GO-11600; PI Weiner), using a very similar observing strategy. We have included the GOODS-N data in our analysis and data release, and throughout this paper we discuss the combined grism data set for all five fields. The footprint of 3D-HST[19] is slightly smaller than that of CANDELS; approximately 70% of the CANDELS WFC3 area is covered by grism spectroscopy from our program or the Weiner program.

Table 1
3D-HST Fields

Field	R.A. (h m s)	Decl. (d m s)	G141 Area (arcmin^2)	G800L Area (arcmin^2)
AEGIS	14:18:36.00	+52:39:00.0	121.9	102.5
COSMOS	10:00:31.00	+02:24:00.0	122.2	112.7
GOODS-N	12:35:54.98	+62:11:51.3	116.0	84.1
GOODS-S	03:32:30.00	−27:47:19.0	147.3	134.6
UDS	02:17:49.00	−05:12:02.0	118.7	107.4
Total			626.1	541.3

In many cases, the grism spectra can stand on their own, particularly for galaxies that have bright (redshifted) emission lines that fall between 1.1 and 1.7 μm (see, e.g., Atek et al. 2010; Straughn et al. 2011). However, the value of the grism spectra can be enhanced by combining them with broad- and medium-band photometry at other wavelengths, which is possible in the CANDELS fields (see Skelton et al. 2014). We have developed an integrated approach, where the ground- and space-based imaging data are optimally combined with the G141 grism spectroscopy. The combined grism and photometric data set was used to derive redshifts, measure emission lines, and determine other parameters of *all* galaxies in a photometric catalog (not just those with bright lines), down to well-defined magnitude limits. These steps can be summarized as follows.

1. We obtained and reduced the available *HST*/WFC3 images in the fields and reduced them using the same pixel scale and tangent point used by the CANDELS team.
2. We used SExtractor (Bertin & Arnouts 1996) to detect objects in deep combined $J_{125} + JH_{140} + H_{160}$ images.
3. We used source catalogs (along with the detection images, associated segmentation maps and point-spread functions (PSFs)) to measure photometric fluxes at wavelengths of 0.3–8 μm from a large array of publicly available imaging data sets. As a result, we derived high-quality spectral energy distributions (SEDs), particularly in fields with extensive optical and near-IR medium-band photometry.
4. The catalogs and segmentation maps were blotted to the original (interlaced) coordinate system of the WFC3 and ACS grism data, and spectra were extracted for each object covered by the grism. No source matching was required, and the photometric SEDs could be combined directly with the grism spectroscopy.
5. The interlaced 2D spectra and SEDs were fitted simultaneously to measure redshifts, allowing a limited range of emission line ratios.

[19] Throughout the text, the terms "3D-HST" and "the 3D-HST data" refer to the combination of the CANDELS imaging and all the other ancillary space- and groud-based imaging data sets as presented in Skelton et al. (2014) and the grism spectroscopy of the 3D-HST Treasury program, GO-11600, and other data sets (Table 2), in all five CANDELS fields, except when explicitly specified otherwise.

6. With the redshifts determined, emission line fluxes were measured from the 2D spectra with no prior on line ratios.
7. Stellar population parameters were determined by fitting stellar population synthesis models to the SEDs, using the redshifts as input.
8. Mid-IR photometry was obtained from Spitzer/MIPS imaging. These data, combined with rest-frame UV emission measurements from the SEDs, were used to determine star formation rates (SFR) of the galaxies.
9. The set of images, PSFs, and catalogs was used to measure structural parameters of the objects in the WFC3 and ACS bands, following the methodology of van der Wel et al. (2012).

Steps 1–3 have been described in Skelton et al. (2014). Step 8 is discussed in Whitaker et al. (2012) and, specifically for the 3D-HST survey in Whitaker et al. (2014). Step 9 is described in van der Wel et al. (2014), who measured the structural parameters of objects in the J_{125} and H_{160} WFC3 bands. All these data sets have been made publicly available. In this paper, we discuss steps 4, 5, 6, and 7, and describe the full 3D-HST spectroscopic data release. These steps are inextricably linked to the previously published data sets. The detection and segmentation maps of Skelton et al. (2014) are used as inputs for the grism reduction and their WFC3 images are used to create the grism image model.

The utility of the data products that are described in this paper has already been demonstrated in a large number of studies, and we highlight several examples that illustrate particular aspects of the grism data. Nelson et al. (2012), Schmidt et al. (2013), and Wuyts et al. (2013) analyze Hα emission line maps of galaxies at $z \sim 1$, which are very difficult to obtain by any other means. Furthermore, Nelson et al. (2013, 2015) show that emission lines in 3D-HST can be traced to large radii by stacking thousands of spectra and find that the Hα emission is more extended than the stellar continuum. This suggests that galaxies grow inside-out. Fumagalli et al. (2012) study the evolution of the Hα equivalent width, and find that it increases rapidly with redshift. Price et al. (2014) show that the Balmer decrement of galaxies at $z \sim 1.5$ increases with stellar mass and derive expressions for the relation between continuum extinction and the extinction toward H II regions. Brammer et al. (2013) use the deepest G141 that were in existence at the time to constrain the spectrum of a $z \sim 12$ galaxy candidate in the Ultra Deep Field.[20] Despite the relatively shallow depth of our survey we also obtain information on absorption lines of galaxies out to fairly high redshift. This is demonstrated in van Dokkum & Brammer (2010) and particularly in Whitaker et al. (2013), who spectroscopically confirm the existence of a large population of galaxies with old stellar populations at $z \sim 2$.

Lastly, we list several 3D-HST results that do not use particular spectral features, but utilize the large, homogeneous data set of galaxies with reliable redshifts that the survey provides. van Dokkum et al. (2013b) and Patel et al. (2013) describe the evolution of Milky-Way-like galaxies from $z \sim 2.5$ to the present, using number density-matched samples. van der Wel et al. (2014) combine the 3D-HST catalogs with CANDELS photometry to study the evolution of the mass–size relation with redshift. Whitaker et al. (2014) provide a new measurement of the relation between star formation and stellar mass (the "star formation main sequence"), and find that there is a turnover in the relation at low masses. Nelson et al. (2014) and van Dokkum et al. (2014, 2015) study the formation and evolution of the cores of massive galaxies. A full list of 3D-HST papers can be found on the 3D-HST website.[21]

The structure of the paper is as follows. In Section 2, we describe the data that are now part of the 3D-HST project. Section 3 details the data reduction, including the interlacing procedure that we use instead of drizzling. The extraction of the two-dimensional (2D) and one-dimensional (1D) spectra is discussed in Section 4. The redshift fits are described in Section 5, along with a review of their accuracy. We fit the spectra twice, once in conjunction with the photometry to determine redshifts, and then a second time to measure emission line fluxes and equivalent widths. Line flux fits are described in Section 6. The catalog entries are explained in Section 7. In Section 8, we highlight the properties of the spectroscopic sample. The paper is summarized in Section 9. Magnitudes throughout are on the AB system.

2. DATA

Most of the data described in this paper were obtained by the 3D-HST Treasury Survey, which was allocated 248 orbits of *HST* time during Cycles 18 and 19. We obtained 2-orbit depth observations using the ACS/G800L and WFC3/G141 grisms in parallel. These observations cover 124 pointings in four of the five deep fields observed by CANDELS (AEGIS, COSMOS, GOODS-S, and UDS) and constitute the largest effort to acquire space-based near-infrared spectra in these fields. A detailed description of the 3D-HST observations is presented in Brammer et al. (2012b). The fifth CANDELS field, GOODS-N, had already been observed with WFC3/G141, in a Cycle 17 program as part of AGHAST (A Grism Hα SpecTroscopic Survey; GO-11600, PI: Weiner), using a very similar observing strategy. ACS/G800L observations in GOODS-N were taken as part of GO-13420 in Cycle 21 (PI: Barro). We have included the ACS/G800L and WFC3/G141 GOODS-N data in our sample, and throughout this paper we discuss the combined grism data set for all five fields. Table 1 lists the coordinates of each field and total areas covered with each instrument.

A comprehensive list of all ACS and WFC3 grism observations in the five 3D-HST/CANDELS fields taken in Cycles 17 through 21 is presented in Table 2. In addition to the ACS/G800L and WFC3/G141 data, we also summarize the available archival WFC3/G102 observations for each field. While the 3D-HST WFC3/G141 observations cover a relatively wide area to a shallow, 2-orbit, depth, other programs have obtained deep observations—up to a depth of 15 orbits (PRIMO, GO-12099, PI Riess, Rodney et al. 2012)—over a single pointing. The major source of these deep grism observations are the CANDELS supernovae follow-up programs (GO: 12099 and 12461; PI: Riess) in addition to data from GO-12190 (CDFS-AGN; PI: Koekemoer), GO-11367, (PI: O'Connell) and GO-12547 (EGS, PI: Cooper). We made a subset of these deep data (in the Hubble Ultra Deep Field) publicly available in a previous data release,[22] but they are not

[20] These deep spectra of galaxies in the Hubble Ultra Deep Field were released in 2013 (van Dokkum et al. 2013a) and are available from the 3D-HST website.

[21] http://3dhst.research.yale.edu/Publications.html

[22] See van Dokkum et al. (2013a).

Table 2
WFC3 and ACS Grism Observations in the 3D-HST/CANDELS Fields in Cycles 17 to 21

Field	Instrument	Number of Orbits			Proposal ID	*HST* Cycle	Survey/Pointing	PI
		G800L	G102	G141				
AEGIS	WFC3	⋯	⋯	2	13063	20	SN EGSA	Riess
	WFC3	⋯	⋯	6[a]	12547	19		Cooper
	ACS, WFC3	62	⋯	62	12177	18	3D-HST	van Dokkum
COSMOS	WFC3	⋯	12	12	12461	19	SN TILE 41	Riess
	ACS, WFC3	56	⋯	56	12328	18	3D-HST	van Dokkum
GOODS-N	WFC3	56	56	⋯	13420	21		Barro
	WFC3	⋯	⋯	4	12461	19	SN COLFAX	Riess
	WFC3	⋯	⋯	56	11600	17	AGHAST	Weiner
GOODS-S	ACS, WFC3	76	⋯	76	12177	18	3D-HST	van Dokkum
	WFC3	⋯	12	12	12190	18	CDFS-AGN 1, 2	Koekemoer
	ACS, WFC3	6	⋯	6 + 15	12099	18	GEORGE, PRIMO	Riess
	WFC3	⋯	2	2	11359	17	ERS	O'Connell
UDS	WFC3	⋯	10	⋯	12590	19	IRC0218A	Papovich
	ACS,WFC3	56	⋯	56	12328	18	3D-HST	van Dokkum
	WFC3	⋯	⋯	18	12099	18	MARSHALL	Riess
Total		312	92	383				

Note.
[a] The full program is 24 orbits in 12 pointings; however, only 6 orbits overlap with the 3D-HST/CANDELS footprint.

part of the data set described in this paper: here we concentrate on the shallow, 2-orbit depth wide-field data.

2.1. WFC3 Observations

The WFC3 G141 grism has spectral coverage from 1.1 to 1.65 μm (at >30% throughput) and a peak transmission of 48% at 1.45 μm. The G141 dispersion is 46.5 Å pixel^{-1} ($R \sim 130$) in the primary (+1st) spectral order. However, in practice, the spectral resolution for each (resolved) object is different because it is largely determined by its morphology. The wavelength zero point and the dispersion uncertainties are 8 Å and 0.06 Å pixel^{-1}, respectively (Kuntschner et al. 2010). The field of view of the WFC3 IR channel is $136'' \times 123''$.

The layout of the WFC3/G141 observations in the CANDELS fields is shown in Figure 1, overlaid on the H_{160} imaging footprint. Across the five fields, 70% of the CANDELS area is covered with at least two orbits of WFC3/G141 data. In AEGIS, COSMOS, and UDS, 60% of the CANDELS imaging area has complementary G141 grism data, while 70% of GOODS-N and 86% of GOODS-S have G141 coverage. The total area of the G141 observations is 626 arcmin^2 (Table 1).

The observations for the 3D-HST survey started on 2010 October 30 and ended on 2012 March 22. Two pointings in AEGIS, 1 and 22, were re-observed on 2013 April 21 and 2012 November 30, respectively. Each pointing of the 124 3D-HST pointings was observed for two orbits, with four paired JH_{140} direct and G141 grism exposures. Typical total exposure times in each pointing are 800 s in JH_{140} and 5000 s in G141. The four pairs of direct+grism exposures are separated by small telescope offsets to improve the sampling of the PSF, to enable the identification of hot pixels and other defects not flagged by the default pipeline processing, and to dither over some WFC3 cosmetic defects such as the "IR blobs" (Pirzkal et al. 2010). The sub-pixel dither pattern used throughout the survey is shown in Figure 3 of Brammer et al. (2012b).

The 56 orbits from the AGHAST program in GOODS-N are divided into 28 pointings, each with two-orbit depths. The observations were carried out between 2009 September 16 and 2010 September 26. Due to high background and scattered light artifacts, nine of the AGHAST pointings were partially re-observed between 2011 April 19 and 24. Analogous to 3D-HST, each two-orbit observation was split into four sets of G141 grism images and JH_{140} direct exposures. The dither patterns of AGHAST and 3D-HST are slightly different, but they both sample the WFC3 PSF on a grid that is 0.5 × 0.5 their native pixel size. The typical exposure time per pointing is 800 s in JH_{140} and ~5200 s in G141. Further information about AGHAST can be found on the survey website.[23]

2.2. ACS Observations

Exposures with the ACS G800L grism, accompanied by I_{814} direct imaging, were taken in parallel with the primary WFC3 exposures. ACS coverage of the GOODS-N fields was done in program GO-13420 (PI: Barro) as parallels to their WFC3/G102 primary observations. The G800L grism has a wavelength coverage from 0.55 to 1.0 μm with a dispersion of 40 Å pixel^{-1} in the primary first order. The total exposure times in each pointing/visit are 480 s in I_{814} (1299 s in GO-13420) and between ~2800 s (GOODS-N) and ~3500 s (AEGIS) in G800L. Figure 2 shows the layout of the pointings in all five fields. Unlike the WFC3 pointings, the ACS pointings do not have a regular pattern, but an effort was made to maximize the overlap between the two grisms. Fully 86.5% of the WFC3 grism observations also have ACS grism coverage. Within each pointing, four pairs of I_{814} direct images and G800L grism images were taken in a sequence. As a result of the larger ACS

[23] http://mingus.as.arizona.edu/~bjw/aghast/

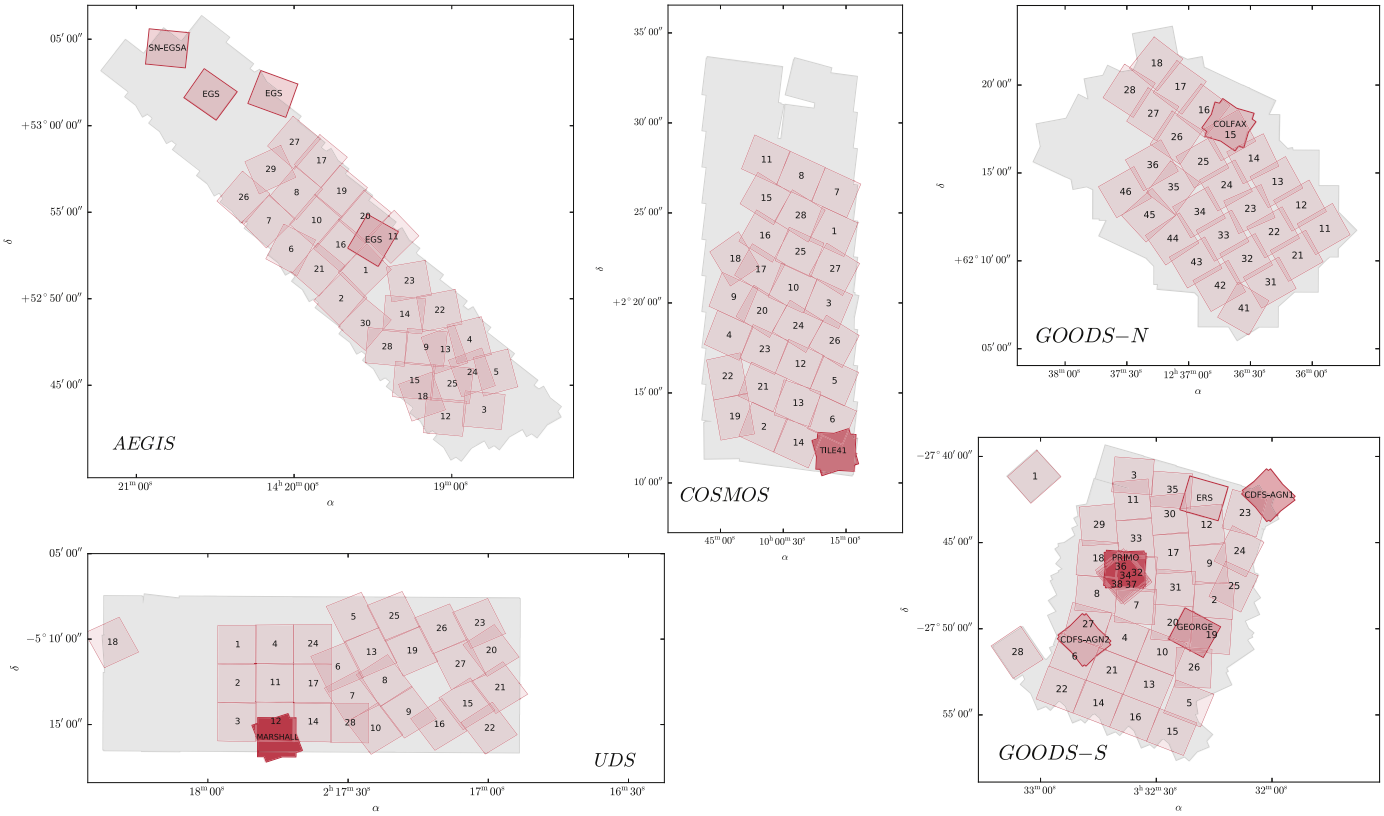

Figure 1. Layout of the WFC3 G141 observations. The primary WFC3 G141 pointings are shown with red outlines with the pointing ID numbers as defined in the Phase II file. Observations in the GOODS-N field are from the AGHAST survey (PI: Weiner). Additional pointings are marked with the program or pointing names. The color is proportional to the grism depth, ranging from ∼5 ks for 3D-HST to 60 ks in the GOODS-S HUDF/PRIMO area. See Table 2 for details.

field of view ($202'' \times 202''$), there is larger overlap between the pointings, with some areas covered up to a depth of eight orbits.

Ten ACS pointings (listed in Table 3) fall completely outside of the footprint of the CANDELS/3D-HST WFC3 imaging. As described below, the WFC3 mosaics are used as world coordinate system (WCS) reference images for aligning the direct I_{814} images. Pointings that fall outside of these mosaics cannot be aligned to the same WCS and cannot be processed in the same manner as the rest of the pointings. Thus, they are processed throughout the preliminary reduction steps only.

3. DATA REDUCTION

An early version of our reduction pipeline was described in Brammer et al. (2012b). At that time, the pipeline used custom pre-processing steps such as alignment, flat-fielding, and sky-subtraction followed by extraction of the grism spectra using the `aXe` software package (Kümmel et al. 2009). We have made major changes to the reduction procedures since then, and here we describe our final approach to the data processing.

Some changes have been made to the pre-processing steps, as discussed below, but the main difference between our current pipeline and that described by Brammer et al. (2012b) takes place after the pre-processing. In particular, we no longer use the `aXe` package. `aXe` drizzles the data onto a grid with linear sampling in the wavelength and spatial direction prior to extraction of 2D and 1D spectra. Drizzling introduces correlations between pixels and smooths the data, and in an effort to optimally use the information in the grism spectra we have developed an approach that uses the original WFC3/IR pixels without resampling. We do not drizzle the data, but instead place the original pixels of four dithered exposures on a (distorted) output grid for which the pixels are exactly half the native pixel size. This interlacing approach, discussed in detail in Section 3.5, retains the independence of adjacent pixels and the full resolution of the data. The distortions are encoded in the software that is used to analyze the spectra, and in a pixel-to-wavelength conversion table that is unique to each object and supplied with our data release.

3.1. WFC3 JH_{140} Images

We downloaded the raw (RAW) images, the calibrated (flat-fielded or FLT) images, and the association tables (ASN) for all observations from the Mikulski Archive for Space Telescopes (MAST[24]). The calibrated images are processed with the `calwf3` pipeline (described in detail by Koekemoer et al. 2011). We also obtain the persistence images (PERSIST)[25], which provide estimates for the total IR persistence that affects a given exposure, both from sources internal to the 3D-HST visit and also external sources from prior observations.

The reduction of the direct images is described in detail in Skelton et al. (2014). Here we give a brief summary of the relevant steps. The main difference in the image preparation steps relative to Skelton et al. (2014) is the full integration of `TweakReg` and `AstroDrizzle` (Gonzaga et al. 2012) in the reduction. Previously, it was used only for the final alignment and drizzling steps. We do not apply the advanced processing

[24] http://archive.stsci.edu.

[25] Long et al. (2013); http://archive.stsci.edu/prepds/persist/.

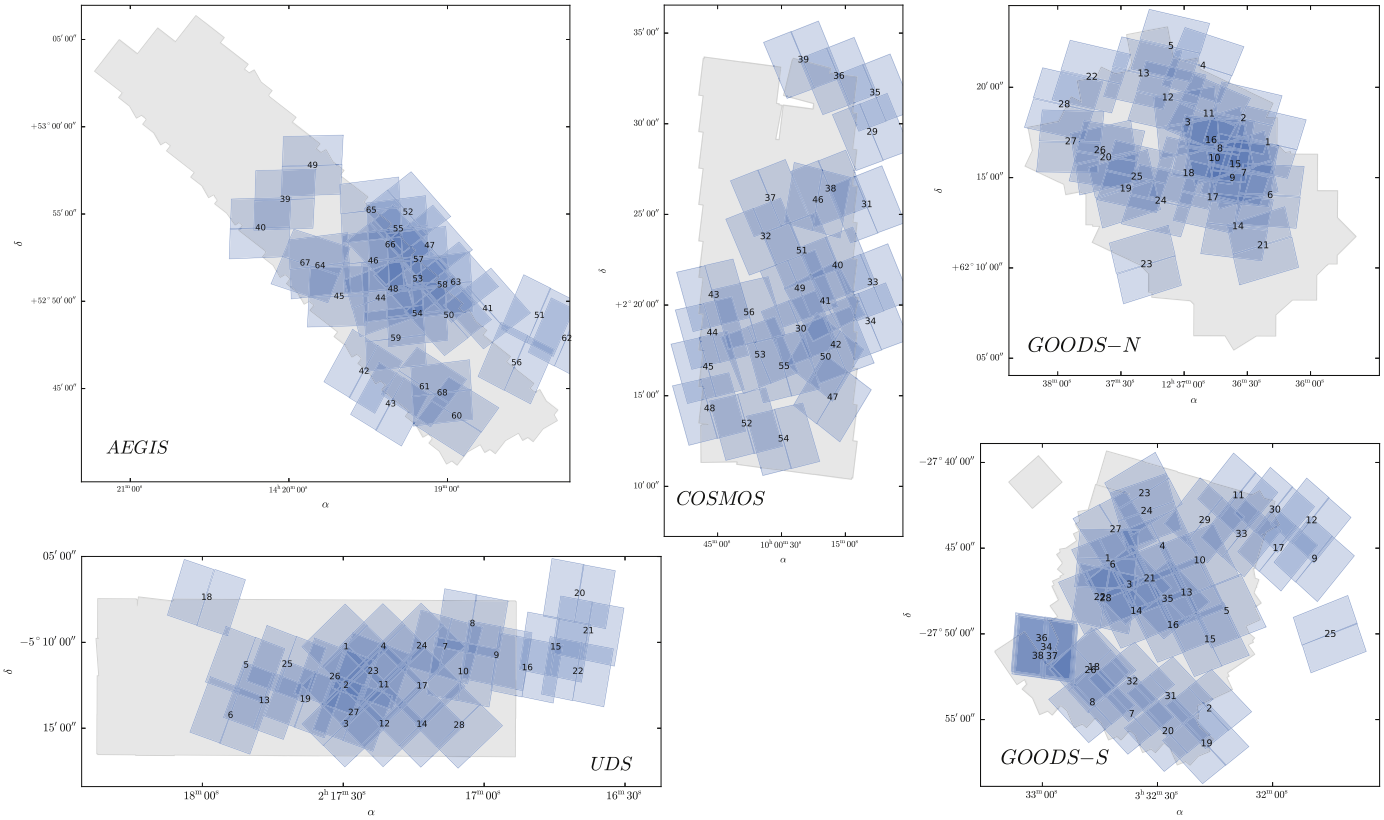

Figure 2. Layout of the ACS G800L observations. The observations in GOODS-N are from GO-13420 (PI: Barro). The AEGIS, COSMOS, and GOODS-N pointings are numbered differently from their WFC3/IR parallels. The pointing numbers shown in this figure are also used in the data release.

Table 3
ACS Pointings Outside of the CANDELS/3D-HST Footprint

Field	R.A.	Decl.	ACS Pointing ID	WFC3 Primary
AEGIS	14:18:46.129	+52:49:27.29	41	3
	14:18:26.632	+52:49:17.39	51	13
	14:18:16.253	+52:47:58.13	62	24
GOODS-S	03:31:49.881	−27:45:28.50	9	9
	03:31:50.754	−27:43:14.19	12	12
	03:31:45.353	−27:49:48.84	25	25
UDS	02:16:45.756	−05:10:12.60	15	15
	02:16:40.606	−05:07:05.53	20	20
	02:16:38.757	−05:09:16.03	21	21
	02:16:40.999	−05:11:38.08	22	22

of individual reads that is described in Section 3.2.1; the direct images are comprised of only four samples and eliminating one or more of these would lead to a significant loss in the integration time.

All FLT images were first inspected for satellite trails and artifacts, as well as for regions of elevated background due to earthshine. Affected regions identified in the inspection were masked and given a data quality flag of 2048 (the cosmic-ray data quality flag, Chapter 2.2.3, WFC3 Data Handbook, Rajan et al. 2010) so that they are treated as pixels without information in the subsequent processing steps. For persistence masking, we apply a conservative threshold, requiring that the predicted persistence is less than 0.6 times the values in the FLT error extension. We grow the persistence-masked area slightly and then set the 4196 bit in the data quality extension for the masked pixels. These are later treated as cosmic rays and are not used in the final mosaics. Finally, we add a component to the FLT uncertainties to account for cross-talk from pixels where the total number of deposited electrons is greater than 2×10^4 e^-. Time-dependent sky flats were created from the science exposures, which account for the appearance of new IR "blobs" with time since the installation of WFC3.

We run `AstroDrizzle` first to identify hot pixels and cosmic rays not flagged by the `calwf3` calibration pipeline. This step produces a preliminary combined JH_{140} science image of each pointing. We subtract the background from this image in the following way. A preliminary source detection is done with SExtractor (Bertin & Arnouts 1996). The detected sources are used to create a mask, and we fit a second order polynomial background and subtract it from the FLT exposure. Using `TweakReg`, we align each FLT image to the reference frame of the Skelton et al. (2014) mosaics by providing a reference list of object positions derived from the Skelton et al. (2014) photometric catalogs. These alignment corrections beyond the commanded dither positions are typically small, of the order of 0.1 pixels.

Figure 3 illustrates the differences between the default FLT image and the final processed FLT. The most notable difference between the science images (top) is the background subtraction, which removes the pedestal of $\sim 2e^-$ s^{-1} in the default image. In the data quality arrays (bottom), the persistence, caused by the spectra of the two bright stars in the frame, has been masked.

3.2. WFC3 G141 Images

Following the `calwf3` pre-processing, we apply several steps to improve the grism data quality. These steps are

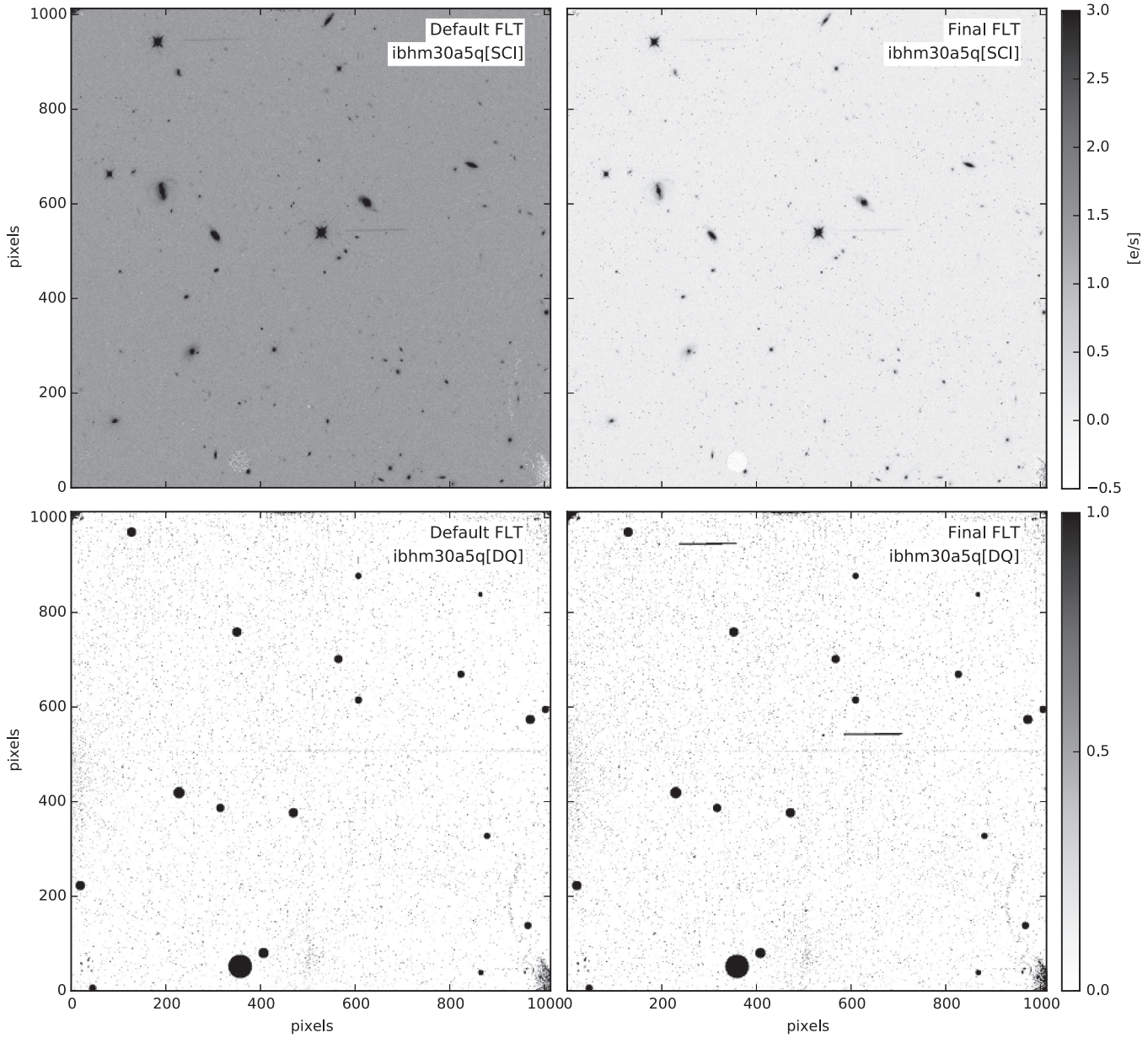

Figure 3. Original default FLT (left) and final processed FLT (right). For each FLT, we show both the science image (top) and the data quality array (bottom). The main difference between the two science images is the background subtraction. The main difference between the data quality arrays is the persistence masking.

removing satellite trails, persistence masking, flat-fielding, sky-subtraction, astrometric alignment, and final cosmic-ray and bad pixel rejection. Some of these steps were also described in Brammer et al. (2012b); these are briefly summarized with an emphasis on any differences that we implemented since that paper. Figure 4 demonstrates various stages of our processing of one of the FLT images (of pointing AEGIS-01); details are provided below.

3.2.1. Removing Satellite Trails and Earthshine

The grism images occasionally contain satellite trails and other cosmetic blemishes, which we identify by visual inspection of all grism exposures in a manner similar to the direct images. A single WFC3 exposure is comprised of multiple samples, which are generated by multiple non-destructive MULTIACCUM reads during the exposure. Therefore, a single WFC3 image is really a sum of independent images, which can be recovered by analyzing the individual reads.[26] While the direct images typically only have four non-destructive reads, the grism images have 12–15 100 s reads.

Satellites move across the WFC3 field-of-view quickly and typically only affect a single read. Rather than masking areas of the detector, we remove the read (or, sometimes, the reads) that is affected by extraneous light. To remove the affected reads, we use the IMA files (intermediate MULTIACCUM files) produced by `calwf3`, which contain the individual calibrated reads from the exposure. We average the count rates in all of the clean reads in the calibrated IMA files and use this averaged image in place of the FLT. As this process bypasses the

[26] We use the terms "sample" and "read" interchangably, although it is more correct to use the term "read" for the process that, through differencing, produces a sample.

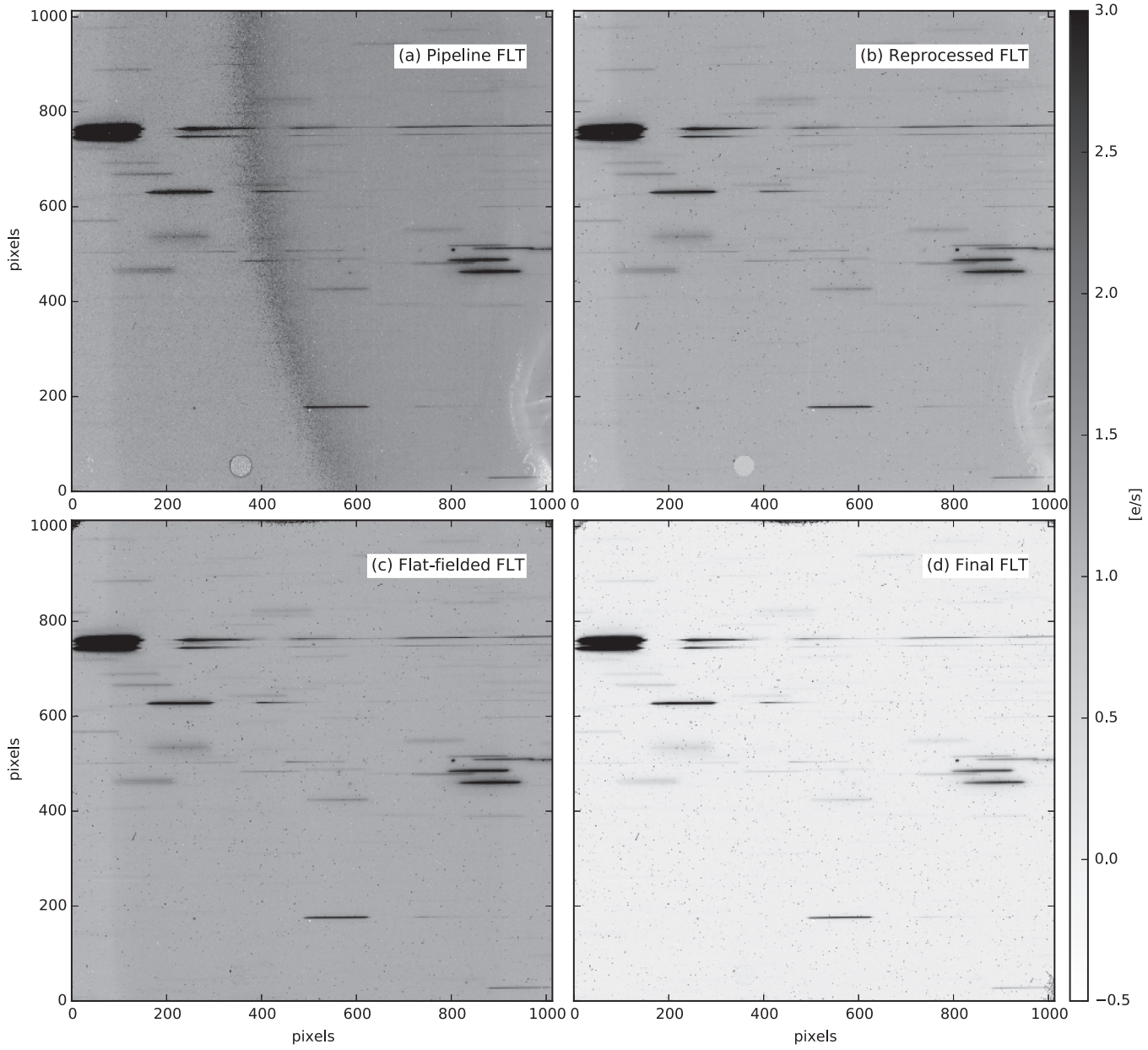

Figure 4. Steps in the reduction of a G141 image. We use as an example one of the FLT images of pointing AEGIS-01: ibhj39uuq_flt.fits (also shown in Figure 29). Shown are (a) the default `calwf3` pipeline-processed image, which, in this case, has a high earthshine background component; (b) the reprocessed `calwf3` image with the last four reads removed; (c) the flat-fielded reprocessed frame; and (d) the background-subtracted final image. The `calwf3` reprocessing is done only for a small subset of all FLT images; most of the pipeline-processed FLT files resemble panel (b), not panel (a).

`calwf3` up-the-ramp cosmic-ray rejection step, the cosmic rays in these (few) reprocessed exposures must be identified separately based on comparison with the other dithered exposures (with `AstroDrizzle`). The final exposure time of the reprocessed exposures are reduced by the duration of the rejected reads (typically by 100 s).

A number of pointings, specifically in GOODS-N, are affected by scattered Earth light or earthshine. This light is observed when the telescope points near the bright Earth limb and its light reaches the detector through an unintended path in the optics. As a result, the background level in the leftmost ∼200 columns of the detector can be increased to levels of up to twice that of the rest of the detector (see Figure 6.17 of the WFC3 Data Handbook Rajan et al. 2010). As with the satellite trails, we visually examine the individual reads and remove those affected by bright earthshine from the sequence (reads with low-level earthshine are not removed because they can be corrected in the background subtraction step, discussed below). Removing reads in this manner decreases the effective exposure time of the observation by ∼100 s per read removed.

Further information about the removed reads is provided in the Appendix. In total, 30 pointings had one read removed and one pointing had two reads removed due to satellites crossing the WFC3 field of view during the exposure. The effect of these readout removals is minimal because satellites affect one of the four FLTs in a pointing and the loss of a single read only constitutes a loss of ∼2% of the total exposure. Earthshine, on the other hand, can have a significant effect on the depth of a pointing because it typically appears at the beginning or at the end of the exposure and lasts for multiple reads. Twenty-three

FLT images in 12 pointings in GOODS-N as well as two of the AEGIS pointings are affected by earthshine. Panels a and b of Figure 4 show the removal of earthshine in pointing AEGIS-01, where 4 of the 12 readouts are removed. Across all FLTs, between one and eight reads have been removed in each affected image, which results in significant loss of depth in some of these pointings. Some were partially re-observed, but, within our framework, the additional data cannot easily be combined with the original observations (see Section 3.5 below). Even though the final depth in these 25 affected pointings is lower than the rest of the survey, they only constitute 8% of the data.

3.2.2. Correcting for the Effects of a Time-variable Background

A significant background component in the G141 images is the emission of metastable He at 1.083 μm (Brammer et al. 2014), which is negligible in the Earth's shadow, but increases sharply when the spacecraft is outside the shadow. Unlike the earthshine, which only appears on the edge of the image close to the bright Earth limb, the He emission elevates the background across the whole detector. The strength of the He line background depends on the position of the telescope relative to the bright Earth limb and can therefore vary significantly within a single exposure. This time-variable component causes a nonlinear increase in the background counts in consecutive reads. The `calwf3` pipeline uses nonlinearity to identify and filter out cosmic rays during the exposure: a cosmic ray hit in between two reads leads to an increase in the flux of a pixel that is inconsistent with the expectation from the gradual accumulation of charge during the exposure. As the nonlinear background variation mimics the behavior of cosmic rays, `calwf3` flags the majority of pixels in these images as cosmic rays and corrupts the FLT products.

To avoid this unintended `calwf3` behavior and mitigate the effects of the time-variable He background, we redistribute the total counts in the IMA files evenly over the individual samples. We then run `calwf3` on the corrected IMA files using only the final cosmic-ray identification step to produce the final FLT. Exposures that were otherwise rendered unusable due to the variable backgrounds are recovered, albeit with somewhat lower signal-to-noise than unaffected exposures because the overall background count rates are higher. By redistributing the charge, we also retain the ability of `calwf3` to identify cosmic rays using the up-the-ramp sampling.

3.2.3. Grism Flat-fielding and Background Subtraction

Following Brammer et al. (2012b), we first divide the G141 exposures by the JH_{140} imaging flat-field. This neglects the wavelength dependence of the flat-field (which is at most a few percent across the field) in favor of greatly reduced computational complexity. Panel (c) in Figure 4 shows the flat-fielded FLT image in our example pointing. The main effect is the removal of the "wagon wheel" in the lower right corner of the frame.

At each pixel in the grism exposures the background is the sum of different spectral orders sampled at different wavelengths. There is significant structure in the background across the detector resulting from vignetting of the spectral orders, and this structure must be removed to enable extraction of clean spectra of objects. Using on-orbit science observations, Kümmel et al. (2011) created a single master sky image that can be used with the `aXe` software to remove the grism sky background. However, Brammer et al. (2012b) noticed significant variation in the spatial structure of the grism backgrounds and created four separate master sky images that helped to account for the variation.

As described in Brammer et al. (2014), we now understand that the observed variation in the background structure is mainly due to three distinct sources: the zodiacal continuum, scattered light, and the He emission line. Brammer et al. (2014) created master sky images[27] for each of these three physically motivated background components. We fit a linear combination of these component images to each exposure, requiring the zodiacal component to be constant throughout a given visit and allowing for a variable contribution from the emission line component. This technique removes much of the background structure in the grism images. Following Brammer et al. (2012b), we subtract a final masked column average to create the final background-subtracted images to remove low-level residuals not accounted for by the three-component fits.

This final step in the pre-processing sequence is shown in panel (d) of Figure 4. The final images have uniform and low background. The final quality of this example FLT file is representative of all the data in the survey.

3.3. ACS I_{814} Images

We download the CTE-corrected (FLC) images and association tables (ASN) for all observations from MAST. The calibrated images were processed on the fly by the `calacs` pipeline, which is described in detail in the ACS Data Handbook (Chapter 3). In brief, the `calacs` pipeline does all the calibration steps including bias-subtraction, cross-talk correction, dark-subtraction, flat-fielding, cosmic-ray rejection, charge transfer efficiency (CTE) correction, shutter shading correction, and masking of bad and saturated pixels. The final images are in units of electrons.

3.4. ACS G800L Images

The CTE-corrected I_{814} FLC images and association tables were obtained from MAST. The images are then processed with AstroDrizzle to identify cosmic rays. A model for the grism background is obtained by carefully masking all detected objects, scaling the individual exposures to match the average sky values and taking the median of the background pixels. This background model is then subtracted from the individual exposures. The ACS grism images are not flat-fielded. Pirzkal et al. (2002) show that applying a direct-imaging flat to the grism observations introduces $\pm$10% large-scale differences. Without the flat-fielding, these differences are much smaller, $\sim$5% across the detector.

The individual exposures for each pointing are combined by rounding the offsets between exposures to the nearest integer. Because no interpolation is used, this step retains the noise properties of the data, at the expense of also retaining the geometric distortions in the frame. In the following section, we discuss the rationale of this approach in the context of the WFC3 G141 grism data.

The reduced ACS data are part of the 3D-HST data products, and are publicly available. However, in this paper, we limit the spectral extractions and redshift fitting to the WFC3 G141 data.

[27] http://www.stsci.edu/~brammer/grism_sky/

The procedures described below can be applied in the same way to the ACS data. Although the ACS data are shifted and summed and the WFC3 data are interlaced, the final product is similar: distorted frames with noise properties that are preserved, with similar pixel size. We note here that a key advantage of the G141 data over ground-based near-IR spectroscopy, namely the low near-IR background from space compared to the ground, does not apply in the same way to the optical ACS spectra.

3.5. Interlacing

The traditional method of combining dithered images from *HST* is through "drizzle" image processing, which allows for the recovery of resolution in under-sampled images as well as the correction of geometric distortions (Fruchter & Hook 2002). The drizzle algorithm works particularly well when a large number of images are combined. However, in the limit of a few images, it is prone to producing correlated noise. The reason for these noise correlations is that pixels from the individual input images contribute to multiple pixels in the resampled output grid. The amount of this resampling "diffusion" can be controlled in the drizzle algorithm; however, some diffusion is usually necessary to avoid uneven coverage of the output grid. The net effect is a slight smoothing, resulting in a loss of resolution and correlations between adjacent output pixels.

Drizzling is particularly problematic for spectra, as correlated noise can mimic emission or absorption features. Furthermore, the correlated noise is difficult to properly take into account when fitting the spectra, again leading to confusion between noise and real spectral features. Lastly, for spectra, correcting the geometric distortions is not strictly necessary, as long as the mapping between pixels and wavelength is known.

Most 3D-HST pointings (exceptions described below) are comprised of eight images—four direct and four dispersed, observed with a 4-point dither pattern (see Figure 3 of Brammer et al. 2012b). The dithers between the images sample the WFC3 pixels at half-pixel intervals. With this optimal sub-pixel sampling, we *interlace*, rather than drizzle, the original exposures into an output mosaic used for the spectral extractions.

We combine the exposures of each visit into a single output frame by placing the original images onto a subgrid of pixels that are exactly half their original size. The procedure is illustrated in detail in Figure 5 using a portion of a JH_{140} image, and compared to a standard drizzling approach in Figure 6. As the input images have exactly half-pixel offsets by design, this results in a one-to-one correspondence between input and output pixels (e.g., van Dokkum et al. 2000) and offers the key benefit of preserving the individual pixel errors. Adjacent pixels come from sections of the original images that are ∼10 pixels apart and are entirely uncorrelated. Furthermore, interlacing improves the sampling of the PSF by a factor of two without having to interpolate; it therefore produces the highest resolution images that are attainable with the WFC3 camera. Both the G141 and the direct JH_{140} images are interlaced in the same manner. The output G141 images have a pixel size of ∼23 Å × 0.″06. Interlacing is possible because (a) the relative pointing errors between the images of a given set are small (∼0.1 pixels) and (b) the dithers between images are small (⩽10 pixels) and the relative distortion on these scales in WFC3 and ACS is small.

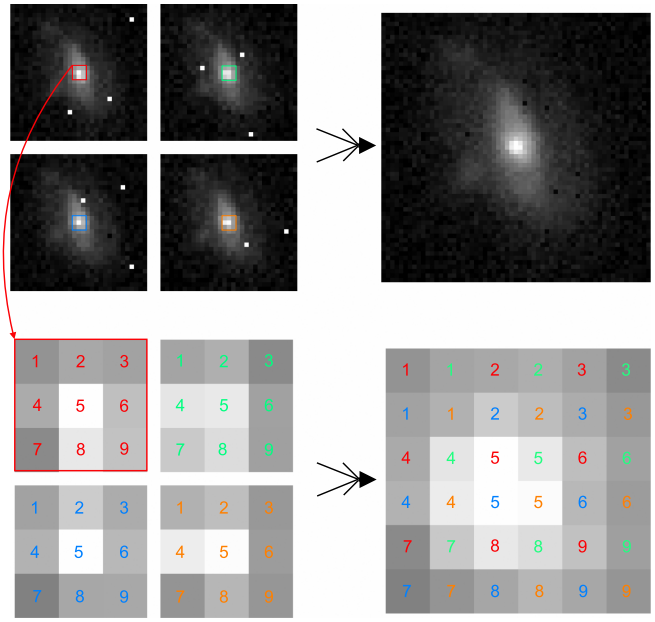

Figure 5. Illustration of interlacing with a small section of a JH_{140} direct image. The same process is also used to interlace the grism exposures. The top left shows the four individual exposures. These are combined to produce the interlaced image in the top right. In the bottom row, we show the same procedure for a 3 × 3 pixel part of the core of the galaxy to demonstrate how the pixels from the individual images are arranged in the final interlaced grid.

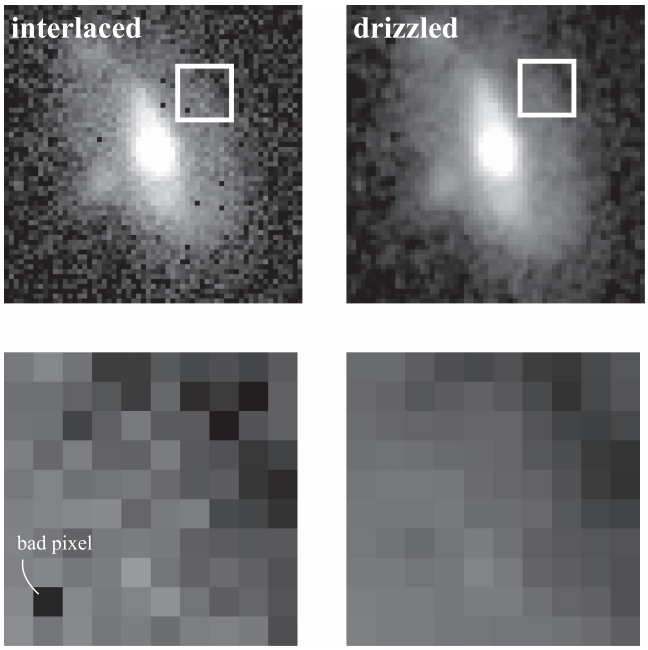

Figure 6. Comparison between interlacing and drizzling. Compared to the drizzled image, the interlaced image has higher resolution, as the pixels were not interpolated. Flagged pixels (due to cosmic rays and chip defects) are retained as single pixels in the interlaced image, whereas they are interpolated over in the drizzled image.

The primary shortcoming of this approach is that if one or more images in a dither sequence are missing, the combined image will have empty pixels.[28] This only affects one of our

[28] Note that this is a shortcoming in the data, not the method, and drizzled images are similarly affected. However, it is less obvious in drizzled images as the missing pixels are effectively interpolated over.

pointings, AEGIS-20, which only has three direct and two grism exposures. Areas that were masked in all four exposures will also have no information. This only affects four 3D-HST pointings: GOODSS-15, where one dither position was repeated, UDS-15, which has many masked pixels affected by the Mars-crossing asteroid 1036 Ganymed, and UDS-25 and UDS-26, which are both affected by long-term persistence. Small portions of these images were masked entirely. Finally, all pixels flagged due to other reasons will also be empty in the final image. These are typically less than 1% of all pixels in the image. We note that empty pixels in the grism images are handled trivially in the fits to the 2D spectra described below: empty pixels do not contribute to the χ^2 or likelihoods of the fits. We stress that drizzled versions of the same data have the same missing information. The only difference is that the missing information is interpolated over in the drizzling process.

Another shortcoming of the interlacing approach is that only observations taken at the same rotation angle can be combined. Observations within a single visit are always taken at the same rotation angle, but re-observations of failed visits are frequently done at a different rotation angle. Within our data set, the re-observations of GOODSN-11, 14, and 23 are done at a different angle from the original visits and cannot be combined with the original observations. Furthermore, the re-observed data in GOODSN-11 contains only two dither positions, which means that only half of the pixels in the image are filled. These data can still be valuable for certain applications and are included in the data release accompanying this paper. However, the lower information content of the spectra from these three pointings decreases the accuracy of the redshift fits. The majority of the spectra from these pointings were flagged in the visual inspections and they are excluded from the analysis in Sections 5 and 6.

3.6. Reference Image, Catalog, and Segmentation Map

Before we can extract spectra from the interlaced frames, we require a reference image, which provides the positions and morphologies of all objects in a given grism pointing. This image sets the wavelength reference for all sources and is used to create a model that accounts for the contamination from overlapping objects. The reference image must be accompanied by a catalog, which defines positions and magnitudes for all objects within the pointing as well as a segmentation map, which defines the pixels that belong to each object in the catalog. Brammer et al. (2012b) used the direct JH_{140} images as a reference and ran `SExtractor` to create a catalog and a segmentation map for each grism pointing. This approach posed two main challenges: (1) repeat objects, which appeared in multiple pointings, could not be directly co-added because they would have different segmentation polygons; and (2) when we later matched the catalog objects to external photometric catalogs the matches were not always unique.

We use the data products of Skelton et al. (2014) to create the reference images, catalogs, and segmentation maps. We make a distinction between the *direct image* (JH_{140} for 3D-HST and AGHAST), which was used to align the exposure WCS, and the *reference image*, which may be a deeper astrometrically aligned image in a different filter. The reference image is used to determine the distribution of light in each source, which defines the spatial morphology of the 2D object spectra. The reference image is a sum of the J_{125}, JH_{140}, and H_{160} WFC3 images. The reference mosaic is created for the whole field by coadding the Skelton et al. (2014) mosaics (before PSF matching) in the three bands: the individual images are scaled to the JH_{140} AB zero point and then co-added with the inverse variance maps used as weights. Although we use the Skelton et al. (2014) catalogs, we do not use their flux measurements as our "standard" magnitudes because these catalogs include areas where the 3D-HST JH_{140} and the CANDELS H_{160} images do not overlap. As a result, there is not one consistent flux/magnitude measurement that exists for all objects in the catalogs. To remedy this, we run `SExtractor` on the co-added image, in dual image mode with the Skelton et al. (2014) detection images and the same settings, in order to determine uniform fluxes and magnitudes that are defined for all objects. These magnitudes, which we refer to as $JH_{\rm IR}$, are used throughout the paper to determine the depth of the fits and to apply magnitude cuts to the final catalogs.

For each grism pointing, we create a reference image by "blotting" the full mosaic to the frame of each of the grism FLT images, where the WCS alignment has been performed in the preparation steps and the `AstroDrizzle.ablot` utility transforms the rectified mosaics into the distorted FLT frame. The individual blotted images are then interlaced in the same manner as the grism exposures. The same procedure is also used to blot the master segmentation map into the FLT frames. Finally, a reference catalog is created for all objects, which fall within the blotted segmentation map, with object pixel positions in the distorted frame computed with the `skytopix.rd2xy` task.

One significant benefit of using an external reference image is that it no longer needs to be limited to the size of the WFC3 field of view. The blue edge of the first order grism spectrum is offset by ~65 (interlaced) pixels to the right from the object position in the direct image and the zeroth order is offset by ~380 (interlaced) pixels to the left of this position (see Figure 8.6 in the WFC3 Instrument Handbook Rajan et al. 2010). Objects that fall off the left (blue) side of the direct image will still be dispersed onto the detector and objects to the right (red) of the image can create zeroth-order spectra within the grism exposures. Such objects need to be taken into account in the contamination model and can also yield scientifically useful spectra. To account for these objects, we make the blotted reference images larger than the original FLT frames by 215 original pixels (430 interlace pixels) on each side along the x-axis. We also add 45 pixels on each side along the y-axis to account for objects along the top and bottom edge of the image. Figures 7(a) and (c) show how the interlaced reference image produced from the mosaic compares to the interlaced grism image. The final interlaced images are 2888 × 2208 pixels.

4. CONTAMINATION MODEL AND SPECTRAL EXTRACTIONS

We extract the 2D spectra of individual objects from the interlaced mosaic images. A key element of the extraction of slitless spectra is creating a model that identifies which pixels constitute the spectrum of a given object, which pixels belong to neighboring sources, and areas where spectra overlap. Our goal is not only to simply identify areas of the image with overlapping spectra, but to create a quantitative model that accurately accounts for overlapping spectra from sources dispersed onto the same or neighboring pixels. The basis of this contamination model is an estimate of the contribution of

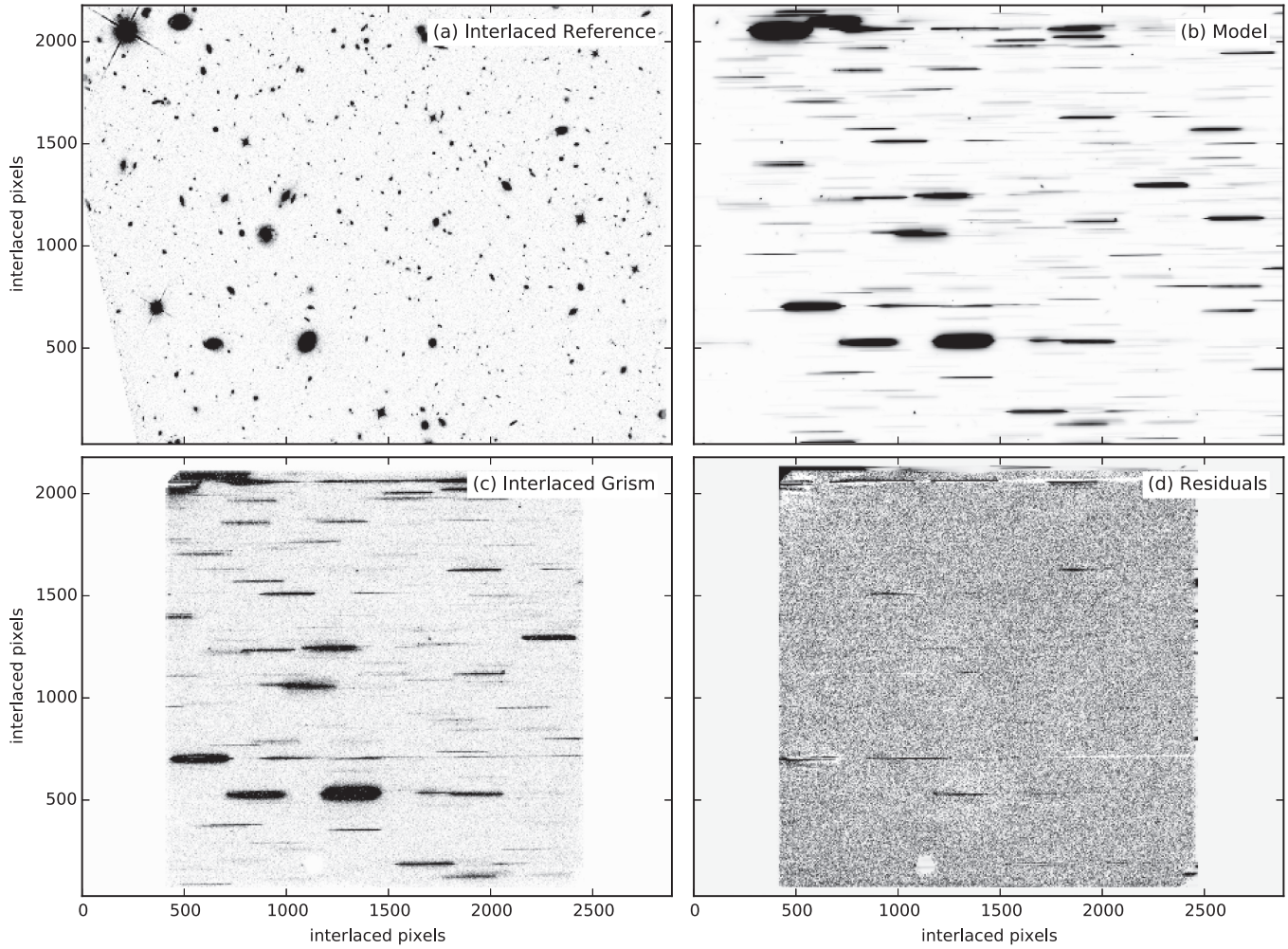

Figure 7. Full contamination model of the COSMOS-04 pointing. The panels show (a) the interlaced direct reference image, created from the CANDELS+3DHST $J_{IR} = J_{125} + JH_{140} + H_{160}$ mosaic of the COSMOS field; (b) the contamination model created using the direct image and model spectra for all the objects; (c) the observed interlaced grism image; and (d) the residuals after subtracting the contamination model from the interlaced grism image.

every source in the direct image to the grism image. The contributions of the individual objects are independent and can be co-added to create a complete model of the grism image. For each object in the grism image then, the contamination model consists of the co-added contributions of all other objects. We refer to this as the *contamination model* to distinguish it from spectral models described in Sections 5 and 6. The accuracy of the contamination model determines the quality of the extracted spectra. Since our goal is to extract high-quality spectra for all objects in the footprint of the survey, the fidelity of the contamination model is of paramount importance. In this section, we describe the approach to creating the quantitative contamination model, the steps of the extraction, and show examples of the final 2D and 1D reduced spectra.

4.1. General Considerations

The grism dispersion varies across the instrument field. The dispersion is described in configuration (CONF) files provided by STScI, such that for a given x and y pixel position in the observed direct image frame, one can determine the trace (position) of the dispersed spectrum for each spectral order in the observed grism exposure, as well as the wavelength along that trace. The position of the trace and the wavelength solution along the trace are described by low order polynomials, where the polynomial coefficients are themselves order-n polynomials that encode the position-dependence of the trace calibration ($n \leqslant 6$). The dispersion varies smoothly across the field of view and the edge-to-edge variations are small. The WFC3 dispersion of the main, +1st order varies between 44.7 Å per pixel and 47.8 Å per pixel across the field (±3% over 1014 pixels). The resulting traces and dispersion are smooth functions of the x and y position in the image and the spectra are slightly tilted (by ∼0.°5) with respect to the detector rows. It is important to note that both the position of the trace and the wavelength along the trace are defined within the coordinate system of the distorted image; adapting them to the interlaced images, which are also distorted (but padded and resampled) is therefore straightforward.

As described in the previous section, we pad the reference images in order to account for objects dispersed within the grism frame. In order to model the spectra for these objects, we assume that the trace and dispersion polynomials continue their smooth behavior outside the field of view of the instrument. We find that this extrapolation is sufficiently stable to enable modeling of the spectra of these outlying objects.

The *HST* grisms are not equipped with order-blocking filters and, therefore, multiple spectral orders are dispersed onto the

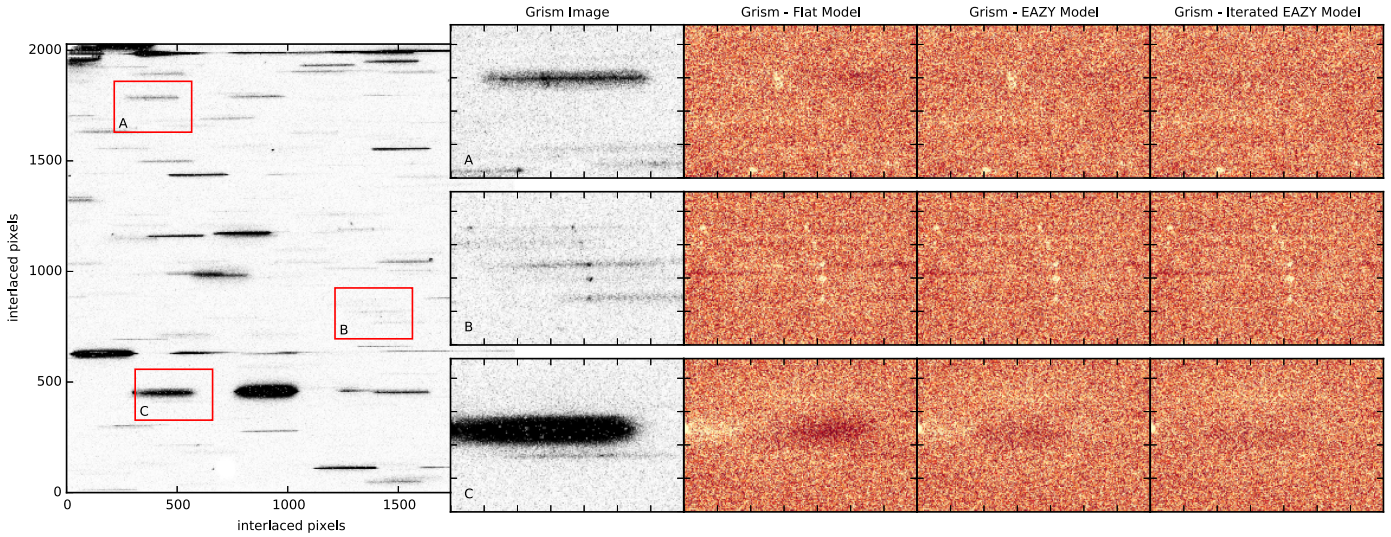

Figure 8. Illustration of the quality of our contamination modeling. For demonstration purposes, we model and subtract all spectra, including that of the object of interest. The pointing is COSMOS-04. The full interlaced grism image is shown for context, with the panels showing zoomed-in portions of the image. The orange panels show the residuals after subtracting three different versions of the contamination model: a flat spectrum (left); the best-fitting EAZY model template (middle); and the best-fitting EAZY template for faint objects plus empirical spectra for bright objects (right). The final model is excellent, with the only significant residual emission lines of faint sources (middle row).

detector for each object (depending on its position in the frame). The primary, +1st, order contains most of the power of the dispersed spectra, followed by the zeroth order, and then by the −1st and by the higher orders. As a result, for most objects, only the zeroth and +1st orders are visible (higher orders are only visible for bright objects such as stars). The zeroth order is only slightly dispersed and appears similar to the undispersed images of objects in the direct images, though offset in position. These spectra can appear much like compact emission line sources and are an important component of the contamination model (the zeroth-order position is fairly well calibrated to a precision of ∼1 pixel in the current configuration files). For the contamination model, we include all of the orders with entries in the configuration files (−1st through +3rd), though we note that the positions and intensities of the higher orders are generally less well calibrated than the zeroth and +1st orders.[29]

4.2. The Contamination Model

The inputs for creating the contamination model are the reference direct image (panel a in Figure 7, see Section 3.6) and the segmentation map, both projected into the distorted interlaced image frame, the `SExtractor` $JH_{\rm IR}$ catalog, also projected into the distorted frame, and the interlaced grism image (panel (c) in Figure 7).

To model the 2D spectrum of a given object, we first compute the trace and dispersion parameters for each spectral order at the center coordinates of that object. These two parameters define the mapping of a single pixel in the reference image into a one-pixel-wide spectrum in the dispersed image. The full 2D model of a given object is then built by shifting and adding this elemental spectrum, scaled by the observed flux in the reference image, for each pixel within the segmentation region. The entire process is analogous to a convolution of the 2D thumbnail in the reference image with an assumed 1D object spectrum.

The two main considerations in creating the models for the individual objects is the treatment of their spatial and spectral light distributions. For modeling the *spatial* distribution, we use a single reference image to define the morphology of an object (constructed from the available $J_{125} + JH_{140} + H_{160}$ mosaics as described in Section 3.6). In favor of computational simplicity, this neglects any wavelength dependence of the source morphology, which may be complex for well-resolved objects (e.g., extended line emission and compact continuum emission). The measurement of the relative sizes and morphologies of continuum and line components of distant galaxies is in itself an important scientific diagnostic largely unique to *HST* slitless spectroscopy (e.g., Nelson et al. 2013). Furthermore, compact objects are susceptible to the change in PSF size as a function of wavelength; the WFC3/IR PSF at 1.7 μm is ∼20% larger than at 1.0 μm. In general, we find that cross-dispersion residuals are small for all but the most compact objects.

We model the *spectral* distribution in the following way. To first order, the full contamination model can be computed by assuming flat source spectra normalized to the observed flux in the reference image. While this would typically be sufficient for contamination *masking* (see Figure 8), our goal is to generate a high-fidelity, *quantitative* contamination model that can be subtracted from the observed spectra. For every object in the 3D-HST photometric catalogs from Skelton et al. (2014), we obtain the best-fit galaxy EAZY SED template determined from the photometric redshift fit (with emission lines removed because these would not be at the correct observed wavelengths based on the imprecise photometric redshift estimates). In some sense, this is similar to the `aXe` fluxcube model that measures fluxes directly from reference images in multiple bands to model the broadband spectrum shape. Here, however, the galaxy template is obtained from the fit to all of the available photometric bands, *HST* and ground-based; the EAZY fits accounts to some extent for line contributions to the broadband fluxes; and, most importantly, the EAZY spectrum is a full

[29] For reference, the dispersion of the ACS G800L grism is 38.8 Å per pixel in the center of the frame; the G800L spectra are tilted by ∼2°; and the contamination model contains the −3rd to +3rd orders.

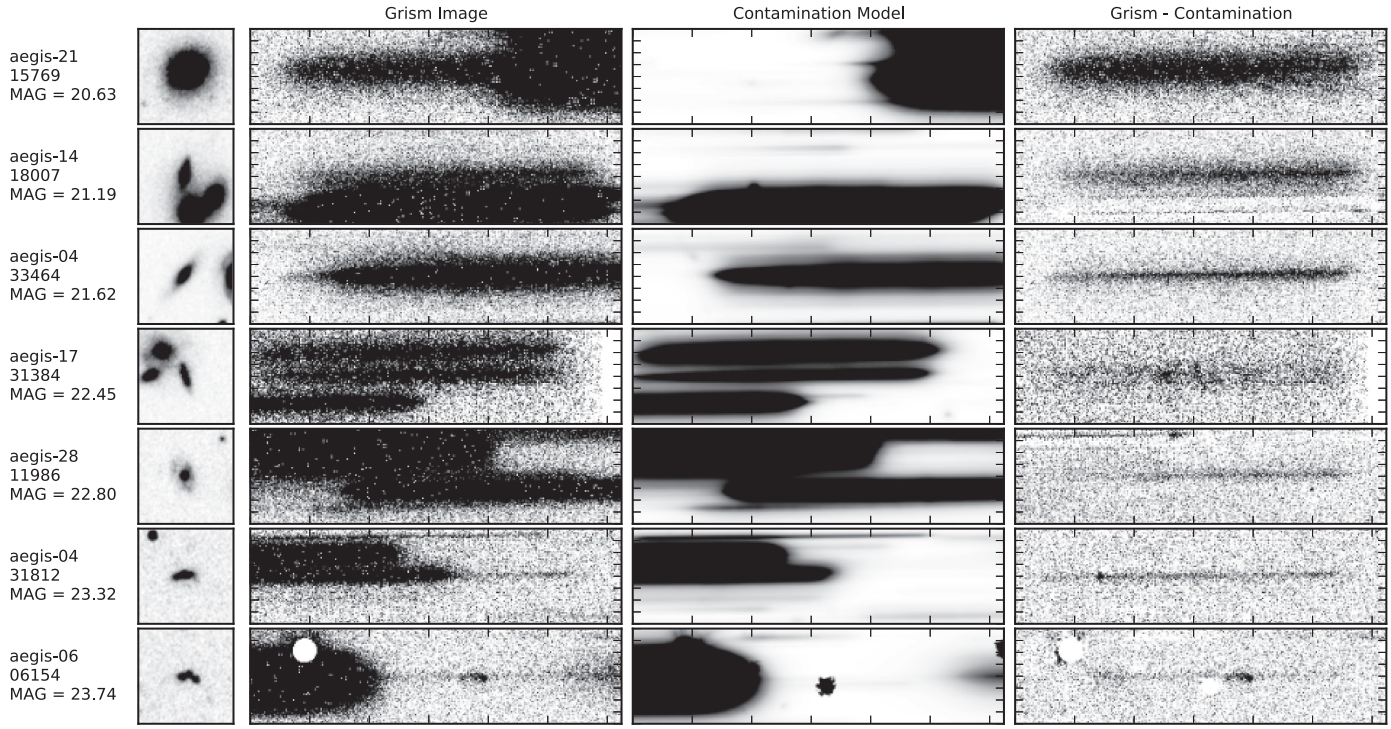

Figure 9. Illustration of contamination subtraction for a range of objects of different magnitudes with substantial contamination. Objects shown here were chosen to have high contamination where the modeled contaminating flux is 70% or more of the expected object continuum flux somewhere along its spectral trace. Creating a quantitatively accurate contamination model is critical because ∼50% of objects at $JH_{\rm IR} \leqslant 24$ have contamination at this level.

stellar population synthesis model, not a polynomial. Once we create a complete contamination model of the pointing from the EAZY templates, we then refine it for objects brighter than $JH_{\rm IR} = 22$ based on the object spectrum itself, extracted along the central pixel of the trace. This refinement step will include emission lines and will correct any mismatch between the EAZY template and the observed spectrum, with the caveat that, if two objects are aligned exactly along the trace, an emission line can be mistakenly ascribed to the wrong object.

The overall quality of the contamination model is illustrated in Figure 7. The full contamination model of the pointing (panel (b)) is subtracted from the interlaced grism image (panel (c)) to highlight the differences between the modeled and observed data (panel (d)). Significant residuals are only seen at the position of the spectra of the two bright stars in the pointing in rows ∼700 and ∼2100, as well as several fainter ones across the image. Overall, the quantitative agreement is excellent. Figure 7 also demonstrates that, as a result of the padding of the reference image, the model extends across the full width of the grism image. Without the padding, the spectra in the leftmost 300 pixels of the grism image would not have been modeled. We note that the model image (panel (b)) is only created for quality control and visualization purposes. The actual contamination model is a set of arrays of wavelength, flux, and position for each object within the pointing, which, combined with the reference image thumbnails and segmentation maps, can be used to reproduce the positions and fluxes of all spectra.

A detailed demonstration of the iterations on the contamination model is shown in Figure 8. To illustrate the differences between the models, we show the residuals for three objects, modeled in three different ways. The first set of panels shows the residuals from a model where flat (f_λ) spectra are assumed for all objects. The second set of panels shows the residuals for models computed with spectra defined by the EAZY templates. The rightmost set of panels show the final model where the spectra of bright objects are refined based on the observed spectrum itself. Three representative galaxies are shown. Galaxy A is a bright galaxy with an emission line. The flat model does not include the slope of the spectrum or the emission line. The EAZY model decreases the residuals in the continuum, but only the iterative approach models both the continuum and the emission line well. Galaxy C is a bright continuum source: the EAZY model decreases the residuals, but the iterative model is slightly better. Panel (B) shows faint emission line sources. In this case all models do equally well. These sources are below the magnitude limit for the iterative model and their lines are not included in any of the models. The combination of the EAZY first-guess approach with the subsequent iteration for bright objects correctly captures the SED for the majority of objects in the survey and accounts for most of the contaminating flux. The fact that emission lines in faint objects are not included in the contamination model could affect redshift fits of overlapping objects. However, since the redshift and emission line fits are done in 2D (Sections 5 and 6 below), this is only a serious issue if the traces of two objects with the same morphology are exactly aligned, which is exceedingly rare.

A final illustration of the contamination model is provided in Figure 9, where we specifically show objects that have modeled contaminating flux that is at least 70% of the expected continuum flux at a point along the trace. Approximately 50% of objects with $JH_{\rm IR} \leqslant 24$ have contamination at this level, and 95% of objects with $24 < JH_{\rm IR} < 26$. If we were unable to subtract the contaminating flux, the vast majority of spectra in the survey, especially at faint magnitudes, would not be usable.

However, as shown, the contamination-subtracted spectra are clean, with little or no residual flux from other objects.

We note here that the contamination modeling procedure as described here is a *static* model. That is, for the redshift and emission line fits for individual objects described in Sections 5 and 6, the static model of all contaminating sources is subtracted before performing the fit. Future generations of the fitting software could perform the fitting in an iterative process, updating the static model of the full field after performing the detailed fits to each spectrum individually.

4.3. Extracting 2D and 1D Spectra

The 2D spectra are the main "basic" data product of the survey. With the contamination model in hand, the spectral extraction is relatively straightforward. For each object, we use the configuration file to determine the pixels that contain the first order dispersed spectrum. We extract the same number of pixels (312) along the dispersion axis for each object, which results in non-uniform wavelength grids for the extracted spectra as the dispersion $\Delta\lambda$/pix varies across the field. The extent of the spectrum in the spatial direction is chosen to be three times the `SExtractor` FLUX_RADIUS (with a minimum value of 26 interlaced pixels or 1.″56 enforced for small objects),[30] and a 2D cutout is extracted from the interlaced grism image. Cutouts are then extracted from the same pixels in the interlaced error image and the static contamination model generated for all objects other than the object of interest, and a separate extension is created for the object model itself. Square cutouts of the interlaced reference image and segmentation map are also extracted with the same cross-dispersion dimensions.

With additional extensions defining the wavelength grid and sensitivity curve, the 2D FITS files, therefore, contain all of the necessary information for detailed modeling of the spectra with the `threedhst`[31] analysis software, such as the redshift and emission line fits described in Sections 5 and 6. We emphasize that modeling the interlaced spectra directly in 2D is critical to the success of the fits and ultimately the quality of the catalogs derived from the spectra.

3D-HST is tiled in such a way that adjacent pointings occasionally overlap. For objects that fall within the overlap region, we can obtain two (or more) independent 2D spectra, typically observed at different rotation angles. Additionally, the HUDF area of the GOODS-S field was observed in five separate visits. We refer to these as *duplicate spectra*. Such repeat spectra can be co-added (see Brammer et al. 2013) and in van Dokkum et al. (2013a) we released the full depth grism spectra in the HUDF. In the current release, all spectra have consistent 2-orbit depth and repeat observations are not co-added.

While the analysis of ground-based slit spectra is typically performed on 1D extractions, collapsing the data to 1D leads to important information losses in the case of the slitless grism spectra. Galaxies often have complex morphologies in the 2D spectra (see, e.g., Nelson et al. 2015), and several distinct clumps of line emission would be degraded to a single broad feature in a 1D spectrum. Although we do not use 1D spectra in the fits, we provide them for convenience and to facilitate plotting and visualization. The 1D spectra are optimally weighted (Horne 1986). We limit the extraction window to the region where the average flux in the spatial direction is greater than 10% of its maximum to stay within the high S/N area of the spectrum.

With the data release described in this paper, we deliver spectra for all objects in the Skelton et al. (2014) catalogs that fall within the 3D-HST grism pointings. A total of 246,052 2D and 1D spectra have been extracted. In Figure 10, we show several examples of 2D and 1D spectra. For each object, we show the reference image as well as the 2D interlaced spectrum, the 2D contamination-subtracted spectrum, and the 2D spectrum with the continuum model subtracted (see below). The latter spectra are essentially 2D maps of line emission (see Nelson et al. 2015). Also shown are the 1D spectrum and the broadband SED.[32] The objects were chosen to span a range in $JH_{\rm IR}$ magnitude, redshift, and color. In all examples, the extractions are clean and contaminating spectra from nearby objects are well subtracted. Note that the emission lines have a similar morphology as the reference images of the objects.

5. REDSHIFT FITS

Most large-scale extragalactic surveys using low resolution slitless spectroscopy have focused on selecting samples of emission line objects (e.g., Zamorano et al. 1994; Gallego et al. 1996; Colbert et al. 2013; Morris et al. 2015). However, 3D-HST has broad science goals that require measurements for galaxies whether or not emission lines are present. In this section, we describe our approach to fitting redshifts in a uniform manner for the complete sample of galaxy spectra.

5.1. Methodology

Due to the low resolution and limited wavelength coverage of the slitless *HST* spectra, the information content of the spectra is often insufficient for robust redshift determinations from the spectrum alone. For example, absorption features are rarely detected with sufficient S/N in individual spectra and common emission line redshift indicators are unresolved (e.g., [O II] $\lambda\lambda$3727, 3729 and Hα+[N II]). Furthermore, because of the fixed bandpass of the grism, the rest-frame wavelength coverage is smaller for higher redshift objects: at $z = 0.7$ the grism captures ∼4100 Å of the optical and NIR spectrum, while at $z = 2.0$, only ∼2300 Å of the UV/optical spectrum falls within the grism wavelength range. With such a limited wavelength coverage, it is common that only one prominent spectral feature appears within the G141 spectral coverage.

For the 3D-HST redshift measurements, we combine the additional information from the (multi-band) photometric catalogs with the information in the (2D) spectrum and fit both components simultaneously. We note that the rich photometric data sets alone often provide photometric redshift precision of ∼1%–3%, thereby leading to little ambiguity in identifying lines detected in the spectra. For the combined photometry + redshift fits, we use a modified version of the EAZY templates. Emission lines are removed from the standard templates, and we supplement these continuum templates with an additional pair of emission line templates. The emission templates contain *just* emission lines, with line

[30] We also provide cutouts with a fixed size in the spatial direction of 80 interlaced pixels. These cutouts are particularly useful for stacking analyses, and the analysis of spectra that are neighboring the object of interest.

[31] https://github.com/gbrammer/threedhst

[32] All data products in Figure 10 are provided in the data release.

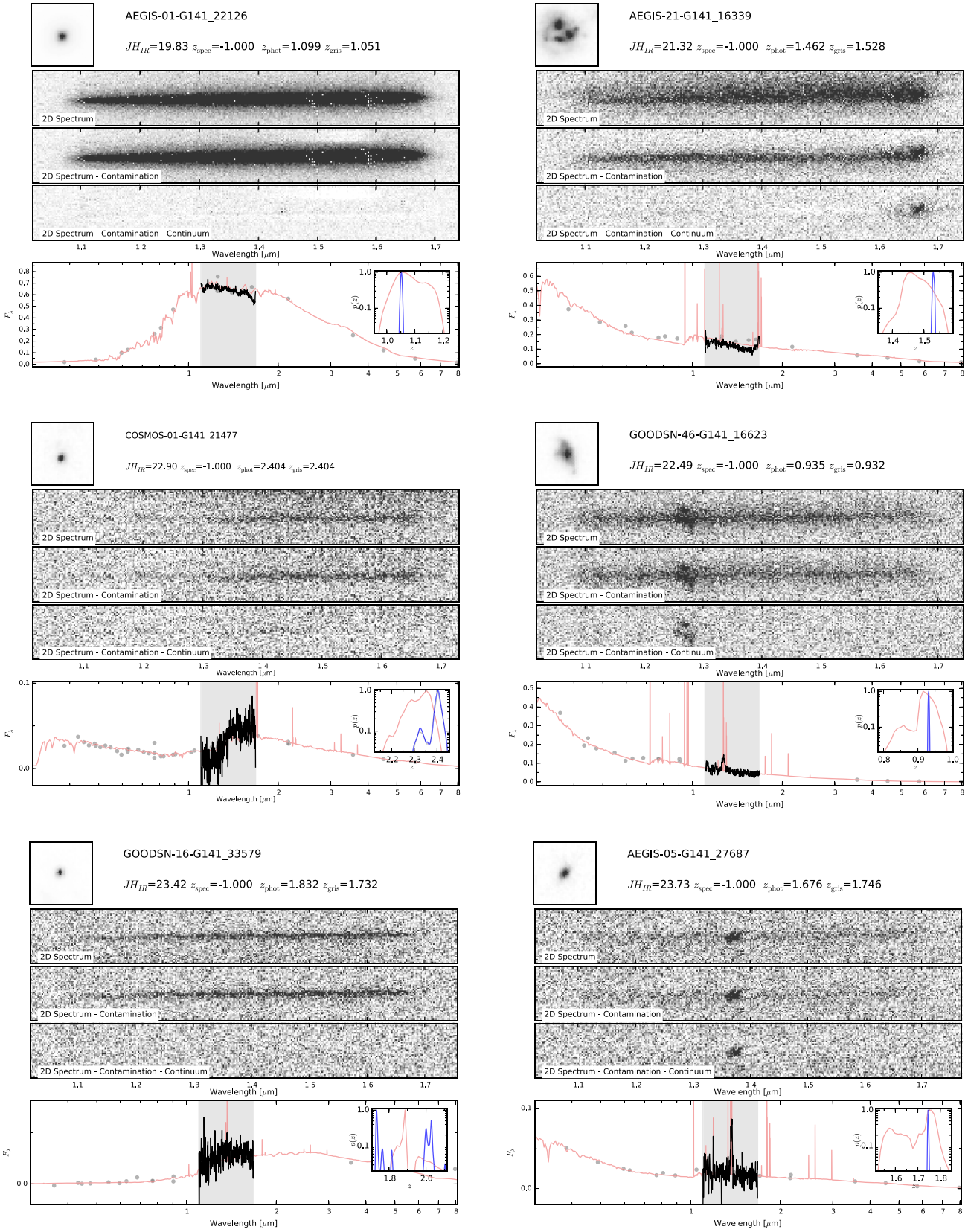

Figure 10. Examples of grism spectra. In the left column, we show three quiescent galaxies with decreasing JH_{IR} magnitudes (top to bottom) and, in the right column, we show three emission line objects. For each object, we show the reference direct image (top), the 2D grism spectrum extracted from the G141 image, the contamination-subtracted 2D spectrum and the continuum-subtracted 2D spectrum (labeled). In the bottom panel, in each plot, we show the optimally extracted 1D spectrum (black line), overlaid on the photometric points (gray circles) and the best-fit redshift template (red line). The inset shows the redshift probability distributions for the photometry fit alone (red) and for the joint photometry+grism fit (blue).

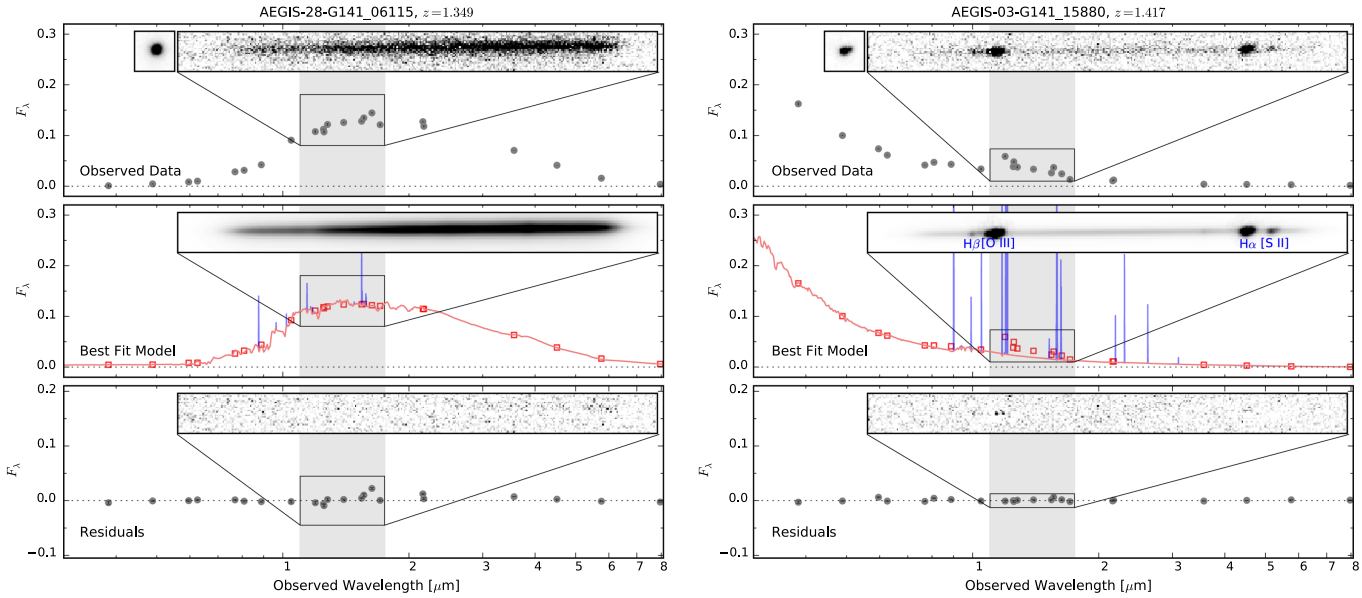

Figure 11. Examples of the redshift fitting procedure for two objects: AEGIS-03-G141_06115, a quiescent galaxy at $z = 1.349$ and AEGIS-28-G141_15880, a star-forming galaxy at $z = 1.417$. In the top panels, we show the observed data: the photometric fluxes (gray circles, errors are smaller than points), the 2D image of the object (square inset), and the 2D grism spectrum (inset). In the middle panels, we show the best-fit linear combination of templates (red lines, with the emission lines added in blue). We also show the best-fit 2D model of the spectrum, which is a convolution of the template and the 2D direct image. In the bottom panels, we show the residuals (gray points, inset), which are minimized in the fitting process. Note that the 2D spectrum and the photometry are modeled simultaneously, in a single fit.

ratios taken from the average SDSS star-forming galaxy spectra of Dobos et al. (2012). The pair of templates brackets extreme values of the Hβ/[O III] ratios to allow for flexibility in the fits, as this ratio has been observed to evolve significantly with redshift (e.g., Steidel et al. 2014).

After computing a scalar normalization factor s between the photometry and spectrum (which are generally offset from one another, due to aperture effects), the data provided to the fit are the individual photometric measurements f_i (with uncertainties σ_i) along with the scaled flux in *each pixel* in the contamination-subtracted 2D grism science and error spectra (G_{xy} and ϵ_{xy}). A 2D spectrum model ($T_{xy,j}$) is computed for each redshifted template, j, as described in Section 4.3, which is also convolved with the photometric bandpasses ($T_{i,j}$). The final likelihood of the fit at redshift, z, is then

$$T_j = \sum_j^{N_{\rm templ}} \alpha_j \cdot T_{i,j}, \quad (1)$$

$$T_{xy} = \sum_j^{N_{\rm templ}} \alpha_j \cdot T_{xy,j}, \quad (2)$$

$$\epsilon^2_{xy} = g^2_{xy} + \gamma \cdot c^2_{xy}, \quad (3)$$

$$\chi^2 = \sum_i^{N_{\rm filt}} \frac{(f_i - T_i)^2}{\sigma_i^2} + \sum_{xy}^{\rm 2D} \frac{(s \cdot G_{xy} - T_{xy})^2}{\epsilon^2_{xy}},$$

$$\mathcal{L} = e^{-\chi^2/2}, \quad (4)$$

where the individual non-negative template normalizations α in Equations (1)–(2) are computed with the "NMF" algorithm of Sha et al. (2007) (see also Blanton & Roweis 2007; Brammer et al. 2008). The per-pixel spectrum uncertainties ϵ_{xy} in Equation (3) include an additional term to reduce the contribution of pixels with estimated contamination count rates c_{xy} (with $\gamma = 1$; higher values would result in more aggressive contamination masking). Equation (4) demonstrates the power of fitting the interlaced 3D-HST spectra: at no point in the processing steps outlined above are the pixels G_{xy} in the 2D spectrum resampled and the per-pixel spectrum uncertainties ϵ_{xy} are preserved from the instrumental noise model. Missing pixels (hot pixels, etc.) in the 2D interlaced spectrum are simply excluded from the fit.

As the generation of the convolved G141 template spectra is the computational bottleneck for the redshift fit, it is done in three steps to reduce the computational time. In the first two steps, we fit the photometry and spectrum separately, over a complete but coarse redshift grid that is just sufficient to identify strong emission lines. We multiply the likelihoods from the separate fits. We finally carry out a third fit to the spectrum and photometry simultaneously (Equation (4)) on a fine wavelength grid, informed by the redshift range where the joint likelihood is the highest. To allow more freedom in this final step, we assume that some small fraction (1×10^{-4}) of the overall likelihood is evenly distributed between $z = 0$ and $z = 4$. In the limit of low S/N in the spectrum (G_{xy}, $T_{xy} \to 0$), the contribution of the spectrum to the likelihood in Equation (4) will be effectively constant with redshift, and the fit will be dominated by the first term from the photometry. That is, it will be similar to the EAZY photometric redshift fit, although not identical due to the adoption of different continuum and emission line templates as described above.

Examples of the redshift fitting procedure are shown in Figure 11 for two objects, a quiescent and a star-forming galaxy, both at similar redshifts, $z \sim 1.4$. In the top two panels, we show the observed data: the photometric fluxes (gray points), the 2D image of the object and the 2D grism spectrum (insets). The model is evaluated on how well it fits both the photometric points and the 2D grism spectrum. In the middle panels, we show the best-fit combination of templates (red lines, with the emission lines added in blue). The template is then projected into the observed space as we derive the observed fluxes in the photometric bandpasses (red squares)

and convolve it with the 2D direct image to create the model 2D spectrum (inset). In the bottom panels, we show the residuals, which are minimized in the fitting process.

The best estimate for the galaxy redshift determined from the photometry+spectrum fit, $z_{\rm grism}$, is taken to be the redshift where $\mathcal{L}$ is maximized. We save the full likelihood distribution $\mathcal{L}(z)$, from which we calculate the central redshift 68% and 95% confidence interval for each object.

The redshift fits have been done for all objects with $JH_{\rm IR} < 26$. We require that a 2D spectrum has at least one non-zero column in order to be fit. This resulted in a total of 98,668 spectral fits for 75,386 unique objects in the Skelton et al. (2014) catalog. The limiting magnitude of these fits is well below where we can expect to detect continuum emission at sufficient S/N (Brammer et al. 2012b). At $JH_{\rm IR} = 26$, a pure emission line object (EW $= \infty$) will have a line flux of $f \sim 9 \times 10^{-17}$ erg s^{-1} cm^{-2}. Similarly, a resolved line detected at S/N > 5 ($f \sim 5 \times 10^{-17}$ erg s^{-1} cm^{-2} at 1.5 μm) will have EW$_{\rm obs} > 5990$ Å. In the remainder of this paper, we focus our attention on the brighter $JH_{\rm IR} < 24$ sample of 30,621 2D spectra, which we have visually inspected in its entirety (Section 5.2) and assigned quality flags to. The catalogs described in Section 7.2 contain the full sample at $JH_{\rm IR} < 26$. While fainter galaxies with $24 < JH_{\rm IR} < 26$ can be useful for scientific analyses, we caution that the quality of the redshift fits for those fainter sources has not been fully validated from visual inspections.

5.2. Visual Inspections

We visually inspected the 30,621 spectra and redshift fits of all 23,564 objects with $JH_{\rm IR} < 24$. The first goal of the visual inspections was to identify any systematic problems in the reductions and redshift fits, which led to the finding that the fits for bright, red, low redshift galaxies are unreliable (see end of this section). No other systematic problems were identified. The secondary goal of the inspections was to assign a quality flag for spectra affected by known failure modes such as the incomplete masking of zeroth-order spectra, which can mimic emission lines, residuals from the spectra of very bright stars, which may not be subtracted properly, and instances where corrupted photometric measurements lead to errors in the spectral fit. The inspections were done by 12 individuals such that each spectrum was seen twice and given two separate flags. The primary criterion of the inspections was "is the redshift clearly affected by a systematic error in the spectrum?" These inspections are therefore somewhat permissive in favor of completeness. For example, a spectrum with substantial contamination residuals from an overlapping object would *not* be marked as "bad" if the fit identifies an unambiguous emission line. Furthermore, reliable redshift measurements can be obtained for objects at the edge of the field with less than a complete spectrum if the available coverage includes an emission line. Redshift fits where the grism spectrum contributes little to the fit and the final redshift probability is identical to that from the photometry alone are also marked as "good." These inspections do not attempt to flag "bad" spectra because the criteria for what constitutes a bad spectrum will heavily depend on the particular application. For these reasons, we caution against blindly using the spectra for objects identified to have reliable redshift measurements for other purposes.

Individual classifications are either "good," "bad," or "unclear." Of the 30,621 spectra classified, 2824 are flagged as "bad" by at least one user and 900 are flagged as "unclear" by at least one user. Overall 3540 (11.5%) spectra have at least one flag set to "bad" or "unclear." This selection is a bit more permissive than other cuts used for grism spectra. For example, Fumagalli et al. (2012) require that more than 75% of the spectrum falls on the detector (~12.5% of all spectra are excluded) and that the average contamination is lower than 10%, which results in removing 60% of the initial sample. In general, we find that a successful redshift fit does not depend on the fraction of the spectrum, which falls on the detector or the amount of contamination (to a point). However, these requirements can be useful in selecting spectra for stacking or other purposes where the successful redshift fit is not the only requirement.

At this stage, we reconcile redshift measurements for duplicate objects, so that for all these objects, we have a primary measurement, which appears in catalogs with one line per object, and a secondary measurement. Out of the 30,621 spectra fit down to $JH_{\rm IR} = 24$, 7057 are repeat observations. If there are two (or more) spectral fits for a given object, we first remove those that have at least one flag set to "bad" or both flags set to "unclear." Among the remaining, we chose the fit with the narrowest $p(z)$ (as measured by the 68% confidence interval) to include in the catalog (if only one spectrum remains, it is the default choice). If all spectra for a given object have at least one flag set to "bad" or both flags set to "unclear," none of them are included in the catalogs (there are only 149 such objects in the full five fields).

The subjective classifications are combined into a single flag in the redshift catalogs (`use_subj`). Objects have `use_subj = 0` if either one of the classifications is "bad," or if both are "unclear." Spectra with no "bad" classifications and at least one "good" classification are assigned `use_subj = 1`. Of all classified objects, only 1686 (7.1%) have `use_subj = 0`. Limiting the sample to the 21,876 non-stellar objects according to the Skelton et al. (2014) criteria (`star_flag = 0`), 1314 objects (6.0%) have `use_subj = 0`.

As shown in the next subsection, redshifts derived from the combined fit to the photometry plus the G141 spectrum ($z_{\rm grism}$) are generally more precise than redshifts derived from the photometry alone ($z_{\rm phot}$). However, there are exceptions: for some bright, red galaxies at low redshift, we find that the photometric redshift provides somewhat more accurate redshifts than the grism redshifts. This is likely due to subtle issues with our reddest, oldest templates; as discussed in Fumagalli (2016) templates based on popular stellar population synthesis libraries do not always accurately reproduce the broad absorption features at $\lambda_{\rm rest} > 0.8$ μm in old galaxies. We find that the photometric redshifts are preferable to the grism redshifts for the small subset of galaxies whose grism redshifts do not fall in the 95% confidence interval of the photometric redshifts and either have $z_{\rm grism} < 0.55$ or have $z_{\rm phot} < 0.65$ and $(U - B)_{\rm rest} > 0.9$.

The final `use_zgrism` flag is set to 1 if the subjective flag is 1, the object has `star_flag = 0`, and the criteria to remove bright, red, low redshift galaxies with incorrectly identified spectral features are met. This results in 18,927 objects with `use_zgrism = 1`.

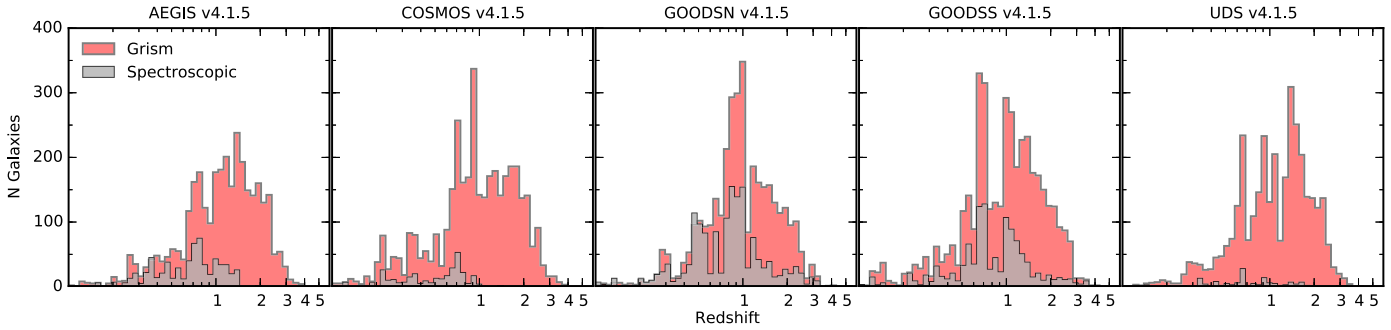

Figure 12. Distributions of grism redshifts for objects with $JH_{\rm IR} \leqslant 24$ (red histogram) and ground-based spectroscopic redshifts (gray histogram) for the five deep extragalactic fields.

5.3. Redshift Accuracy and Precision

Understanding the accuracy of the grism redshifts ($z_{\rm grism}$) is of paramount importance, especially since this is, to our knowledge, the first time joint photometric and spectral fits have been done for a complete, magnitude-limited, sample of objects (irrespective of their spectral characteristics). In this section, we test the grism redshifts in three different ways: by comparing them to spectroscopic redshifts, by comparing the redshifts of projected pairs, and by comparing repeat redshift measurements. The redshift accuracy may depend on the characteristics of the objects (e.g., their magnitudes, colors, and the presence of strong emission lines), and we investigate this as well. The accuracy of the photometry and photometric redshift fits, which we use jointly with the grism spectra, to a large extent predetermines the accuracy of the redshifts in this work. A detailed analysis of the photometric redshift errors is outside the scope of this work and is presented in a companion paper, Bezanson (2016).

5.3.1. Ground-based Spectroscopic Redshifts

Comparison to ground-based spectroscopic redshifts is the most common method for determining the accuracy and precision of a redshift sample. In Figure 12, the spectroscopic redshift distribution is compared to the grism redshift distribution (for objects with $JH_{\rm IR} < 24$). The spectroscopic redshifts come from the entire CANDELS area (see Skelton et al. 2014), and are not limited to the area with grism coverage. There are a total of 5361 ground-based redshifts in these fields (see Section 5.1 of Skelton et al. (2014), for references to the sources of the spectroscopic redshifts). It is clear that the 3D-HST survey provides a major step forward in the spectroscopic coverage of the CANDELS fields. In all five fields the 3D-HST survey now provides the vast majority of redshifts $z > 1$, and in COSMOS and UDS it is even the most important source of redshifts at low redshift. The field with the largest number of ground-based spectroscopic redshifts is GOODS-N, followed by GOODS-S. It is reassuring that the 3D-HST grism redshifts show the same redshift peaks as the ground-based spectroscopic redshifts.

In Figure 13, we compare the ground-based spectroscopic redshifts to the grism redshifts for the 3278 objects where both exist. There is an excellent overall agreement between the two. We specifically focus our attention on the $z > 0.7$ sample where Hα enters the G141 coverage and the redshift accuracy may be expected to be the highest. We use the normalized median absolute deviation (NMAD), $\sigma_{\rm NMAD}$, to quantitatively characterize the scatter in $\Delta z/(z+1)$ (for a definition see Brammer et al. 2008). For a Gaussian distribution, NMAD is equal to the standard deviation, but it is less sensitive to outliers. In all fields, the $\sigma_{\rm NMAD}$ for the $z > 0.7$ sample is between 0.0023 and 0.0032. At 1.4 μm, this scatter corresponds to 42 Å, i.e., approximately one original grism pixel (46 Å). Uncertainties at this level are expected from morphological effects alone: if the luminosity-weighted center of the gas distribution in galaxies differs from that of the stars by ∼1 kpc it will introduce an error in the emission line redshift of ∼1000 km/s (see, e.g., Nelson et al. 2015). The scatter in $\Delta z/(z+1)$ is centered at zero and symmetric about zero, indicating that there are no systematic offsets between our redshift measurements and the ground-based samples.

In each field, there are clear outliers that are many σ removed from the one-to-one line. This appears particularly pronounced in GOODS-N, although that is largely simply due to the large number of ground-based spectroscopic measurements in that field. These outliers are a mix of objects with large formal uncertainties in their grism redshifts, possible object confusion, errors in the ground-based spectroscopic redshifts, AGN, and genuine outliers with no straightforward explanation. Object confusion may occur due to the matching of our *HST* catalogs to ground-based spectroscopic measurements. Errors in spectroscopic redshifts may come from mis-identified lines or from low S/N spectra. It is difficult to disentangle these effects, particularly since the original ground-based catalogs and spectra are usually not available.

Three surveys have used preliminary catalogs from 3D-HST to select galaxies for ground-based near-infrared spectroscopic follow-up: MOSDEF (Kriek et al. 2015), KMOS^{3D} (Wisnioski et al. 2015), and VIRIAL (Mendel et al. 2015). MOSDEF uses the multi-object slit spectrograph MOSFIRE on the Keck telescope; KMOS^{3D} and VIRIAL use the multi-object integral field spectrometer KMOS on the Very Large Telescope. Since the selection of these objects is not independent of the grism redshifts, they were not included in Figure 13. Nevertheless, a comparison between their redshifts and those from the grism is informative, as they rely strictly on rest-optical features observed in the near-infrared, and probe a regime of galaxy parameter space that is underrepresented in optically selected ground-based spectroscopic surveys. In Figure 14, we show a comparison between the grism redshifts and those from the first year campaigns of these ground-based collaborating surveys. The sample comprises 581 objects with $JH_{\rm IR} \leqslant 24$. The scatter, $\sigma_{\rm NMAD}$, is ∼0.0015–0.0045, comparable to the the scatter from the larger ground-based sample. All three surveys only use the

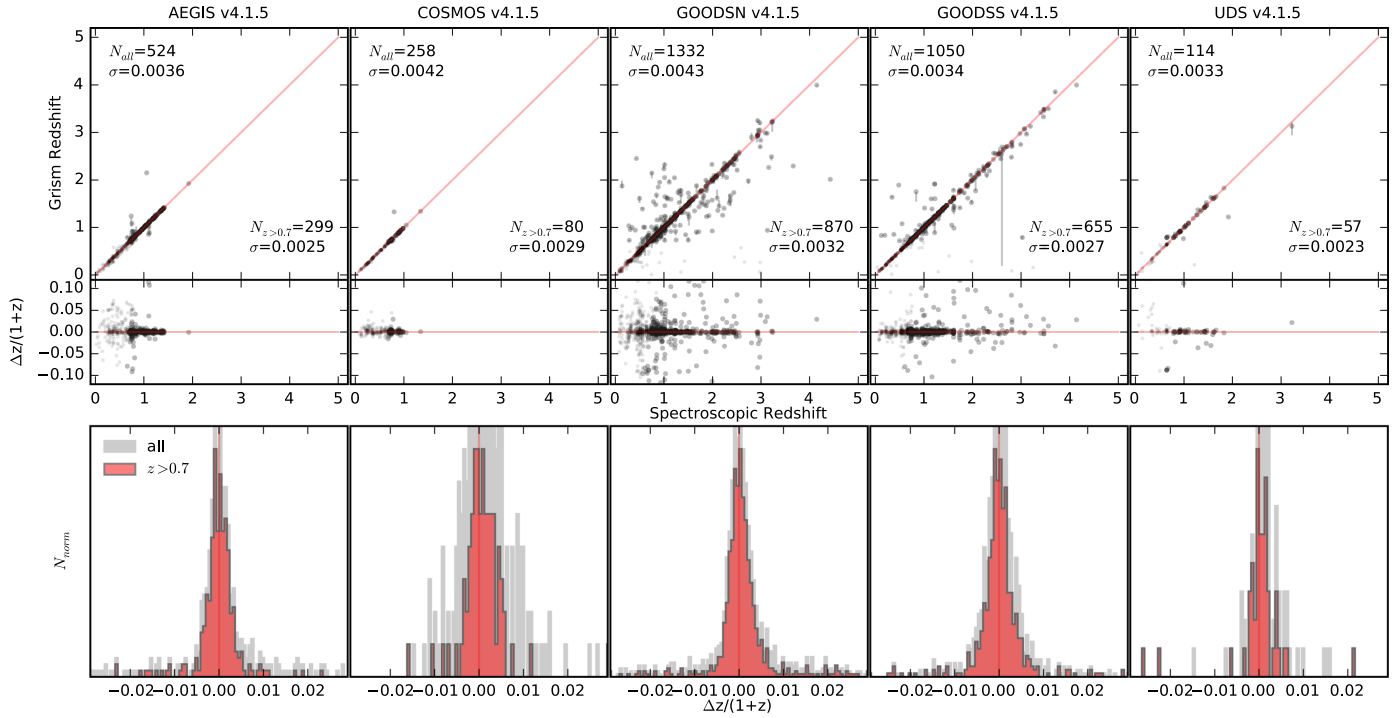

Figure 13. Comparison between grism and spectroscopic redshifts. Top row: spectroscopic vs. grism redshifts for 3278 objects. Objects with $z_{\rm spec}$, $z_{\rm grism} < 0.7$ are plotted with smaller symbols. In the top left corner, the total number of objects in each field is listed, along with the NMAD scatter. In the lower right corner of each panel the same statistics are listed for the $z > 0.7$ sample. Middle row: $\Delta z/(z+1)$ as a function of redshift. Bottom row: distributions of $\Delta z/(z+1)$ for the full sample (gray) and for the $z > 0.7$ sample (red). The distributions are centered on zero and symmetric in all fields.

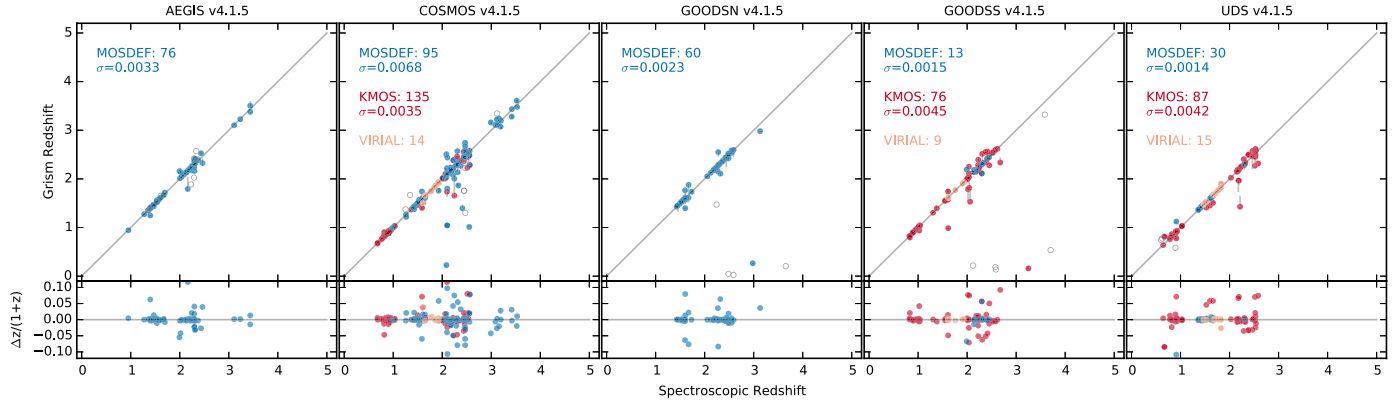

Figure 14. Comparison between grism redshifts and spectroscopic redshifts from surveys that use the 3D-HST catalogs for target selection. Shown are 274 measurements from the MOSDEF survey (Kriek et al. 2015), 298 from KMOS^{3D} (Wisnioski et al. 2015), and 38 from VIRIAL (Mendel et al. 2015). Only objects with $JH_{\rm IR} < 24$ are shown. Measurements are color-coded to indicate the survey. Open points have `use_grism` = 0. Top row: spectroscopic vs. grism redshift. In each panel, we show the number of objects from each survey and the NMAD scatter in $\Delta z/(1+z)$. The outlier clipped NMAD scatter is in brackets. Bottom row: $\Delta z/(z+1)$ as a function of redshift.

3D-HST preliminary grism redshift and do not specifically target objects with detected emission lines in the grism, hence it is expected that the scatter is slightly larger relative to the ground-based sample. The scatter in the COSMOS MOSDEF sample is particularly large, likely because this sample has a larger fraction of faint $JH_{\rm IR} > 23.5$ targets (40% versus 15%–30% for the other fields). If we limit the COSMOS sample to $JH_{\rm IR} < 23.5$, the scatter is consistent with the other fields. While both KMOS^{3D} and MOSDEF target primarily star-forming galaxies, VIRIAL is specifically focused on quiescent objects without emission lines. Strikingly, even in that comparison our redshift accuracy is excellent, $\sigma_{\rm NMAD} = 0.0034$ for the full sample.

5.3.2. Pair Analysis

While comparison to spectroscopic redshifts is the most straightforward method of determining redshift accuracy, such a comparison can be limited because the spectroscopic sample is not fully representative of the parent grism sample. The sample used in Section 5.3.1 is dominated by low redshift bright objects: 70% of objects are at $z < 1$ and 56% are brighter than $JH_{\rm IR} = 22$. Furthermore, they are also likely dominated by bluer galaxies (emission lines make for easier redshift identification) and biased toward specific sub-populations (Lyman break galaxies, for example). Therefore, the redshift accuracy we quoted in the previous section may not apply to the full grism

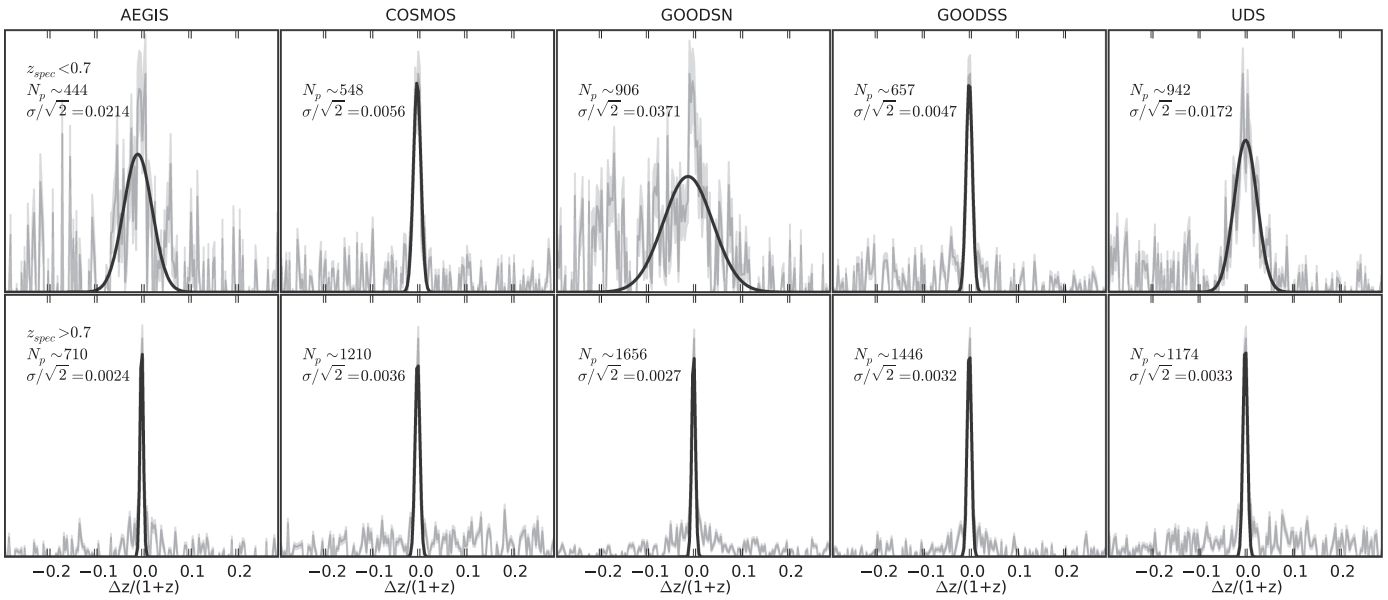

Figure 15. Distribution of $\Delta z_{\rm grism}/(1+z_{\rm mean})$ after correcting for projected pairs for each of the five fields in our sample, divided into two bins of redshift: $z < 0.7$ (top) and $z > 0.7$ (bottom). Tracer galaxies are limited to those within the redshift bin, with magnitudes of $JH_{\rm IR} < 24$, and with grism use flags set to one. The companion galaxies have the same selection, except that there is no constraint on their redshifts. The number of pairs that contribute to each measurement is calculated by integrating under the best-fit Gaussian (solid black line). As a result of the different geometries of the fields, they produce different numbers of pairs. We remove the *y*-axis labels and scale the plots by the maximum histogram value for clarity. The high redshift objects have redshift accuracies of $\sigma \sim 0.0025$–0.0035, consistent with the comparison to spectroscopic redshifts. The low redshift objects have significantly larger scatter.

sample. In this section, we use an empirical method to determine the redshift accuracy, first proposed by Quadri & Williams (2010). The method uses pairwise redshift differences to estimate the width and shape of the error distribution for the full sample. In brief, the method takes advantage of the fact that galaxies are strongly clustered in real space and projected pairs have a high likelihood of being at the same redshift. The errors in grism redshifts, limited by the WFC3/IR grism resolution, are larger than the true virial motions of galaxies within structures; this aspect makes this method (originally used by Quadri & Williams 2010, on photometric catalogs) applicable to our sample.

Following the methodology described in Quadri & Williams (2010), we choose primary (tracer) galaxies with given characteristics (magnitude range, redshift range, color) as well as a secondary (companion) population. For each tracer galaxy, we identify all galaxies within the companion population that are within $r = 25''$ and determine $(z_t - z_s)/(1 + z_{\rm mean})$, where z_t and z_s are the corresponding redshifts of the tracer and companion. In the analysis presented here, both z_t and z_s are grism redshifts. To subtract the background contribution of uncorrelated pairs, we assign the companions random positions from our master catalog. We then identify again all pairs for each tracer galaxy. The distribution of $(z_t - z_s)/(1 + z_{\rm mean})$ for the randomized companions gives us the background level of projected pairs, which are not physically associated. This background is then subtracted from the observed distribution. The physically bound pairs comprise the residual above the background. The characteristic error of the grism redshifts, $\sigma_{\rm girsm}$, is approximated by $\sigma_{\rm Gauss}/\sqrt{2}$, where $\sigma_{\rm Gauss}$ is the width of the Gaussian that is the best fit to the residual distribution.

In Figure 15, we show the distribution of $\Delta z_{\rm grism}/(1 + z_{\rm mean})$ after correcting for projected pairs for each of the five fields in our sample, divided into two bins of redshift: $z < 0.7$ (top) and $z > 0.7$ (bottom). At $z > 0.7$, we find that the grism redshifts produce narrow peaks with $\sigma_{\rm grism} \sim 0.0025$–$0.0035$, consistent with the results from Section 5.3.1. This is a factor of 3 (COSMOS) to 12 (AEGIS) improvement in redshift accuracy over the photometric redshifts in these fields.

At $z < 0.7$, before Hα enters the wavelength coverage of the G141 grism, there are no strong features in the NIR spectra of galaxies.[33] As a result, we would expect that the grism spectra will add very little to the overall fits in constraining the redshift. In fact, as seen in Figure 15, in AEGIS, GOODS-N, and UDS, the pairwise grism accuracy $\sigma_{z<0.7}$ is the same as for the photometric redshifts in those fields. In COSMOS and GOODS-S, the fields with the best photometric redshifts, the addition of the grism improves the redshifts by ~30%. In the following, we limit the analysis of the $z_{\rm grism}$ accuracy to objects with $z > 0.7$.

The accuracy of photometric redshifts depends on the color of the galaxy. As shown in Quadri & Williams (2010) and Bezanson (2016), red galaxies have more accurate photometric redshifts as a result of the strong breaks present in their SEDs, while blue SEDs are relatively featureless. However, the addition of low resolution spectroscopic data can counteract this trend with the addition of emission lines for blue galaxies. In Figure 16, we show the distributions of $\Delta z_{\rm grism}/(1 + z_{\rm mean})$ for galaxies at $z > 0.7$ in the five fields, divided into star-forming and quiescent samples using the Whitaker et al. (2012) *UVJ* selection. The accuracy of the redshift measurements of quiescent galaxies is $\sigma_{\rm grism} \sim 0.006$, which is a factor of several better than the photometric redshift accuracy in most fields. For star-forming galaxies, the characteristic 1σ errors are ~0.002–0.0035, a factor of two lower than those for quiescent galaxies across all fields, independent of the accuracy of the photometric redshifts. As discussed above, based on

[33] There are relatively strong absorption features (the TiO bands), but the weak S III [9068, 9530] doublet lines are typically the only emission features that can be detected.

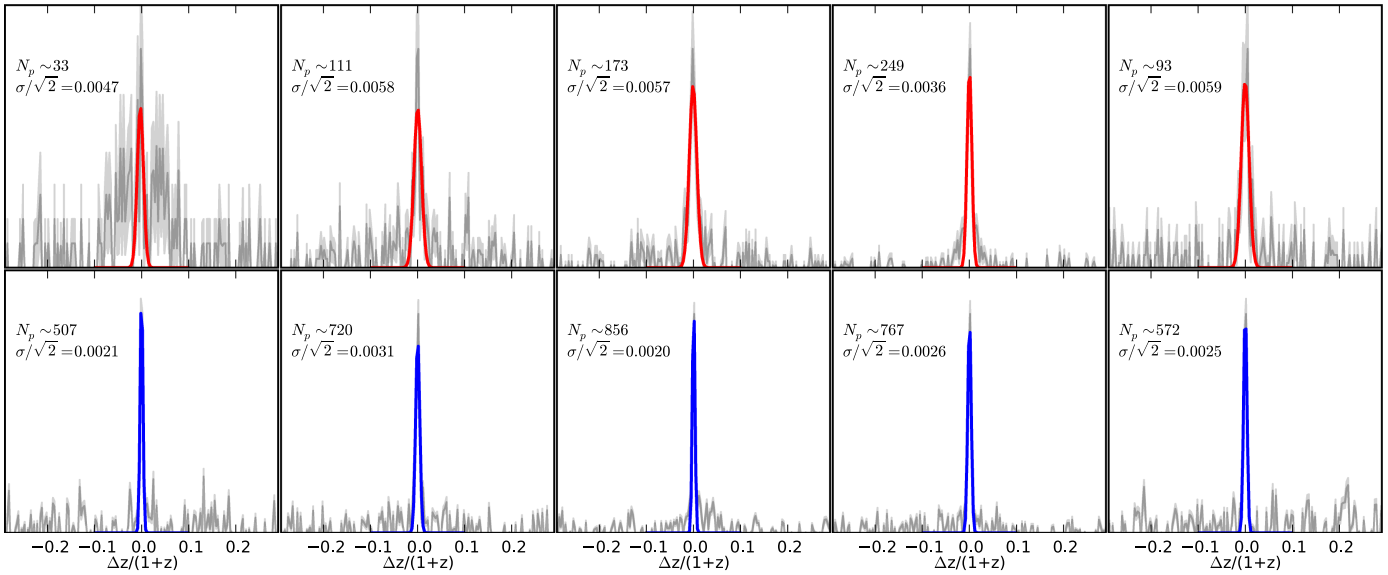

Figure 16. Distribution of $\Delta z_{\rm grism}/(1+z_{\rm mean})$ after correcting for projected pairs for each of the five fields in our sample separated into quiescent (top row) and star-forming (bottom row) by *UVJ* rest-frame color (Whitaker et al. 2012). The tracer and companion galaxies are selected in the same way: we require $z_{\rm grism} > 0.7$, $JH_{\rm IR} < 24$, *use_grism* = 1 and rest-frame color such that it falls within appropriate color selection. The number of pairs that contribute to each measurement is calculated by integrating under the best-fit Gaussian (solid blue/red line). We remove the *y*-axis labels and scale the plots by the maximum histogram value for clarity.

morphological considerations alone, this is the likely limit of the accuracy of grism redshifts for extended sources.

5.3.3. Duplicate Grism Redshifts

The third and final method we use to assess the accuracy of the grism redshifts is repeat measurements. As discussed above, duplicate spectra are not co-added and each 2D spectrum is fit separately in combination with the corresponding photometry. There are a total of 4520 objects, which have at least one repeat observation and 476 objects with more than two observations down to $JH_{\rm IR} = 24$. The maximum number of repeat observations is four for most fields and as many as six in GOODS-S. We limit this analysis to $JH_{\rm IR} < 24$ and we use the visual inspections to remove objects with `use_grism = 0`. These selections leave 4185 redshift measurement pairs. A full list of repeat observations is provided as part of the data release (see below).

The top panels of Figure 17 compare the two duplicate redshift measurements. The agreement is excellent, with very few outliers. Next, we use the duplicate measurements to assess the quality of the redshift uncertainties, that is, whether the differences between two repeat observations can be explained by the formal uncertainties in these measurements. This can be quantified by the ratio between the redshift difference of repeat observations Δz and the total error in Δz, $\sigma_{\Delta z} = \sqrt{\sigma_{z1}^2 + \sigma_{z2}^2}$. The individual redshift errors are asymmetric because the $p(z)$ distributions are typically asymmetric. We define a symmetric error by taking the mean of the lower and upper 1σ errors. We show the distribution of $\Delta z/\sigma_{\Delta z}$ for each of the five fields in the bottom row of Figure 17. If the redshift errors are correct, we would expect that these distributions would have a width of unity. We specifically focus on the $z > 0.7$ regime where the grism spectra likely dominate the redshift estimates. In all fields, we find that the width of the distribution is close to unity. In AEGIS and UDS, the widths are slightly larger, suggesting that errors are underestimated by 5%–20%.

We caution that the two redshift measurements are not entirely independent because the underlying photometric catalogs are the same; it is therefore unsurprising that we see very few "catastrophic" outliers in the top panels of Figure 17. However, the grism spectra themselves are fully independent, and generally come from opposite sides of the detector; they will therefore have a different effect on the fit depending on factors such as the background level, contamination, coverage fraction, etc.

In light of these considerations, we further explore the redshift errors in Figure 18. Here we consider how the accuracy of the redshift errors depends on the properties of the galaxy, in particular, its redshift, magnitude, the error in the redshift ($\sigma/(1+z)$), and the number of pixels that the spectrum covers. In each panel, we see that in the parameter space, where we expect that the grism spectra would contribute most information to the redshift fits (bright magnitude, $z > 0.7$, small error), the scatter in $\Delta z/\sigma$ is ~1–1.2, indicating again that the grism errors underestimate the true errors by at most ~20%. Where the redshifts are dominated by the photometry (at faint magnitude, low redshift, large error) $\Delta z/\sigma < 1.0$, indicating that the photometric data begin to dominate the redshift measurement.

5.3.4. Summary

We find that all methods for determining the redshift accuracy are in good agreement: the comparison to spectroscopic redshifts, the pair analysis, and the analysis of duplicates all imply typical redshift uncertainties of $\approx 0.003 \times (1+z)$. Quiescent galaxies have larger redshift uncertainties than star-forming galaxies, but even the errors for quiescent galaxies (typically $0.006 \times (1+z)$) are well below those from the photometry alone. Most importantly, the duplicate analysis shows that the formal redshift uncertainties are generally very good (within 10%–20% of the actual error), independent of magnitude.

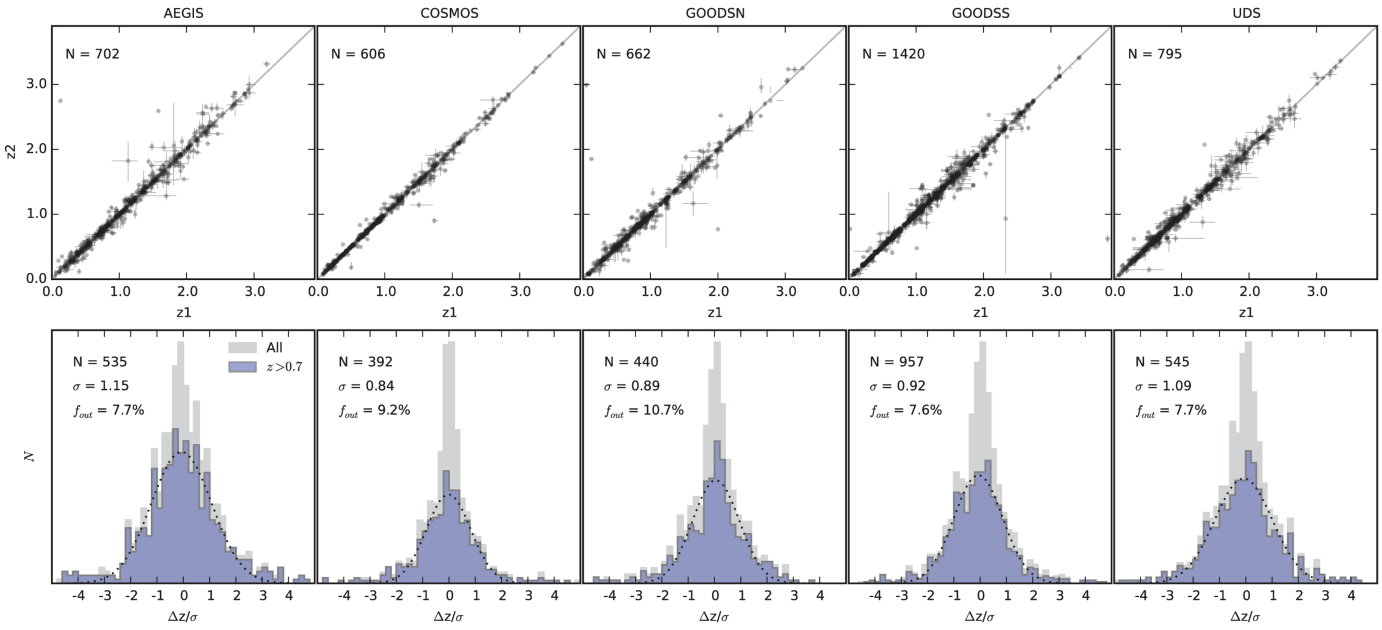

Figure 17. Top: comparison between duplicate redshift measurements. Bottom: distributions of $\Delta z/\sigma$, where Δz is the redshift difference between two duplicate measurements and σ is the total error in Δz. We show both the distribution for all objects (gray histogram) and the distribution for $z > 0.7$ objects (blue histogram). The number of $z > 0.7$ objects for each field is listed in the upper left corner of the panel, along with the best-fit Gaussian σ for the distribution. Overall, these distributions indicate that the redshift errors are accurate and possibly underestimated by 5%–20%.

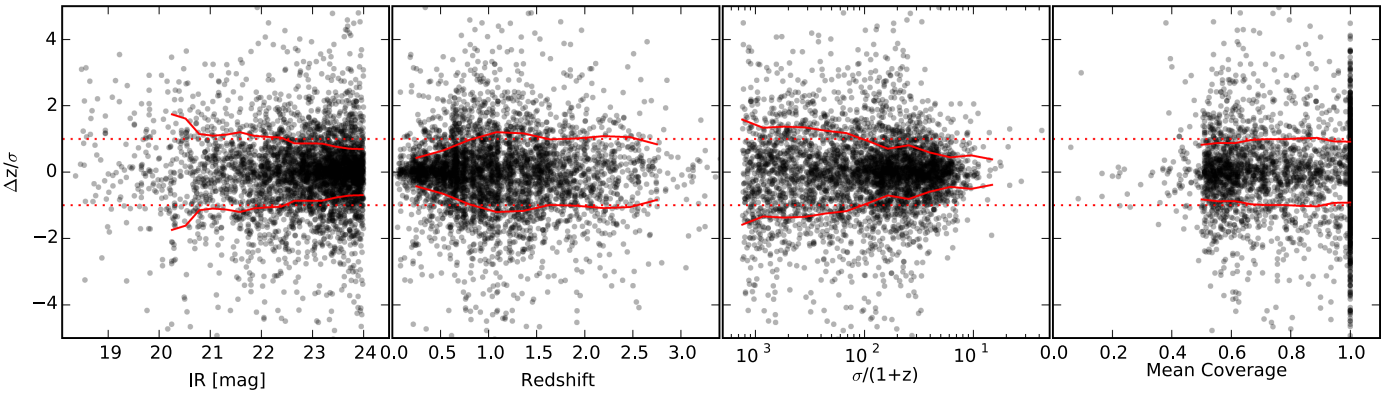

Figure 18. $\Delta z/\sigma$ as a function of $JH_{\rm IR}$ magnitude, redshift, redshift error $\sigma/(1+z)$, and mean spectral coverage for all duplicate pairs in 3D-HST. If the redshift errors are correct, we would expect that the widths of these distributions would be unity (dotted lines). The solid lines indicate that the sliding box NMAD scatter is in each panel.

6. EMISSION LINE FITS

6.1. Methodology

Following the approach of the redshift fits, we also fit the emission line fluxes directly in the 2D spectrum. With the redshift fixed to $z_{\rm grism}$, we generate a 2D model spectrum for each of the emission lines listed in Table 4 that would fall within the grism passband (with unresolved line widths $\sigma = 100$ km s^{-1}), and we adopt the 2D continuum template determined from the earlier redshift fit (Section 5). With parameters for the individual template normalizations, we use the `emcee` sampler (Foreman-Mackey et al. 2013) to determine the marginalized posterior distribution functions of parameters for the individual template normalizations, which can be converted directly into line fluxes and observed-frame equivalent widths in physical units (i.e., erg s^{-1} cm^{-2} and Å, respectively). As with the redshift fit in Equation (4), this method provides the benefit of fitting in the natural units of instrumental count rates and preserves per-pixel uncertainties.

The Markov chain Monte Carlo (MCMC) fit sequences (a.k.a. "chains") provide a robust estimate of the uncertainties in the fit, which are primarily determined by the wavelength dependence of the grism throughput and by the object size (i.e., the area of the effective aperture of the 2D spectrum fit). The dependence of the derived uncertainties provided in the emission line catalog on these two characteristics is shown in Figure 19: the G141 grism is somewhat more sensitive at red wavelengths and line sensitivity rapidly decreases for large

Table 4
Emission Lines

Line	Catalog ID	Rest Wavelength [Å]	Ratio
Lyα	Lya	1215.400	···
C IV	CIV	1549.480	···
Mg II	MgII	2799.117	···
Ne V	NeV	3346.800	···
Ne VI	NeVI	3426.850	···
[O II]	OII	3729.875	···
[Ne III]	NeIII	3869.000	···
He I	HeIb	3889.500	···
Hδ	Hd	4102.892	···
Hγ	Hg	4341.680	···
[O III]	OIIIx	4364.436	···
He II	HeII	4687.500	···
Hβ	Hb	4862.680	···
[O III]	OIII	5008.240, 4960.295	2.98:1
He I	HeI	5877.200	···
[O I]	OI	6302.046	···
Hα	Ha	6564.610	···
[S II]	SII	6718.290, 6732.670	1:1
S III	SIII	9068.600, 9530.600	1:2.44

extended galaxies. Overall, the line sensitivity can be parameterized by

$$1\sigma = 8 \times 10^{-18}\left(\frac{G(\lambda)}{G(1.5\ \mu\mathrm{m})}\right)^{-2}\left(\frac{R}{5\ \mathrm{pix}}\right)\ \mathrm{erg\ s^{-1}\ cm^{-2}},$$

$G(\lambda)$ is the wavelength dependent throughput of the G141 grism[34] and R is the SExtractor FLUX_RADIUS in pixels. Using Figure 19, we can determine the emission line point source sensitivity. For a point source (4 pixels) the 1σ limiting flux is $0.7 \times 10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ at 1.5 μm, therefore, a 3σ detection will have a flux of $2.1 \times 10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$.

The line fluxes are implicitly normalized to the broadband photometry of Skelton et al. (2014), as the spectra are scaled to match the photometric data. The fluxes are therefore "total," and do not refer to a particular aperture, though an implicit assumption is that the equivalent widths of the lines do not increase or decrease strongly outside of the segmentation map. No absorption corrections are necessary, and in that sense our methodology is different from most measurements in the literature (e.g., Steidel et al. 2014; Kriek et al. 2015). The standard method is to measure the flux of a bright line with respect to an idealized continuum, parameterized by a constant or a linear function defined in a narrow wavelength region to the blue and red of the line. For Hα, Hβ, and other Balmer lines a correction needs to be made after the measurement, to account for absorption in the stellar continuum. In our methodology, the stellar continuum model is not a low order polynomial but the best-fitting stellar population synthesis model that came out of the redshift fit. It therefore uses all the information in the broadband photometry and the grism spectrum. No post-measurement absorption corrections are necessary, as the Balmer absorption lines are present in the model, at the appropriate resolution.

[34] The G141 throughput curve can be obtained with PySynphot: http://ssb.stsci.edu/pysynphot/docs/.

6.2. Results

Emission lines are fit for all objects down to $JH_{\mathrm{IR}} = 26$; however, the analysis in this section is limited to $JH_{\mathrm{IR}} = 24$, where we take advantage of the grism redshift use flags. If none of the emission lines in Table 4 fall within the grism wavelength range (for the best-fit redshift), an emission line fit is not produced. We test the accuracy and precision of the emission line flux measurements by comparing the fits for duplicate objects within the survey and by comparing our measurements to those from ground-based surveys.

We begin by comparing the emission line fluxes measured from repeated observations of the same object. Unlike the redshift fits, the emission line measurements are based on the 2D grism spectra alone and they are, therefore, truly independent measurements. Repeat spectra are typically taken at different angles and the spectra fall on different parts of the detector. Mismatch of repeat line flux measurements can indicate problems with the background subtraction, the flat-fielding, and a number of other effects. In Figure 20, we show the flux measurements from objects with multiple grism spectra. The measurements follow the 1:1 line tightly, with the scatter and errors increasing with decreasing flux (left panel). In the right panel of Figure 20, we analyze the errors of the line fluxes in a manner similar to the analysis of the redshift errors in Section 5 and Figure 18. The sliding box NMAD scatter, σ_{NMAD} (solid black line), is approximately unity across all fluxes, demonstrating that the formal errors are an excellent approximation of the actual uncertainties. The emission line flux errors are calculated on the basis of the interlaced G141 image background errors, which include terms for the contamination subtraction. The fact that $\sigma_{\mathrm{NMAD}} \sim 1$ shows that these errors properly account for the line flux uncertainties and that there are no systematic errors introduced in our data reduction. The outlier fraction in Figure 20 is strikingly small. Even though we apply no quality flags to select spectra with clean emission lines other than the line $S/N > 2$ cut and the redshift use flag, the fraction of objects with $|\Delta F_\lambda/\sigma| > 3$ is only 3.9%.

An external check on the emission line fluxes, shown in Figure 21, is provided by a comparison between our measurements and those from MOSDEF and KMOS^{3D}, as well as from the SINS/zC-SINF survey with VLT/SINFONI (Förster Schreiber et al. 2009; Mancini et al. 2011; Newman et al. 2014). Compared to the 3D-HST spectra, ground-based spectra are affected by (rapidly varying) atmospheric emission and absorption. While the integral field unit (IFU) data from KMOS^{3D} and SINS/zC-SINF allows one to recover well the full 2D spatial emission, slit losses may affect the multi-slit spectra from MOSDEF. Significant effort has been made by the survey teams to correct the ground-based line fluxes for such losses. Close pairs of lines in the ground-based spectra are co-added to compare to the lower-resolution 3D-HST measurements. The grism and ground-based line fluxes match well and follow the 1:1 line overall (left panel) with a scatter of a factor of 1.8 for lines brighter than $10^{-16}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ in both data sets (right panel). This agreement can be considered good when considering the large differences in observing and analysis methods; attributing an equal uncertainty to both data sets, the per-measurement error is a factor of 1.5.

At low line fluxes, there may be a systematic effect, such that the 3D-HST line measurements are slightly higher than the ground-based ones. This could be due to the effects of the large

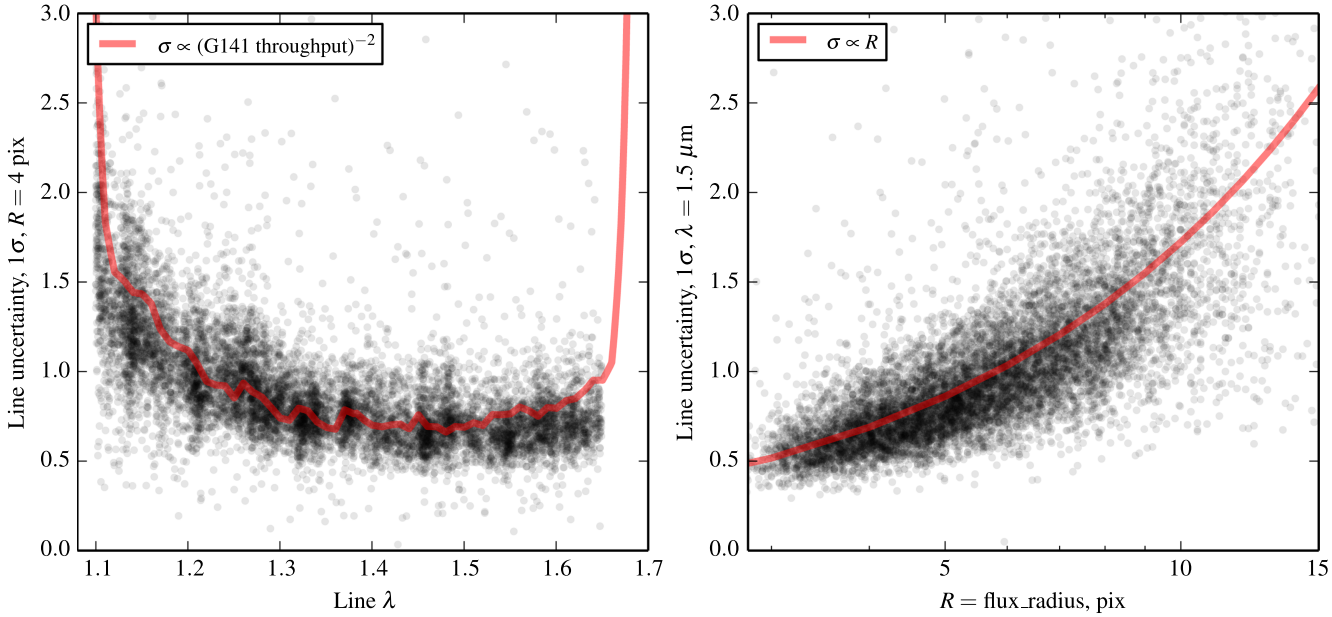

Figure 19. Emission line sensitivity determined from the individual MCMC line fits as a function of wavelength (left) and object size R parameterized by the SExtractor "FLUX_RADIUS" value from the catalog (right). The values on the vertical axes are given in units of 10^{-17} erg s^{-1} cm^{-2}. In both panels, the gross trends caused by the effect in the opposite panel have been divided out. The red curves are normalized to the data, but are not fits to the data points. They show that these trends are as we expect: the line uncertainty is inversely proportional to the grism throughput (left) and proportional to the size of the object (right). For a typical resolved galaxy ($R = 5$ pix) a line at 1.5 μm will have a 1σ line uncertainty of 8×10^{-18} erg s^{-1} cm^{-2} in the 2-orbit 3D-HST G141 grism spectra.

errors in this regime: these objects were selected in 3D-HST and subsequently observed from the ground, which may produce the asymmetry. The effect could also be due to uncertainties in aperture corrections, the fact that our line measurements are corrected for the underlying stellar absorption, or other effects.

7. CATALOGS

The results from the redshifts and emission line fits are assembled into several different catalogs. For the majority of users, these catalogs will probably be the main, or only, gateway to the 3D-HST data set. In this section, we describe the catalogs produced from the survey and the applications for which each of them may be appropriate.

7.1. Redshift and Emission Line Catalog

The first type of catalogs we produce are simply concatenations of the outputs of all redshift and emission line fits. These catalogs contain repeat fits for the same photometric object. The fits are done for each extracted 2D spectrum of each object separately (in conjunction with the photometric information) for a total of 98,668 individual spectra down to $JH_{\rm IR} = 26$ (except for the UDS field where the fits reach approximately 0.5 mag fainter). In these catalogs, each row corresponds to the outputs from a single spectrum. Each spectrum has a unique identifier of the format `aegis-01-G141_00001`, listing the field, the pointing number (zero-padded two-digit integer), the grism name (G141 for 3D-HST), and the photometric identification number of the object (padded five-digit integer). The objects are ordered by pointing number and, within that pointing, by photometric identifier. A list of all duplicate spectra is also provided (see Section 7.2).

Two concatenated catalogs are produced, one containing the redshift fits and one containing the emission line fits. Both catalogs have the same length. The column names and the corresponding descriptions are listed in Tables 5 and 6. We produce catalogs for each field separately as well as a master catalog, which contains all objects in the survey.

The concatenated catalogs provide information for *all* objects fitted as part of the current release. The $JH_{\rm IR}$ magnitude is included as a column in the catalog; however, we do not preselect objects in any way for this catalog and we specifically do not exclude duplicate observations. These catalogs can be used to identify all the information available for a given object in the photometric catalogs.

7.2. Line-matched Catalogs

We also produce redshift and emission line catalogs that are matched to the photometric catalogs of Skelton et al. (2014). These catalogs, one for each field, as well as a master catalog containing all fields, have the same length as those in the v4.1 photometric release with one entry per object from Skelton et al. (2014). The column names and the corresponding descriptions are listed in Tables 5 and 6. The rows corresponding to objects in the photometric catalog that do not have grism spectra are set to default values. Duplicate objects appear only once in these catalogs; the selection of the primary object among duplicates is described in Section 5.2.

The line-matched catalog contains a total of 79,609 unique objects with fits or 38.2% of the photometric catalog. Of these, 22,548 objects have magnitudes brighter than $JH_{\rm IR} = 24$ and it is only these $JH_{\rm IR} < 24$ objects that have been visually inspected, and have a `use_grism` flag assigned as described in Section 5.2. The bright $JH_{\rm IR} < 24$ objects constitute 10.8% of the objects in the photometric catalog. We caution against blindly using our redshifts and emission line fits for objects with $24 < JH_{\rm IR} < 26$. Even though checks of faint subsamples have allowed us to verify that our methods are reliable

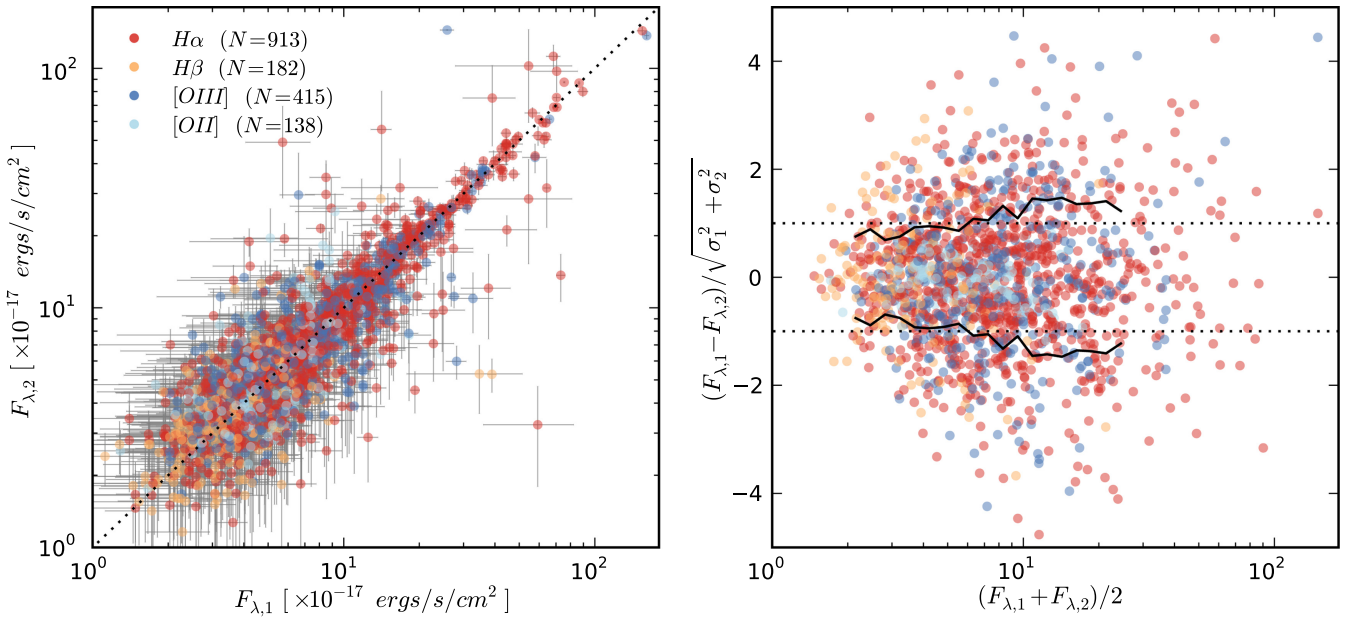

Figure 20. Left: comparison of line flux measurements for objects with multiple grism spectra. Points are color-coded by the emission line with the number of measurements for each line listed in the legend. We require $JH_{\rm IR} < 24$, $S/N > 2$ in the line in both spectra and apply the redshift use flags. Error bars are 1σ. Right: $\Delta F_\lambda/\sigma$ for the emission line flux measurements as a function of mean line flux. The sliding box NMAD scatter as a function of flux (solid black line) is ∼1 at all fluxes, indicating that the line flux errors are correct.

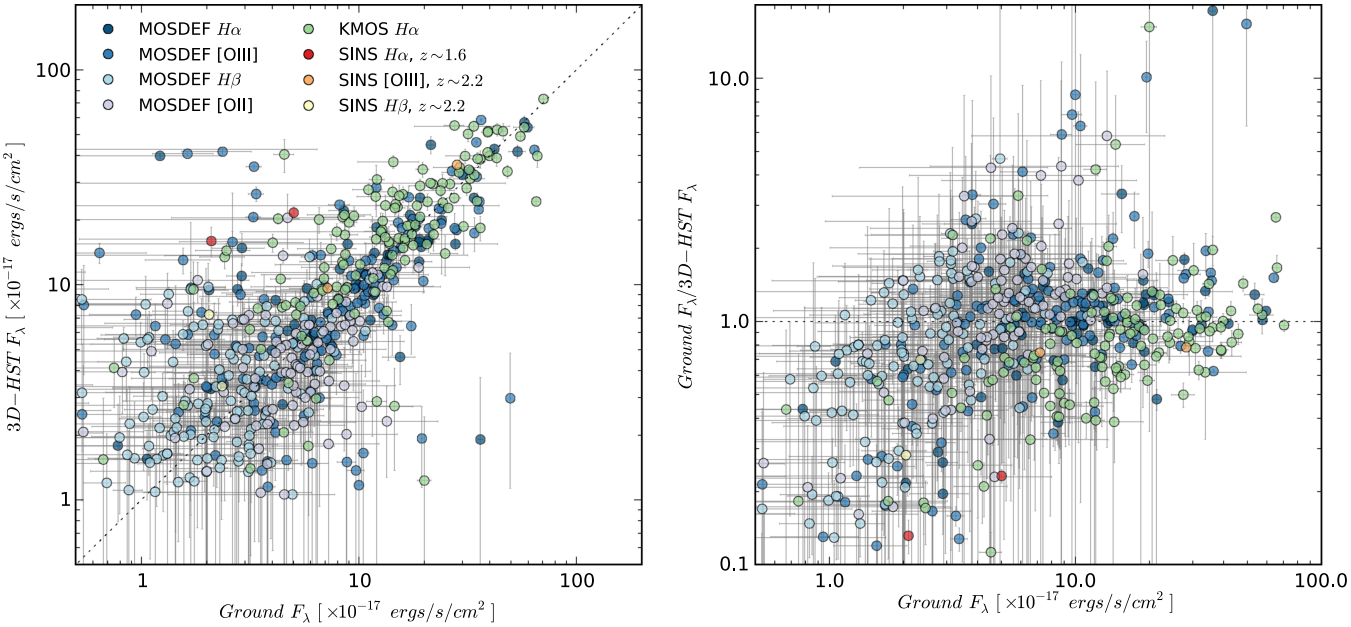

Figure 21. Emission line flux measurements from 3D-HST compared to those from KMOS^{3D}, MOSDEF, and SINS/zC-SINF. For the comparison, Hα and N II from the ground-based observations are co-added, as are the [O II] and [O III] doublets. Left: ground-based flux vs. 3D-HST flux for emission lines in common between the surveys. Right: ratio between the ground-based and 3D-HST grism fluxes. The scatter for bright lines ($>10^{-16}$) is a factor of 1.8, and the mean ratio is 0.9.

in this parameter space, the vast majority of these spectra have not been inspected.

In addition to the redshift and emission line catalogs, we create a row-matched listing of all duplicate spectra of a given object. We also make available the SExtractor catalog with $JH_{\rm IR}$ fluxes measured from the $J_{125} + JH_{140} + H_{160}$ images.

Using the grism redshift fits in the line-matched catalogs, we refit the stellar population parameters, rest-frame colors, and SFRs as described in Skelton et al. (2014) and Whitaker et al. (2014). The outputs from these fits are made available as part of the release.

7.3. "Best" Catalog

Finally, we create a "best" redshift catalog, by merging the grism redshift fits with the photometric redshifts from Skelton et al. (2014). The best redshift is

Table 5
Redshift Catalog Columns

Column Name	Default	Description
phot_id	⋯	Unique identifier from Skelton et al. (2014)
spec_id	00000	Unique identifier for the spectrum that was used in this measurement
jh_mag	⋯	SExtractor MAG_AUTO JH_{IR} magnitude of the objects, described in Section 3.6
z_spec	−1	Spectroscopic redshift, when available, see Skelton et al. (2014) for sources and quality
z_peak_phot	−1	Photometric redshift: same as z_peakfrom the EAZY catalogs of Skelton et al. (2014)
z_phot_l95	−1	Photometric redshift at the lower 95% confidence limit
z_phot_l68	−1	Photometric redshift at the lower 68% confidence limit
z_phot_u68	−1	Photometric redshift at the upper 68% confidence limit
z_phot_u95	−1	Photometric redshift at the upper 95% confidence limit
z_max_grism	−1	The redshift where the $p(z \mid \text{grism, phot})$ is maximized should be used as default grism redshift
z_peak_grism	−1	Integral of $p(z \mid \text{grism, phot})^* z^* dz$, integrated over the whole redshift range
z_grism_l95	−1	Grism redshift at the lower 95% confidence limit
z_grism_l68	−1	Grism redshift at the lower 68% confidence limit
z_grism_u68	−1	Grism redshift at the upper 68% confidence limit
z_grism_l95	−1	Grism redshift at the upper 95% confidence limit
f_cover	−1	Fraction of spectrum within the image (0 = bad, 1 = good)
f_flagged	−1	Fraction of flagged pixels (1 = bad, 0 = good)
max_contam	−1	Maximum contamination
int_contam	−1	Contamination integrated over the spectrum (= flux_contam/flux_object)
f_negative	−1	Fraction of pixels with negative flux after contamination correction (if big could indicate a problem with the contamination correction)
flag1	−1	User assigned flag for the redshift quality
flag2	−1	User assigned flag for the redshift quality
use_grism[a]	−1	Flag defining objects with the most reliable grism-derived redshifts (see Section 5.2)
use_phot[a]	⋯	Photometric use flag from Skelton et al. (2014): 1 = use; 0 = do not use
z_best_s[a]	⋯	Source of the best redshift: 1 = ground-based spectroscopy; 2 = grism; 3 = photometry; 0 = star
z_best_best[a]	−1	Best available redshift measurement (−1 for stars)
z_best_l95[a]	−1	Lower 95% confidence limit derived form the z_best $p(z)$
z_best_l68[a]	−1	Lower 68% confidence limit derived form the z_best $p(z)$
z_best_u68[a]	−1	Upper 68% confidence limit derived form the z_best $p(z)$
z_best_u95[a]	−1	Upper 95% confidence limit derived form the z_best $p(z)$

Note.
[a] This column is only present in the line-matched catalogs.

Table 6
Emission Line Catalog Columns

Column Name	Default	Description
number	⋯	Unique identifier from Skelton et al. (2014)
gris_id	00000	Unique identifier for the spectrum that was used in this measurement
jh_mag	⋯	SExtractor MAG_AUTO JH_{IR} magnitude of the objects, described in the text
z	−1	Grism redshift used in the emission line fit, identical to z_max_grism in the redshift catalog
s0	−99	Normalization coefficient s0, see description in the text
s0_err	−99	Error for normalization coefficient s0
s1	−99	Normalization coefficient s1, see description in the text
s1_err	−99	Error for normalization coefficient s1
X_FLUX	−99	Emission line flux in units of 10^{-17} erg s^{-1} cm^{-2}
X_ERR	−99	Error in the emission line flux in units of 10^{-17} erg s^{-1} cm^{-2}
X_SCALE	−99	Multiplicative scaling factor to correct the flux of the emission line to the photometry
X_EQW	−99	Emission line equivalent width in Å

Note. X = emission line name, as given in Table 4.

1. z_spec if it exists from the Skelton et al. (2014) compilation of spectroscopic redshifts.
2. z_max_grism if there is no spectroscopic redshift and use_grism = 1.
3. z_phot if there is no spectroscopic redshift and use_grism >1.

We emphasize that we only use the photometric redshift if there is no grism spectrum that can be used (either because an object was not observed or because the spectrum has problems, as detailed above). Even if a grism spectrum appears to contain only noise, we use it in the fit; as discussed earlier, the error weighting in the fitting procedure ensures that the resulting redshift is nearly

Table 7
Best Redshift Catalog

Column Name	Description
field	Field identifier (aegis/cosmos/goodsn/goodss/uds)
phot_id	Unique identifier from Skelton et al. (2014)
z_best_s	Source of the best redshift: 1 = ground-based spectroscopy; 2 = grism; 3 = photometry; 0 = star
use_phot	Photometric use flag from Skelton et al. (2014): 1 = use; 0 = do not use
use_grism	Grism use flag as defined in Section 5.2
z_best	Best available redshift measurement (−1 for stars)
z_l95	Lower 95% confidence limit derived form the z_best $p(z)$
z_l68	Lower 68% confidence limit derived form the z_best $p(z)$
z_u68	Upper 68% confidence limit derived form the z_best $p(z)$
z_u95	Upper 95% confidence limit derived form the z_best $p(z)$

completely determined by the photometry in such cases. Using the best redshifts, we also create merged catalogs of the stellar population parameters, rest-frame colors and SFRs.

8. PROPERTIES OF THE 3D-HST DATA PRODUCTS

Here we briefly summarize what 3D-HST contributes to existing data sets and catalogs that are based on deep, "blank" fields. The immediate contributions of the grism spectroscopy are a uniform, complete redshift catalog with relatively small and well understood uncertainties; emission line fluxes; and 2D emission line maps. Furthermore, the combination of these data with stellar masses determined from SED fits, UV+IR SFRs, and WFC3 morphologies constitutes the most complete data set to date for studies of "normal" galaxies out to $z \sim 3$.

8.1. Redshifts and Redshift Distribution

The accuracy of the redshifts (determined in detail in Section 5.3) is $\approx 0.003 \times (1 + z)$ for most galaxies, with some dependencies on magnitude and rest-frame color (mostly reflecting an underlying dependency on whether a bright emission line is in the observed wavelength range). Crucially, the formal uncertainty in the redshift is generally an excellent measure of the actual error (see Figures 17 and 18). The error corresponds to a velocity uncertainty of $\sim 1000\ \mathrm{km\ s^{-1}}$.

The redshift accuracy that is achieved makes it possible to identify overdensities and characterize the environment of galaxies, with much better contrast than with photometric redshifts alone. This is illustrated in Figure 22, which shows the spatial distribution of galaxies in the UDS in small redshift bins between $z = 1.07$ and $z = 1.11$. The left panels show smoothed fifth nearest neighbor density maps based on photometric redshifts (top) and grism redshifts (bottom), and the right panels show the corresponding redshift histograms. The structure at $z = 1.09$ is clearly defined as a narrow grism redshift peak, but is spread out in the photometric redshifts.

Figure 23 shows the redshift distributions based on all catalogs presented in this paper. The distribution shows a broad peak between $z = 1$ and $z = 2$, due to a combination of the observed-frame magnitude limits, the luminosity function of galaxies, and volume effects. For a particular magnitude limit the distributions of grism redshifts (red or pink) is always below that of photometric redshifts (black or gray), due to the fact that not all objects have a usable grism spectrum. The grism redshift distribution for $JH_{\mathrm{IR}} < 24$ shows more pronounced peaks than the photometric redshift distribution; this is because physically associated galaxies in groups and clusters have more accurate redshifts in the grism catalog. The same behavior is seen in the fainter sample with $JH_{\mathrm{IR}} < 26$, but we emphasize that the grism redshifts for these faint objects were not inspected. Strikingly, the pronounced photometric redshift peak at $z \sim 1.6$ in both the bright and the faint sample is not visible in the grism redshift distribution. This should be regarded as a success of our methodology: this peak is a well-known (but not well understood) artifact in photometric redshift measurements (see, e.g., Brammer et al. 2008; Skelton et al. 2014).

The differences between photometric redshifts and grism redshifts are shown explicitly in Figure 24. Horizontal features in this figure are overdensities that are more clearly identified in the grism redshift distribution. Vertical features indicate "attractors" in photometric redshift; the most prominent of these is the broad peak at $z \sim 1.6$. Note that galaxies with $z_{\mathrm{phot}} = z_{\mathrm{grism}}$ do not necessarily have highly accurate photometric redshifts; these can also be cases where the grism spectrum does not add significant information to the fit and both redshifts are essentially determined by the photometry alone. The accuracy of photometric redshifts is discussed in Bezanson (2016).

8.2. Spectral Features

As discussed in Section 6, the catalogs contain flux and EW measurements, with well calibrated uncertainties, for every emission line of Table 4 that falls in the observed wavelength range for a particular object. The emission lines are measured for every extracted spectrum down to $JH_{\mathrm{IR}} = 26$, but we only supply use flags for galaxies with $JH_{\mathrm{IR}} < 24$. Several papers and projects have used early versions of these catalogs; as an example, both the MOSDEF (Kriek et al. 2015) and KMOS^{3D} (Wisnioski et al. 2015) surveys have used 3D-HST line measurements to select objects for follow-up spectroscopy with ground-based spectrographs.

A comprehensive study of the line fluxes is beyond the scope of the present paper; the Hα emission line luminosities and stellar absorption features are analyzed in Fumagalli et al. (2012), Fumagalli (2015), and Whitaker et al. (2013) respectively. Here we illustrate the relation between the strengths of various emission lines and other galaxy properties in a series of 2D stacks. These stacks are created by ordering the G141 spectra by a particular parameter, such as redshift or stellar mass. Then a 2D surface is generated with (observed or rest-frame) wavelength on the x-axis and the sorting parameter on the y-axis.[35] Rather than showing all spectra, we bin them in small intervals of the sorting parameter, such that each line corresponds to the median of many spectra.

Figure 25 shows the "basic" 2D stack where redshift is the sorting parameter and the horizontal axis is rest-frame wavelength. The selection is $0.15 < z < 3.3$ and $H_{160} < 25$, with some additional constraints on the quality of the spectra. Each line is the median of 100 individual 1D spectra. Each of the 100 spectra was normalized by the object's JH_{140} flux prior

[35] See http://www.sdss.org/science/ for an example of such a 2D stack of 46,420 SDSS quasars, created by X. Fan.

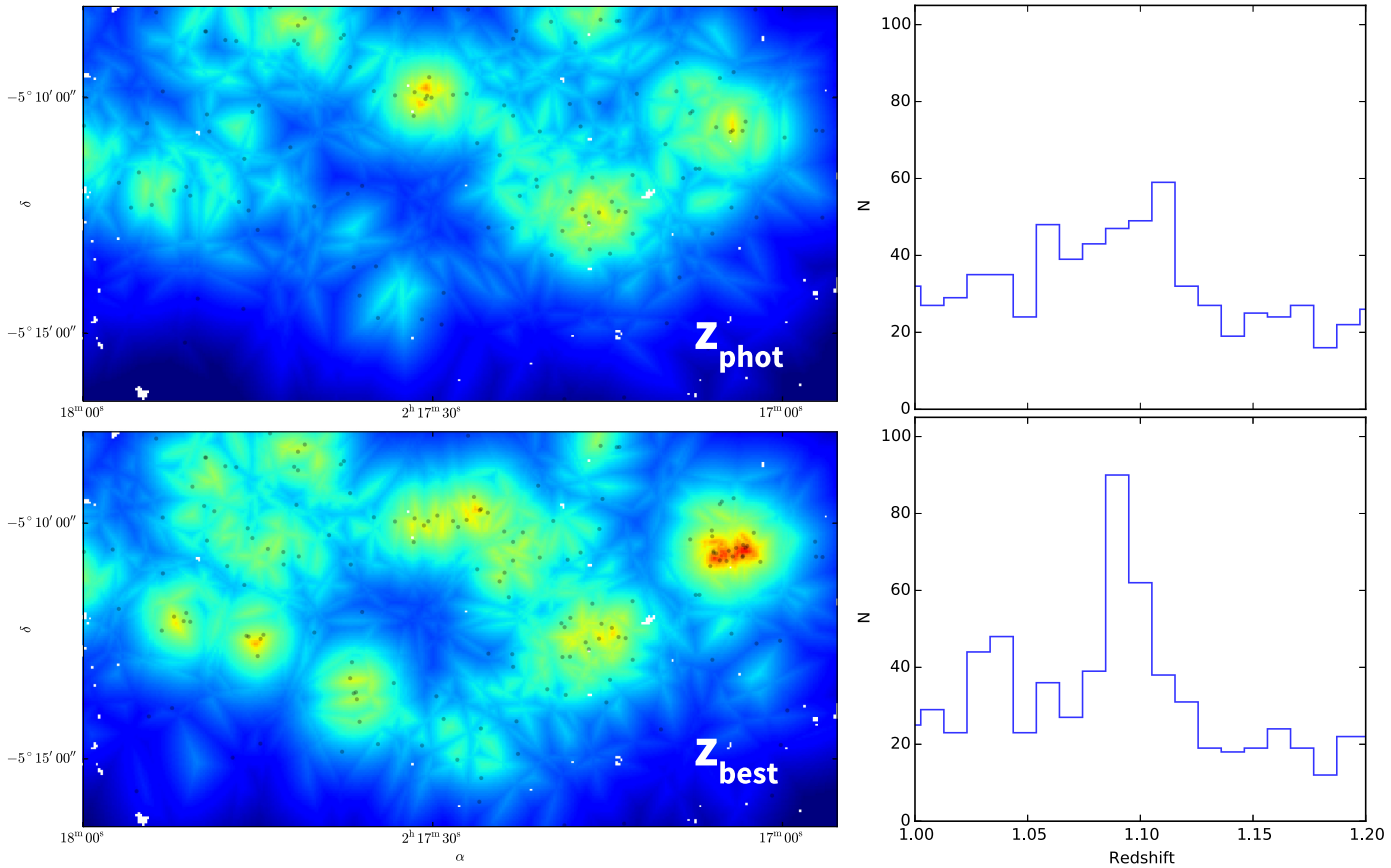

Figure 22. Example of the ability of the grism redshifts to identify overdensities and characterize the environment of galaxies. The panels show the distribution of galaxies in the UDS field, in a narrow redshift bin between 1.07 and 1.11. Left: smoothed fifth nearest neighbor maps using the `z_phot` and `z_best` redshifts for the $JH_{\rm IR} < 24.$ sample. Right: redshift histograms. The overdensity at $z = 1.09$ is clearly defined in the grism redshift distribution.

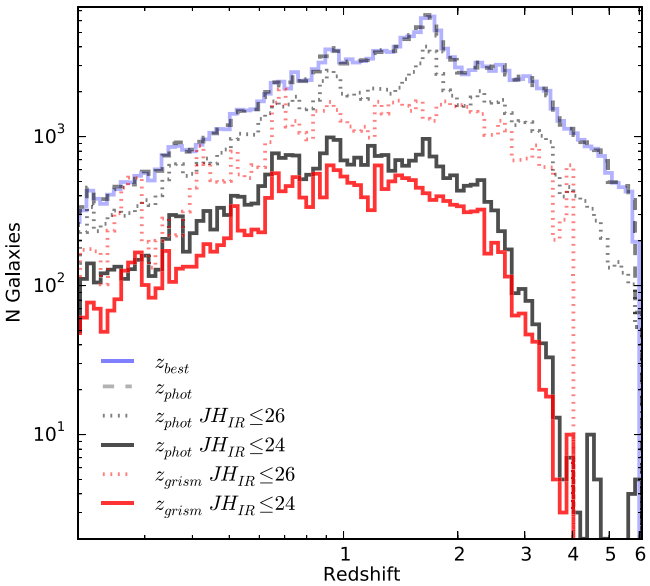

Figure 23. Redshift distributions of the catalogs in this paper. Distributions that are derived from the full photometric + grism fits are shown in red/pink. Distributions that are based on the photometry only are shown in black/gray. The grism data produces more pronounced peaks in the redshift distributions, as expected. Note that the (spurious) broad peak at $z \sim 1.6$ in the photometric redshift distributions is not present in the grism redshift distributions.

to taking the median. Therefore, the intensity of emission and absorption lines in Figure 25 (and subsequent figures) corresponds approximately to their equivalent widths, and not to their fluxes or luminosities. Redshift is shown on the left vertical axis; the cumulative number of spectra is shown on the right axis. Figure 26 shows the same spectra as Figure 25, but ordered by their photometric redshift rather than grism redshift. The differences between this figure and Figure 25 is a qualitative demonstration of the improved redshift accuracy that is enabled by the grism spectroscopy (see also Bezanson 2016).

In Figure 27, galaxies are split by their emission line properties. Galaxies in the left panel have $z > 0.605$, $H_{160} < 24$, and the S/N of [O II], [O III], and/or Hα greater than three. In the right panel, galaxies have $z > 0.605$, $H_{160} < 23$, and an S/N of all lines <2. Several well-known emission lines can readily be identified in the left panel of Figure 27 as they shift into and out of the observed wavelength range of the G141 grism. At $z \gtrsim 0.7$, the prominent Hα line and the S II [6718, 6733] doublet are visible. Over the redshift range $0.7 < z < 1.5$ the Hα equivalent width gradually increases, as discussed by Fumagalli et al. (2012). At $z \gtrsim 1.3$, Hβ and the [O III] [4959, 5007] doublet fall in the observed wavelength range. Again, we see that the equivalent widths of emission lines (in this case [O III]) increase with redshift in broadband flux-limited samples. The small redshift range where Hα and Hβ are both covered by the G141 grism was utilized by, e.g., Price et al. (2014), who study the reddening of H II regions as measured by the Balmer decrement in 3D-HST galaxies. Finally, at $z \gtrsim 1.9$ the [O II] doublet enters the grism wavelength range.

Absorption features (right panel of Figure 27) are not measured automatically in our analysis. Such measurements are highly dependent on the spectral resolution and the precise

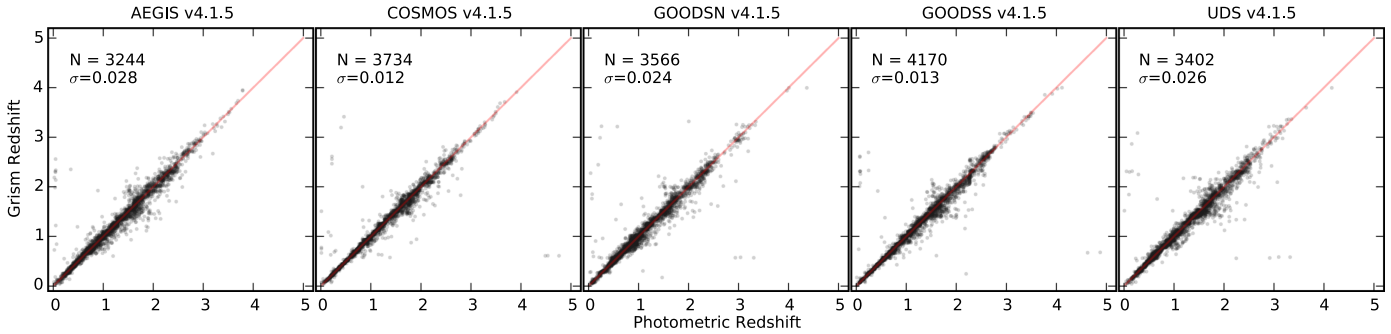

Figure 24. Comparison between grism and photometric redshifts. Horizontal features are real structures in redshift space. Vertical features indicate spurious "attractors" in photometric redshift determinations.

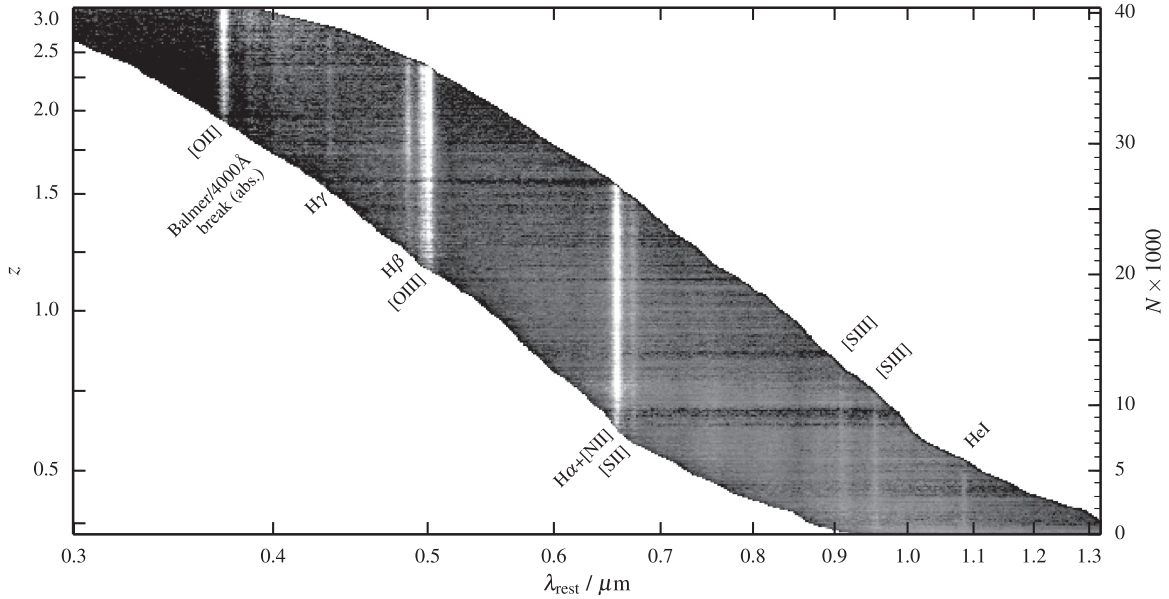

Figure 25. Overview of ∼40,000 3D-HST G141 grism spectra with $H_{160} < 25$. Each pixel row shown is the median of 100 individual 1D spectra sorted by redshift and shifted to the rest frame; ticks on the right axis mark every 1000 galaxies, and tick labels on the left axis indicate the corresponding redshift. Each spectrum is normalized by the object's JH_{140} flux. Absorption and emission lines that move through the G141 passband at different redshifts are indicated.

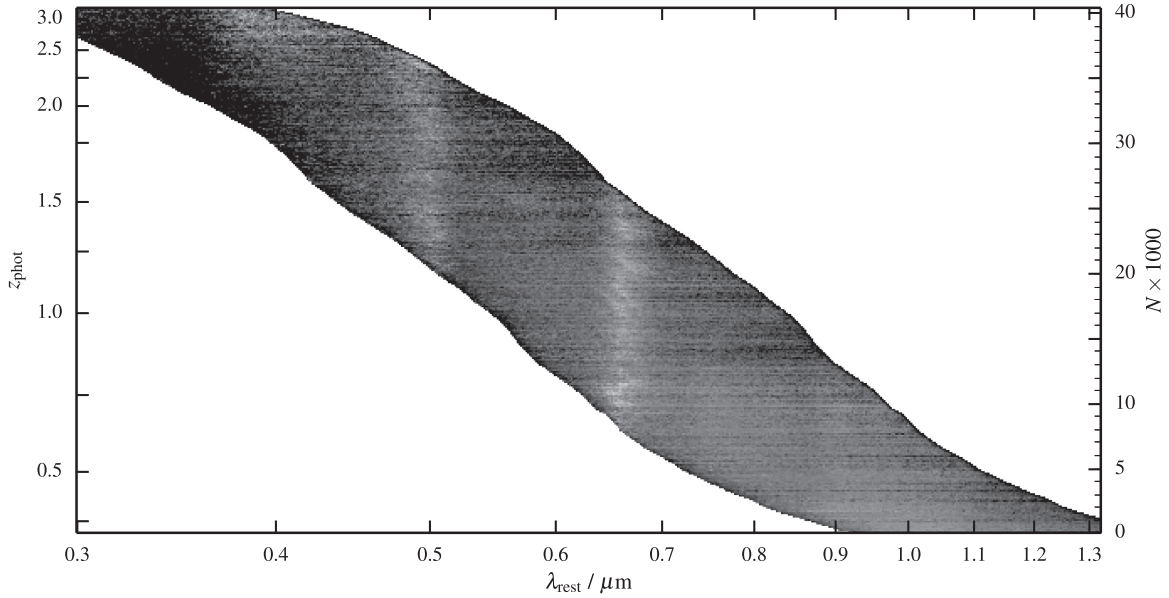

Figure 26. Same as Figure 25, but using photometric redshifts rather than grism redshifts to order the spectra. The differences between this figure and Figure 25 graphically illustrate the improvement in the redshift accuracy when going from photometric redshifts to grism redshifts.

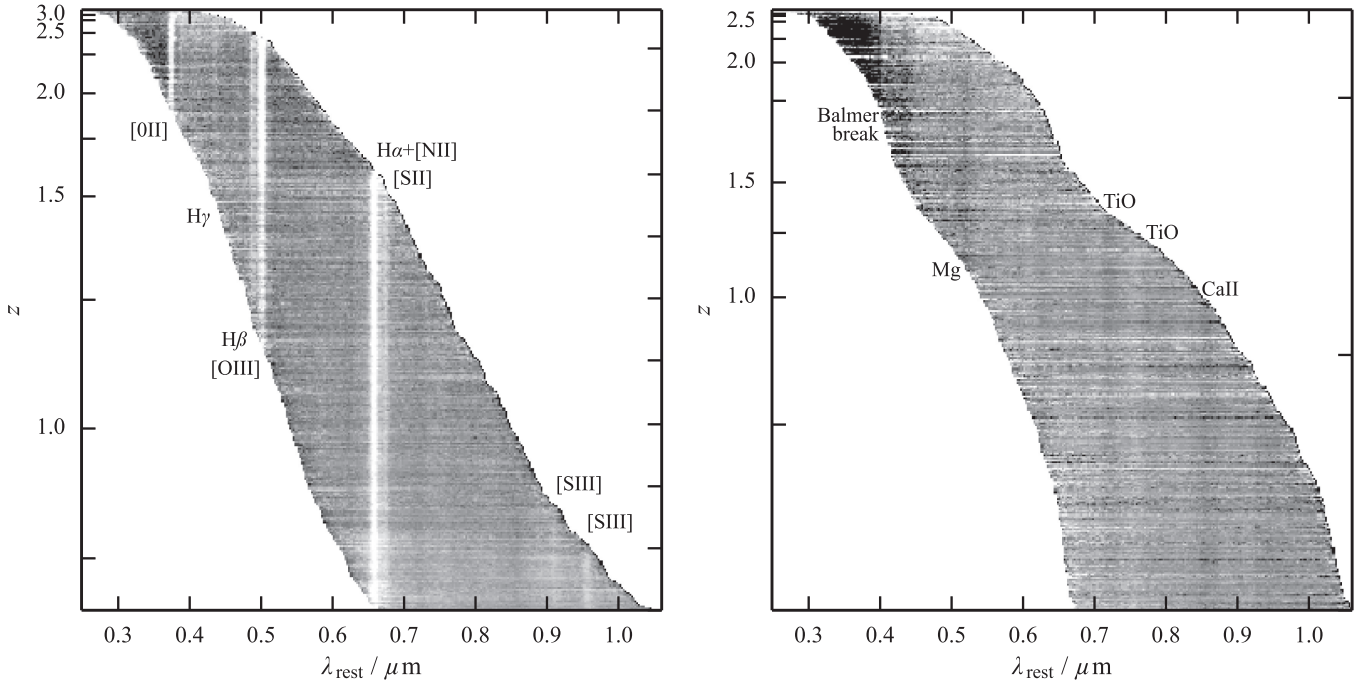

Figure 27. Same as Figure 25, but split by emission line properties and only showing galaxies with $z > 0.605$. Galaxies in the left panel have at least one emission line with an S/N greater than three. Galaxies in the right panel have a relatively bright magnitude limit ($H_{160} < 23$) and no detected emission lines with an S/N greater than two. As in Figure 25, each tickmark on the right vertical axis corresponds to 1000 spectra. The survey contains >2000 spectra of relatively bright quiescent galaxies.

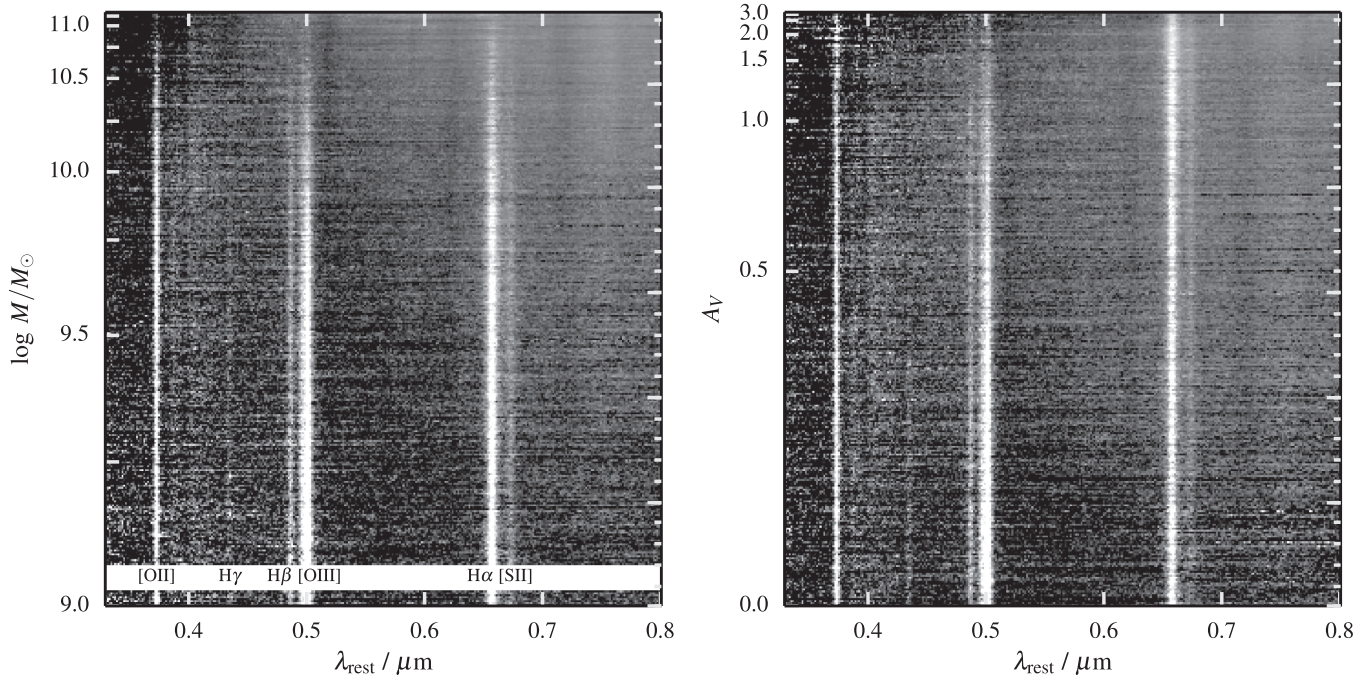

Figure 28. Same as Figure 25, but with objects sorted by M_* (left panel) and continuum dust extinction (A_V, right panel), both determined from stellar population synthesis fits to the broadband photometry. Here galaxies with a range of redshifts contribute to each row, providing rest-frame spectra from 3300 to 8000 Å. There are clear trends: higher mass galaxies have weaker emission lines and stronger absorption lines, and galaxies with higher continuum extinction have stronger Balmer decrements (see the text).

definitions for the line and continuum wavelength regions. We note that our spectral resolution is too poor to use common definitions such as the Lick system (Worthey et al. 1994). We *can* use our data for full spectrum fitting techniques (see Conroy et al. 2014), as demonstrated in van Dokkum & Brammer (2010) and Whitaker et al. (2013). Prominent absorption features include multiple TiO molecular bands at low redshift, the Mg_2 λ5170 feature at $1.2 \lesssim z \lesssim 2.2$, and the Balmer break at the highest redshifts.

We show 2D stacks with a physical galaxy property (rather than redshift) as the sorting parameter in Figure 28. Each line in these stacks was created from spectra at a wide range of redshifts; this is

the reason why the rest-frame wavelength coverage is much larger than in Figure 25. In the left panel, the galaxies are sorted by their stellar mass, as determined from stellar population synthesis models (see Skelton et al. 2014, for a detailed discussion). The dependence on mass is striking: at low masses the galaxies have strong emission lines, and at high masses the emission lines are weak or absent. The spectra also become gradually redder with increasing mass, and the Balmer break becomes more prominent. In the right panel, the sorting parameter is the continuum extinction A_V, again determined from fits to the broadband SEDs. The spectra again become redder with increasing A_V, as expected. Rather strikingly, the Balmer decrement Hα/Hβ increases strongly with A_V. Price et al. (2014) demonstrate this effect using earlier 3D-HST catalogs and a narrow redshift range, and it is very clear in this 2D stack, which uses a larger number of spectra and combines data from a wide range of redshifts.

8.3. Spatially Resolved Emission Lines

Arguably the most unique contribution of 3D-HST is the fact that all emission lines are *imaged* at *HST*'s superb resolution. For each object, the grism effectively places images at different wavelengths next to each other on the detector, with each subsequent image 23 Å (in interlaced space) separated from its neighbors. As a result, if an object is particularly bright in a single emission line, the grism will produce a complete image of the object in the light of that line (see Nelson et al. 2015, for a more in-depth explanation). The only data sets that can achieve something comparable are obtained with laser guide star assisted adaptive optics (AO) observations with IFUs on large telescopes (e.g., Genzel et al. 2006). These AO observations yield only one object at a time, and even though the diffraction-limited performance of VLT and Keck is superior to that of *HST*, the AO-delivered PSF generally has a much poorer Strehl ratio than the *HST* PSF.

Because the lines are broad for large galaxies, it is generally not trivial to obtain these maps from extracted spectra (see Nelson et al. 2012). We therefore provide continuum-subtracted maps in the data release, which can be directly used. Examples of these maps are shown in the bottom panels of the 2D grism spectra shown in Figure 10. The power of these spatially resolved line emission maps is demonstrated in several 3D-HST papers. Nelson et al. (2012, 2013, 2015) study the spatial distribution of Hα emission in galaxies at $z \sim 1$, in different bins of mass and SFR. Brammer et al. (2012a) show the G141 spectrum of a spectacular lensed galaxy with very strong emission lines. The 2D spectrum (Figure 2 in that paper) demonstrates that the grism provides images of the arc in the light of a range of different emission lines. Wuyts et al. (2013) compare the spatial distribution of Hα emission in a sample of relatively bright galaxies to that of the rest-frame UV emission.

9. SUMMARY

In this paper, we have described the observations and data products of the 3D-HST Treasury program. This is a companion study to the photometric analysis presented in Skelton et al. (2014), and together these two papers present a comprehensive photometric and spectroscopic wide-field data set for studies of the distant universe. All data products are available through the 3D-HST website.[36]

[36] http://3dhst.research.yale.edu/

Table 8
Removed Reads

Field	Pointing	FLT	Removed Reads	Defect
AEGIS	01	ibhj39uuq	[9,10,11,12]	E
		ibhj39viq	[11,12,13]	E
	04	ibhj42oqq	[10]	S
	21	ibhj59cvq	[3]	S
	28	ibhj66dfq	[3]	S
	30	ibhj69hgq	[10,11,12]	E
COSMOS	03	ibhm31rcq	[3]	S
	08	ibhm36ksq	[11]	S
	10	ibhm38dfq	[5]	S
	17	ibhm45y1q	[6]	S
	23	ibhm51x2q	[5]	S
	25	ibhm53o3q	[1]	S
GOODS-S	10	ibhj10vmq	[9]	S
	30	ibhj30bzq	[6]	S
	36	ibhj36j3q	[10]	S
	37	ibhj37uvq	[11]	S
UDS	05	ibhm05fjq	[7]	S
	13	ibhm13kfq	[6,7]	S
	14	ibhm14vhq	[9]	S
	18	ibhm18oeq	[2]	S
	19	ibhm19pjq	[9]	S
	20	ibhm20l1q	[10]	S
	24	ibhm24f8q	[3]	S
	26	ibhm26n0q	[9]	S
	27	ibhm27qiq	[11]	S
GOODS-N	11	ib3701s4q	[1,2,3]	E
		ib3701skq	[1,2,3]	E
	12	ib3702u8q	[1,2,3,4,5,6,7]	E
		ib3702uoq	[1,2,3,4,5,6,7,8]	E
	13	ib3703uzq	[1,2,3,4,5,6,7]	E
		ib3703vfq	[11,12,13]	E
		ib3703vmq	[1,2,3,4,8]	E
	14	ib3704wrq	[1,2]	E
		ib3704x8q	[6,7,8,9,10,11,12,13]	E
	17	ib3719v7q	[5]	S
	21	ib3705y1q	[1,2,3,4,5,6]	E
		ib3705y5q	[4]	S
		ib3705ylq	[4,5,6,7,8,9,10]	E
	22	ib3706b2q	[2,3,4,5,6,7]	E
		ib3706biq	[12,13]	E
		ib3706bpq	[1,2,3]	E
	23	ib3707caq	[1,2]	E
		ib3707cqq	[7,8,9,10,11,12]	E
		ib3747zyq	[11]	S
		ib3747a5q	[5]	S
	24	ib3708i5q	[7,8,9]	E
		ib3708ipq	[2,3,4,12]	E
	28	ib3724riq	[13]	S
	31	ib3709j3q	[7,8,9,10]	E
		ib3709joq	[8,11,12,13]	E
		ib3749o5q	[3]	S
		ib3749oqq	[11]	S
	32	ib3710nmq	[12,13]	E
	33	ib3711bkq	[4]	E
	36	ib3726bpq	[13]	S
		ib3726c5q	[10]	S
	44	ib3716psq	[12,13]	E

Note. E = earthshine and S = satellite.

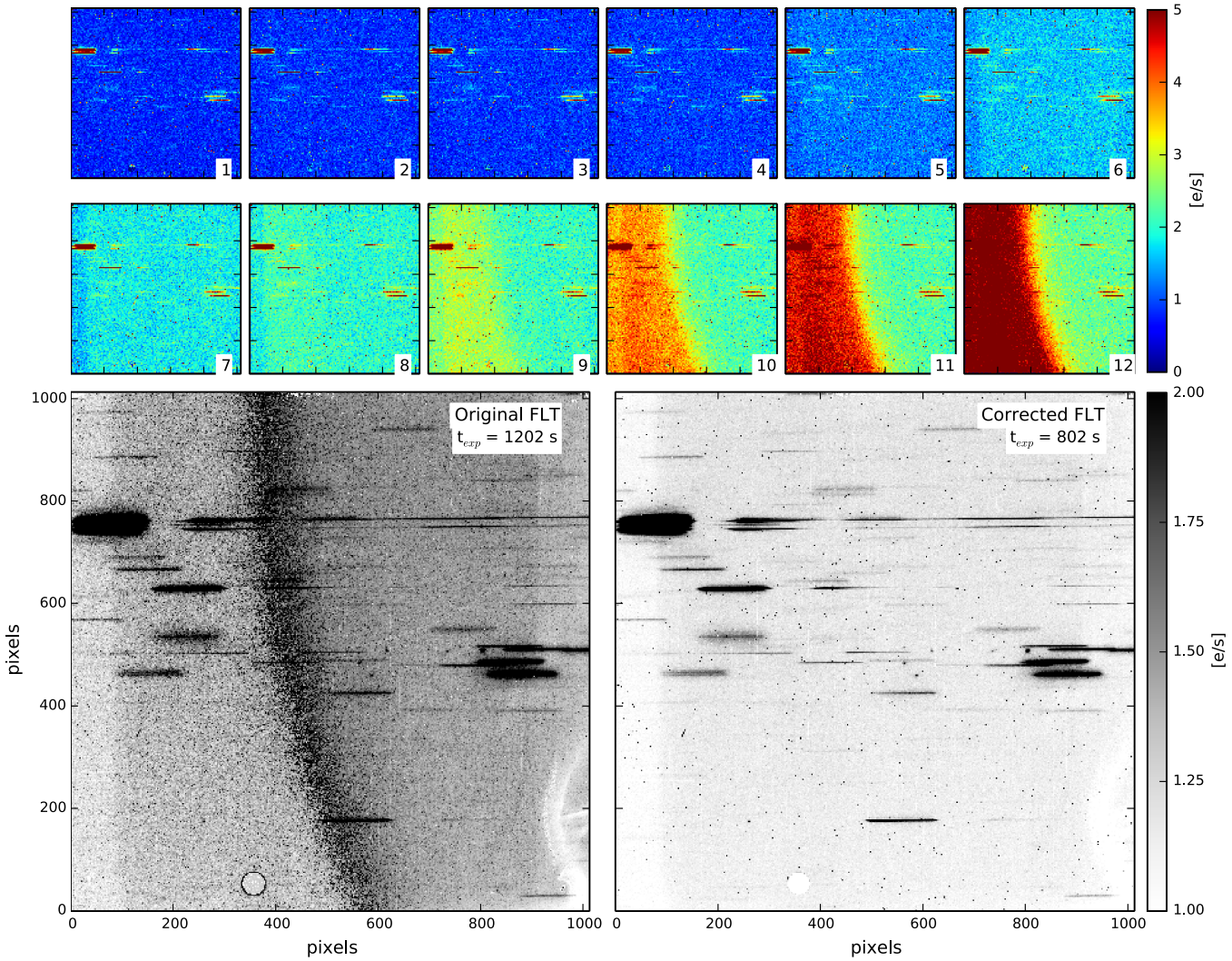

Figure 29. Main left panel: a pipeline-processed FLT file that is affected by earthshine. Top rows: individual reads, which show that only the last four reads in the exposure are affected. Main right panel: corrected FLT file after removing the problematic reads.

These data sets, together with structural parameters and SFRs presented elsewhere (van der Wel et al. 2014; Whitaker et al. 2014), accomplish an important goal of observational extragalactic astronomy: a census of stars and star formation in reasonably bright galaxies out to $z \sim 2.5$. The main source of uncertainty is shifting from errors in counting to errors in interpreting: systematic uncertainties in stellar masses, SFRs, and gas-phase metallicities are beginning to dominate in the regime discussed in this paper. There are excellent prospects to extend the work described in this paper to fainter objects and larger areas: the *James Webb Space Telescope*, *WFIRST*, and *Euclid* will use multi-slit and slitless near-IR spectroscopy to characterize the galaxy population in regions of parameter space that are beyond the capabilities of the WFC3 camera on *HST*.

This work is based on observations taken by the 3D-HST Treasury Program (GO 12177 and 12328) as well as GO 11600 and GO 13420 with the NASA/ESA *HST*, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. R.B. and K.E.W. gratefully acknowledge support by NASA through Hubble Fellowship grants #HST-HF-51318.001 and #HST-HF2-51368 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. We thank the anonymous referee for a careful reading of the manuscript.

APPENDIX
REMOVING INDIVIDUAL READS FROM FLT FRAMES

Our process for removing bad reads is detailed in Section 3.2.1 in the main text. Briefly, we use the fact that the WFC3 IR camera does multiple non-destructive reads per exposure (12 in the case of our G141 exposures). Particular problems such as satellites passing in the field of view, or earthshine, only affect one or a few of these 12 independent samples, and we can reconstruct a clean exposure by removing the offending reads. Here we provide a list of all the removed reads (Table 8), as well as examples of the removal of the effects of earthshine and satellite trails.

Figure 29 shows the effect of earthshine on an exposure, and demonstrates how we remove it. The example in this figure is one of the FLT images of pointing AEGIS-01: ibhj39uuq_flt.fits. The original `calwf3` pipeline-processed FLT exposure is in the

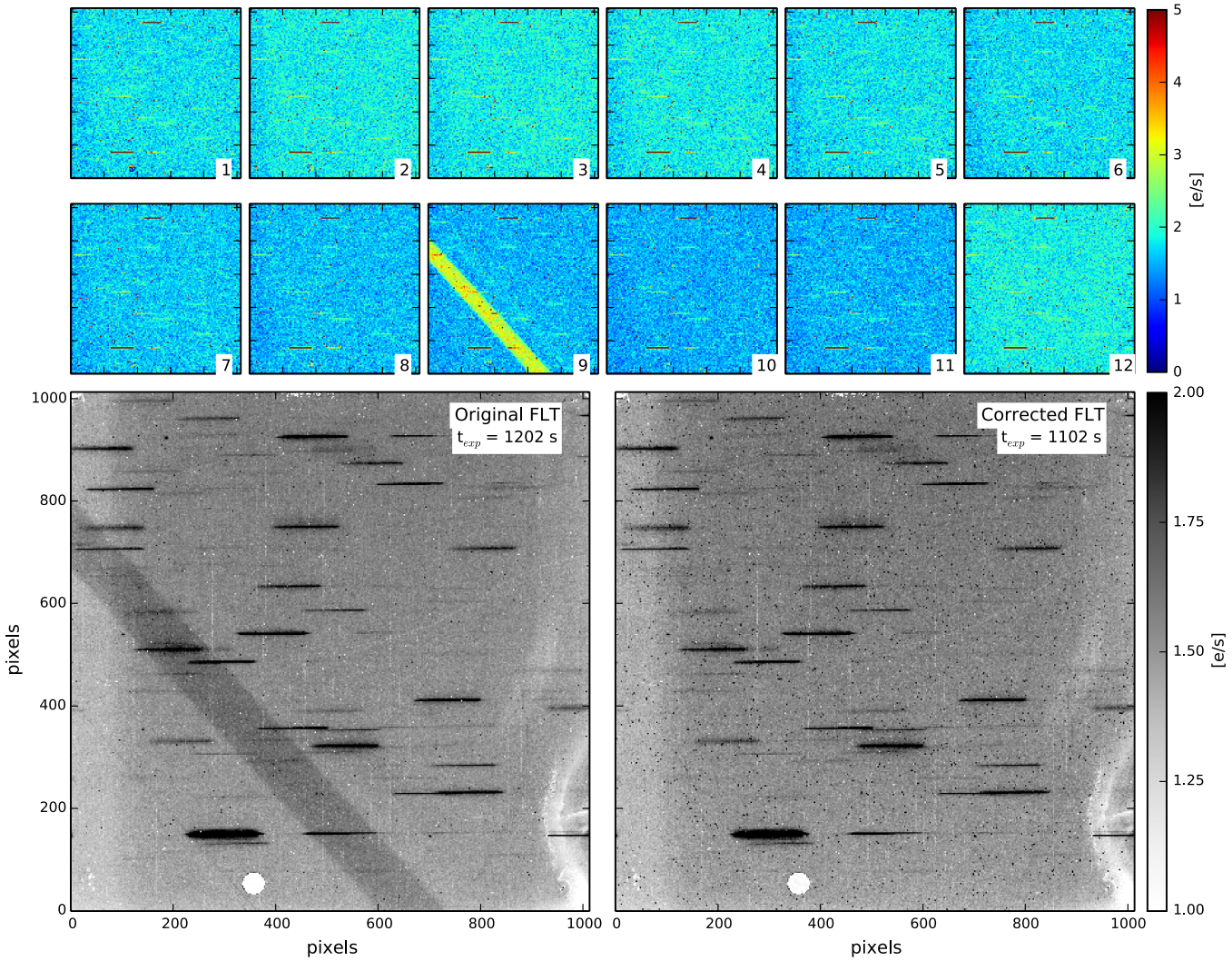

Figure 30. Example of satellite removal. The structure of the figure is the same as Figure 29. The sample obtained in the ninth read was removed in the corrected frame.

left main panel: the earthshine produces a highly structured background, and an apparently unusable exposure. The top two rows show the differences between each sequential pair of non-destructive reads of the WFC3 detector, converted to units of e^{-} s^{-1}. It is clear that only the last four reads are affected. In this example, we remove the last four reads in the sequence. The right main panel shows the corrected FLT image after removing the last four reads. The corrected FLT has 30% lower exposure time; however, without the correction, the full exposure would have been unusable.

The process of removing satellite trails is illustrated in Figure 30, which has the same structure as Figure 29. Here the example is one of the FLT images of pointing GOODSS-10: ibhj10vmq_flt.fits. A satellite moved across the observed field between reads 8 and 9. After removing the sample obtained in read 9 the corrected FLT file shows no trace of the satellite trail. As only a single read was removed the exposure time of the corrected frame is only 100 s shorter than the uncorrected one.

REFERENCES

Atek, H., Malkan, M., McCarthy, P., et al. 2010, ApJ, 723, 104
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bezanson, R., Wake, D. A., Brammer, G. B., et al. 2016, ApJ, 822, 30
Blanton, M. R., & Roweis, S. 2007, AJ, 133, 734
Brammer, G., Pirzkal, N., McCullough, P., & MacKenty, J. 2014, Time-varying Excess Earth-glow Backgrounds in the WFC3/IR Channel, Tech. Rep.
Brammer, G. B., Sánchez-Janssen, R., Labbé, I., et al. 2012a, ApJL, 758, L17
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012b, ApJS, 200, 13
Brammer, G. B., van Dokkum, P. G., Illingworth, G. D., et al. 2013, ApJL, 765, L2
Colbert, J. W., Teplitz, H., Atek, H., et al. 2013, ApJ, 779, 34
Conroy, C., Graves, G. J., & van Dokkum, P. G. 2014, ApJ, 780, 33
Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, ApJ, 626, 680
Dobos, L., Csabai, I., Yip, C.-W., et al. 2012, MNRAS, 420, 1217
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Förster Schreiber, N. M., Genzel, R., Bouché, N., et al. 2009, ApJ, 706, 1364
Fruchter, A. S., & Hook, R. N. 2002, PASP, 114, 144
Fumagalli, M. 2015, ApJ, submitted
Fumagalli, M., Franx, M., van Dokkum, P., et al. 2016, ApJ, 822, 1
Fumagalli, M., Patel, S. G., Franx, M., et al. 2012, ApJL, 757, L22
Gallego, J., Zamorano, J., Rego, M., Alonso, O., & Vitores, A. G. 1996, A&AS, 120, 323
Genzel, R., Tacconi, L. J., Eisenhauer, F., et al. 2006, Natur, 442, 786
Gonzaga, S., et al. 2012, The DrizzlePac Handbook (Baltimore, MD: STScI)
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Horne, K. 1986, PASP, 98, 609
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36

Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, ApJS, 218, 15
Kümmel, M., Rosati, P., Fosbury, R., et al. 2011, A&A, 530, A86
Kümmel, M., Walsh, J. R., Pirzkal, N., Kuntschner, H., & Pasquali, A. 2009, PASP, 121, 59
Kuntschner, H., Bushouse, H., Kümmel, M., Walsh, J. R., & MacKenty, J. 2010, Proc. SPIE, 7731, 3
Long, K. S., Baggett, S., & MacKenty, J. W. 2013, Characterizing Persistence in the WFC3 Channel: Observations of Omega Cen, Tech. Rep.
Malhotra, S., Rhoads, J. E., Pirzkal, N., et al. 2005, ApJ, 626, 666
Mancini, C., Förster Schreiber, N. M., Renzini, A., et al. 2011, ApJ, 743, 86
McLean, I. S., Steidel, C. C., Epps, H. W., et al. 2012, Proc. SPIE, 8446, 0
Mendel, J. T., Saglia, R. P., Bender, R., et al. 2015, ApJL, 804, L4
Morris, A. M., Kocevski, D. D., Trump, J. R., et al. 2015, AJ, 149, 178
Nelson, E., van Dokkum, P., Franx, M., et al. 2014, Natur, 513, 394
Nelson, E. J., van Dokkum, P. G., Brammer, G., et al. 2012, ApJL, 747, L28
Nelson, E. J., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2015, arXiv:1507.03999
Nelson, E. J., van Dokkum, P. G., Momcheva, I., et al. 2013, ApJL, 763, L16
Newman, S. F., Buschkamp, P., Genzel, R., et al. 2014, ApJ, 781, 21
Patel, S. G., Fumagalli, M., Franx, M., et al. 2013, ApJ, 778, 115
Pirzkal, N., Pasquali, A., & Walsh, J. 2002, STECF, 32, 7
Pirzkal, N., Viana, A., & Rajan, A. 2010, Instrument Science Report WFC3, 6
Pirzkal, N., Xu, C., Malhotra, S., et al. 2004, ApJS, 154, 501
Price, S. H., Kriek, M., Brammer, G. B., et al. 2014, ApJ, 788, 86
Quadri, R. F., & Williams, R. J. 2010, ApJ, 725, 794
Rajan, A., et al. 2010, WFC3 Data Handbook, Version 2.1
Rodney, S. A., Riess, A. G., Dahlen, T., et al. 2012, ApJ, 746, 5
Schmidt, K. B., Rix, H.-W., da Cunha, E., et al. 2013, MNRAS, 432, 285
Sha, F., Lin, Y., Saul, L. K., & Lee, D. D. 2007, Neural Computation, 19, 2004
Sharples, R., Bender, R., Agudo Berbel, A., et al. 2013, Msngr, 151, 21
Sharples, R., Bender, R., Agudo Berbel, A., et al. 2014, Proc. SPIE, 9147, 0
Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24
Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, ApJ, 795, 165
Straughn, A. N., Kuntschner, H., Kümmel, M., et al. 2011, AJ, 141, 14
Straughn, A. N., Meurer, G. R., Pirzkal, N., et al. 2008, AJ, 135, 1624
van der Wel, A., Bell, E. F., Häussler, B., et al. 2012, ApJS, 203, 24
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
van Dokkum, P., Brammer, G., Momcheva, I., Skelton, R. E., & Whitaker, K. E. & for the 3D-HST team. 2013a, arXiv:1305.2140
van Dokkum, P. G., Bezanson, R., van der Wel, A., et al. 2014, ApJ, 791, 45
van Dokkum, P. G., & Brammer, G. 2010, ApJL, 718, L73
van Dokkum, P. G., Brammer, G., Fumagalli, M., et al. 2011, ApJL, 743, L15
van Dokkum, P. G., Franx, M., Fabricant, D., Illingworth, G. D., & Kelson, D. D. 2000, ApJ, 541, 95
van Dokkum, P. G., Leja, J., Nelson, E. J., et al. 2013b, ApJL, 771, L35
van Dokkum, P. G., Nelson, E. J., Franx, M., et al. 2015, ApJ, 813, 23
Whitaker, K. E., Franx, M., Leja, J., et al. 2014, ApJ, 795, 104
Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, ApJL, 754, L29
Whitaker, K. E., van Dokkum, P. G., Brammer, G., et al. 2013, ApJL, 770, L39
Wisnioski, E., Förster Schreiber, N. M., Wuyts, S., et al. 2015, ApJ, 799, 209
Worthey, G., Faber, S. M., Gonzalez, J. J., & Burstein, D. 1994, ApJS, 94, 687
Wuyts, S., Förster Schreiber, N. M., Nelson, E. J., et al. 2013, ApJ, 779, 135
Zamorano, J., Rego, M., Gallego, J. G., et al. 1994, ApJS, 95, 387