






















# A Two-dimensional Spectroscopic Study of Emission-line Galaxies in the Faint Infrared Grism Survey (FIGS). I. Detection Method and Catalog

Norbert Pirzkal<sup>1</sup>, Barry Rothberg<sup>2,3</sup> , Russell E. Ryan<sup>1</sup>, Sangeeta Malhotra<sup>4,5</sup> , James Rhoads<sup>4,5</sup> , Norman Grogan<sup>1</sup> ,  
Emma Curtis-Lake<sup>6</sup>, Jacopo Chevallard<sup>6</sup>, Stephane Charlot<sup>6</sup> , Steven L. Finkelstein<sup>7</sup> , Anton M. Koekemoer<sup>1</sup> ,  
Parviz Ghavamian<sup>8</sup> , Myriam Rodrigues<sup>9</sup>, François Hammer<sup>9</sup> , Mathieu Puech<sup>9</sup>, Rebecca L. Larson<sup>7</sup> , Lise Christensen<sup>10</sup> ,  
Andrea Cimatti<sup>11,12</sup> , Ignacio Ferreras<sup>13</sup>, Jonathan P. Gardner<sup>5</sup>, Caryl Gronwall<sup>14,15</sup>, Nimish P. Hathi<sup>1</sup> , Bhavin Joshi<sup>4</sup> ,  
Harald Kuntschner<sup>16</sup>, Gerhardt R. Meurer<sup>17</sup>, Robert W. O’Connell<sup>18</sup> , Goeran Oestlin<sup>19</sup>, Anna Pasquali<sup>20</sup>, John Pharo<sup>2</sup> ,  
Amber N. Straughn<sup>5</sup>, Jeremy R. Walsh<sup>16</sup>, Darach Watson<sup>10</sup> , Rogier A. Windhorst<sup>2</sup> , and Nadia L Zakamska<sup>21</sup> 

<sup>1</sup> Space Telescope Science Institute, Baltimore, MD 21210, USA

<sup>2</sup> Large Binocular Telescope Observatory, University of Arizona, AZ 85721, USA

<sup>3</sup> George Mason University, Fairfax, VA 22030, USA

<sup>4</sup> Arizona State University, School of Earth and Space Exploration, Tempe, AZ 85287, USA

<sup>5</sup> NASA’s Goddard Space Flight Center, Astrophysics Science Division, Code 660, Greenbelt MD 20771, USA

<sup>6</sup> Sorbonne Université, UPMC-CNRS, UMR7095, Institut d’Astrophysique de Paris, F-75014, Paris, France

<sup>7</sup> University of Texas at Austin, Austin, TX 78712, USA

<sup>8</sup> Department of Physics, Astronomy and Geosciences, Towson University, Towson, MD 21252, USA

<sup>9</sup> GEPI, Observatoire de Paris, PSL University, CNRS, 5 Place Jules Janssen, F-92190 Meudon, France

<sup>10</sup> Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, DK-2100, Denmark

<sup>11</sup> University of Bologna, Department of Physics and Astronomy, Via Gobetti 93/2, I-40129, Bologna, Italy

<sup>12</sup> INAF—Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50125, Firenze, Italy

<sup>13</sup> Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey, RH5 6NT, UK

<sup>14</sup> Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA

<sup>15</sup> Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA

<sup>16</sup> European Southern Observatory, D-85748, Garching, Germany

<sup>17</sup> International Centre for Radio Astronomy Research, The University of Western Australia, Crawley, WA 6009, Australia

<sup>18</sup> The University of Virginia, Charlottesville, VA 22904-4325, USA

<sup>19</sup> Stockholm University, SE-10691, Stockholm, Sweden

<sup>20</sup> Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, D-69120, Heidelberg, Germany

<sup>21</sup> Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA

Received 2018 June 1; revised 2018 September 25; accepted 2018 September 26; published 2018 November 20

## Abstract

We present the results from the application of a two-dimensional emission line detection method, EMission-line two-Dimensional (EM2D), to the near-infrared G102 grism observations obtained with the Wide-Field Camera 3 (WFC3) as part of the Cycle 22 *Hubble Space Telescope* Treasury Program: the Faint Infrared Grism Survey (FIGS). Using the EM2D method, we have assembled a catalog of emission line galaxies (ELGs) with resolved star formation from each of the four FIGS fields. Not only can one better assess the global properties of ELGs, but the EM2D method allows for the analysis and improved study of the individual emission-line region *within* each galaxy. This paper includes a description of the methodology, advantages, and the first results of the EM2D method applied to ELGs in FIGS. The advantage of 2D emission line measurements includes significant improvement of galaxy redshift measurements, approaching the level of accuracy seen in high-spectral-resolution data, but with greater efficiency; and the ability to identify and measure the properties of multiple sites of star formation and over scales of  $\sim 1$  kpc within individual galaxies out to  $z \sim 4$ . The EM2D method also significantly improves the reliability of high-redshift ( $z \sim 7$ ) Ly $\alpha$  detections. Coupled with the wide field of view and high efficiency of space-based grism observations, EM2D provides a noteworthy improvement on the physical parameters that can be extracted from grism observations.

**Key words:** galaxies: distances and redshifts – galaxies: star formation – galaxies: structure – methods: data analysis

## 1. Introduction

One indicator of (relatively) recent star formation in galaxies is the presence of strong emission lines (see Kennicutt 1998, and references therein). Recombination lines, such as Ly $\alpha$ , H $\alpha$ , H $\beta$  in the rest-frame UV and optical, and the prominent rest-frame optical forbidden emission lines [O II] and [O III], are all integral in tracing the ionizing flux produced by short-lived ( $\sim 10$  Myr), massive ( $> 10 M_{\odot}$ ) stars. It is important to note that these lines are produced not only in the central nuclear regions, but in star-forming regions throughout galaxies. Emission-line galaxies (ELGs) are predominantly identified in narrowband

photometric (e.g., Djorgovski et al. 1985; Boroson et al. 1993) or spectroscopic grism surveys (e.g., Mayall 1936; Markarian 1967; Smith 1975). ELGs are easily identifiable because a significant amount of the photons radiated by them originate in star-forming regions producing strong emission lines. More recently, with the availability of slitless grism modes on the *Hubble Space Telescope* (HST), with the Near Infrared Camera and Multi-Object Spectrometer (NICMOS), the Advanced Camera for Surveys (ACS), the Wide-Field Camera 3 (WFC3), several projects, such as the NICMOS/HST Grism Parallel Survey (McCarthy et al. 1999), the ACS Pure Parallel Ly $\alpha$  Emission Survey (Pasquali et al. 2003), the Grism ACS

Program for Extragalactic Science (Pirzkal et al. 2004), Probing Evolution And Reionization Spectroscopically (PEARS; Pirzkal et al. 2009), the WFC3 Infrared Spectroscopic Parallel Survey (Atek et al. 2010), the Grism Lens-Amplified Survey from Space (Treu et al. 2015), 3D-*HST* (Momcheva et al. 2016), and the Faint Infrared Galaxy Survey (FIGS; Pirzkal et al. 2017), have identified a large population of star-forming galaxies (SFGs) over a nearly contiguous range of redshifts ( $0 < z < 3.5$ ), including the epoch of peak star formation at  $1.5 < z < 2.5$  (Madau et al. 1998; Hopkins 2004; Madau & Dickinson 2014).

The presence of bright, easily identifiable emission lines makes the spectroscopic determination of the redshift of these individual ELGs straightforward. These same emission lines also allow for the direct measurement of physical properties, such as star formation rates, ages, and metallicities, of both the star-forming regions within ELGs, as well as inferring the overall global properties of galaxies (e.g., Aller 1942; McGaugh 1991; Kewley & Dopita 2002; Nagao et al. 2006). This includes variations in kinematics and star-forming properties across a galaxy (e.g., Rubin et al. 1970, 1972).

For the most part, space-based *HST* grism surveys have focused primarily on the *integrated* properties of SFGs and ELGs. These programs have preferred to sample wider angular coverage on the sky or random parallel fields in order to increase the potential sample size and mitigate cosmic variance. The downside is that such surveys often use only one or two orientations on the sky in an effort to increase efficiency and reduce overheads. Yet, this method also increases the effects and the impact of contamination from other sources in or near the field of view (i.e., light dispersed from foreground objects in the field, or just outside the field, which contaminates the dispersed light of targets of interest). Most surveys also sacrifice depth per field in favor of larger angular sky coverage. While space-based slitless grism observations are often more effective than ground-based counterparts, particularly at longer optical and near-infrared wavelengths (where telluric emission and absorption lines reduce any sensitivity gains from larger aperture mirrors), they still must be carefully planned to avoid contamination from non-telluric sources.

In an earlier series of papers (Straughn et al. 2008, 2009; Pirzkal et al. 2013), we successfully demonstrated how ACS slitless grism observations obtained using multiple orientations on the sky, could be used to significantly improve the wavelength accuracy of emission lines, as well as actually identifying multiple emission line sources *within* a galaxy by avoiding self-contamination (i.e., the dispersed light of one part of the galaxy, contaminating the dispersed light from an adjoining part of the galaxy when using a single orientation on sky). However, our earlier series of papers were restricted to the redshift range of  $0 < z < 1.5$  using rest-frame H $\alpha$ , [O III], or [O II] obtained with ACS on *HST*. While the results were interesting, they fell just short of probing the important peak of star formation ( $1.5 < z < 2.5$ ).

The use of the WFC3 camera and its near-IR grisms offers a comparably large field of view and flux sensitivities as the ACS grism, while allowing us to detect ELGs at redshifts up to  $z = 2$ . The FIGS program obtained deep G102 observations of four distinct fields (two in GOODS-N and two in GOODS-S), each using five distinct orientations on sky (Pirzkal et al. 2017). While the choice of multiple position angles was driven primarily by the need to mitigate the amount of contamination by other dispersed nearby sources, it also created an advantageous situation that allowed to us to search for ELGs

using the methods laid out in Pirzkal et al. (2013), namely detecting the presence of emission lines and their source star formation regions without having to first find or detect the underlying host galaxy. The results from our work at  $0 < z < 1.5$  in PEARS (Pirzkal et al. 2013) resulted in three key findings: (1) the computed line luminosities showed evidence of a flattening in the luminosity function slope with increasing redshift; (2) the star-forming systems showed evidence of complex morphologies with star formation occurring predominantly within one effective (half-light) radius. However, the morphologies showed no correlation with host stellar mass; and (3) the number density of SFGs with  $M_* \gg 10^9 M_\odot$  decreases by an order of magnitude at  $z \leq 0.5$  relative to the number at  $0.5 < z < 0.9$ , supporting the argument of galaxy downsizing (Cowie et al. 1996).

In this paper, we probe up to and including the peak epoch of star formation using near-IR observations obtained with the G102 grisms using the WFC3/IR camera. First, we discuss the EM2D methodology, then apply it to the WFC3/IR G102 science data in Section 3.1. Section 3.2 explores emission line identification, and Section 3.3 describes how line fluxes are measured. Section 3.4 describes the physical location of star-forming regions in individual galaxies. Section 3.5 shows how we identified a significant number of galaxies with diffuse emission and how emission line maps can be created. In Section 3.7 we provide a discussion of the completeness of our EM2D survey. Finally, Sections 4.1–4.3 discuss high-EW galaxies, high-redshift galaxies, and the luminosity functions we derive for the FIGS EM2D galaxy sample. The goal of this paper is to demonstrate the effectiveness of two-dimensional emission line detection, diagnostics, and their application to discerning new insights into the key epoch of star formation in the universe. All calculations in this paper assume  $H_0 = 67.3 \text{ km s}^{-1} \text{ Mpc}^{-1}$  and  $\Omega_M = 0.315$ ,  $\Omega_\Lambda = 0.685$  (Planck Collaboration et al. 2014). All magnitudes are given in the AB system (Oke & Gunn 1983).

## 2. Observations

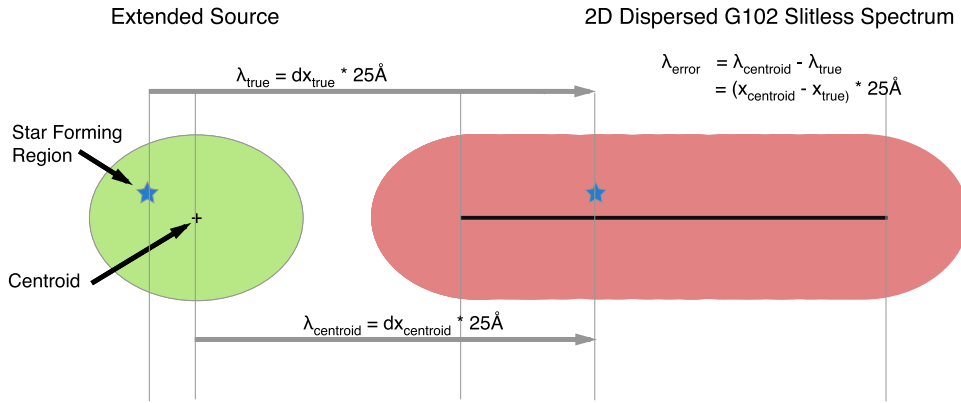
The data presented here are from the *HST* Cycle 22 Treasure Program Faint Infrared Galaxies (FIGS, Proposal ID: 13779, PI: S. Malhotra), which was awarded 160 orbits ( $\sim 100$  ks total exposure time) with the WFC3/IR instrument using the G102 grism filter. The field of view of the WFC3/IR channel is  $2''2 \times 2''2$  and the G102 grism has a resolution of  $25 \text{ \AA}$  per pixel ( $\approx 36 \text{ \AA}$  effective resolution since the WFC3 PSF is  $\approx 1.5$  pixel wide). FIGS focuses on four distinct fields in the Great Observatories Origins Deep Survey (GOODS) North and South fields (Giavalisco et al. 2004): two fields in GOODS-North and two fields in GOODS-South (for details, see Pirzkal et al. 2017). Each field was observed with five distinct position angles (PAs) in order to minimize contamination and maximize the signal-to-noise ratio (S/N). The FIGS survey reaches a  $3\sigma$  continuum depth of  $\approx 26 \text{ AB}$  magnitudes and probes emission lines down to  $\approx 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$  (Pirzkal et al. 2017).

## 3. Data Reduction and Analysis

### 3.1. The Emission-line 2D (EM2D) Method

#### 3.1.1. Methodology

In “classical” slit (or multi-slit) spectroscopy, only the the light from the object within the slit or slits is dispersed across



**Figure 1.** Illustration of how the difference in position between the true and estimated location of the origin of an emission line leads to an error in the observed wavelength estimate for that emission line.

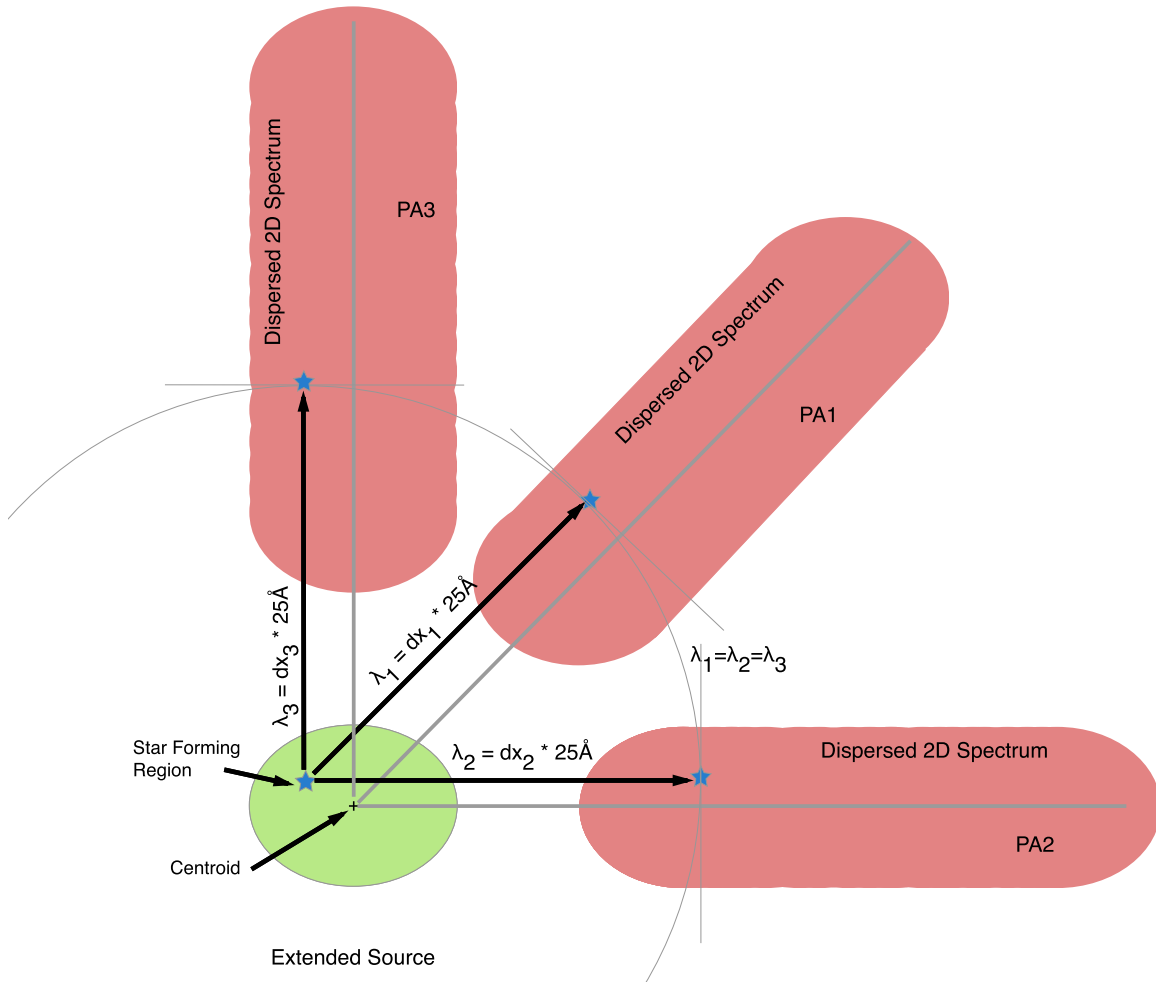
the detector. While this avoids contamination from other nearby objects in the field, it reduces the survey efficiency significantly. Slitless grism spectroscopy is more effective in terms of the number of objects for which light can be dispersed, and it is ideally suited for survey work in which one does not know a priori which objects in the field are of particular interest. However, in a grism survey, each astronomical source acts as a dispersing object, potentially contaminating itself and other sources. Moreover, light from higher grism orders can also be dispersed onto the detector, allowing each object to contaminate many other objects many times, even if the source is outside the nominal field of view as projected onto the detector. Furthermore, a resolved object can “self-contaminate,” that is, its own dispersed light falls on top of adjacent pixels in the target, which are themselves dispersing light. At best, this results in the morphology of the source in the dispersion direction acting similar to a convolution kernel, which blurs the final spectrum in the dispersion direction. At worst, if the object is spectrally inhomogeneous, as in cases of an SFG with knots and clumps containing emission lines, determining the exact wavelength of the emission line requires detailed knowledge of where the signal originates within the object. Traditionally, both slit and slitless spectroscopy assume that light is produced at the center of the source and that the actual shape of the source has the same effect as a simple convolution kernel. While valid for slit spectroscopy (assuming the source is not spatially resolved *within* the slit), this assumption is not valid for slitless spectroscopy in which objects are resolved (even partially). Figure 1 demonstrates that this assumption can result in a systematic error in the determination of the observed wavelength of an emission (or absorption) line from a region that is not centrally located (i.e., a star-forming knot, star-cluster, second nucleus in an unrelaxed merger, etc.) within a galaxy.

In a case where a component is resolved within a galaxy, rotating the field (i.e., changing the *HST* PAs) and dispersing the spectrum of the source in a different direction projected on the sky allows one to recover *both* the exact observed wavelength and the physical source of an emission line. This is because the emission line is assumed to be monochromatic and is dispersed in a specific direction and at a specific distance from the source. Figure 2 illustrates this, and shows how observations taken at multiple PAs of a resolved galaxy containing a star-forming region producing an emission line are related. The EM2D method inverts this approach: starting with emission line candidates identified directly in dispersed 2D

images, the dispersion solution of the grism is inverted to determine where the source of the emission line might be located. For a single PA, the source of the emission line candidate can be anywhere along the spectral trace that goes through the detected emission line (shown as thick black lines in Figure 2). When two PAs are available, the intersections of these multiple traces points to the actual source of the emission lines. Each time an intersect is found, it can be used to compute the observed wavelengths of the emission lines, which, if real, should be identical, as shown in Figure 2. In practice, the derived wavelengths are not identical, but they should be consistent when considering both position measurement errors of each emission line in the dispersed images and the accuracy of the existing grism calibration. This method automatically rejects zeroth spectra orders since these would result in widely inconsistent wavelength estimates when observed in different PAs. When more than two PAs are available, pairs of observations can be used to determine these intersects. Multiple intersects at the same positions indicate more robust detections based on the emission line measured in independent observations with different PAs. Thus, the more intersects (or the more PAs) are available, the more refined and the smaller the uncertainty/error in the computed observed wavelength. Figure 3 demonstrates this with an example showing an object with two distinct star-forming regions that would appear blended in the extracted spectra if using traditional slitless extraction methods.

### 3.1.2. Detection of Emission Lines

The first step of the reduction process is to use the reduced, background-subtracted, and astrometrically registered FLT images and their associated simulation-based extraction (SBE), which are described in Section 3 of Pirzkal et al. (2017). For each of the five available PAs of a given FIGS field, we used DrizzlePac’s *astrodrizzle* (Avila et al. 2015) to produce a combined G102 dispersed mosaic and a combined simulated G102 mosaic by combining all of the observed data and simulated data, respectively. The simulated mosaic was then subtracted from the combined dispersed mosaic to produce a deep dispersed observation where the continuum light was removed. The initial emission line candidate list at each PA was generated by running SeXtractor (Bertin & Arnouts 1996) on the continuum-subtracted mosaics. All spatial features  $2.5\sigma$  above the local background were selected as candidate targets, yielding a few



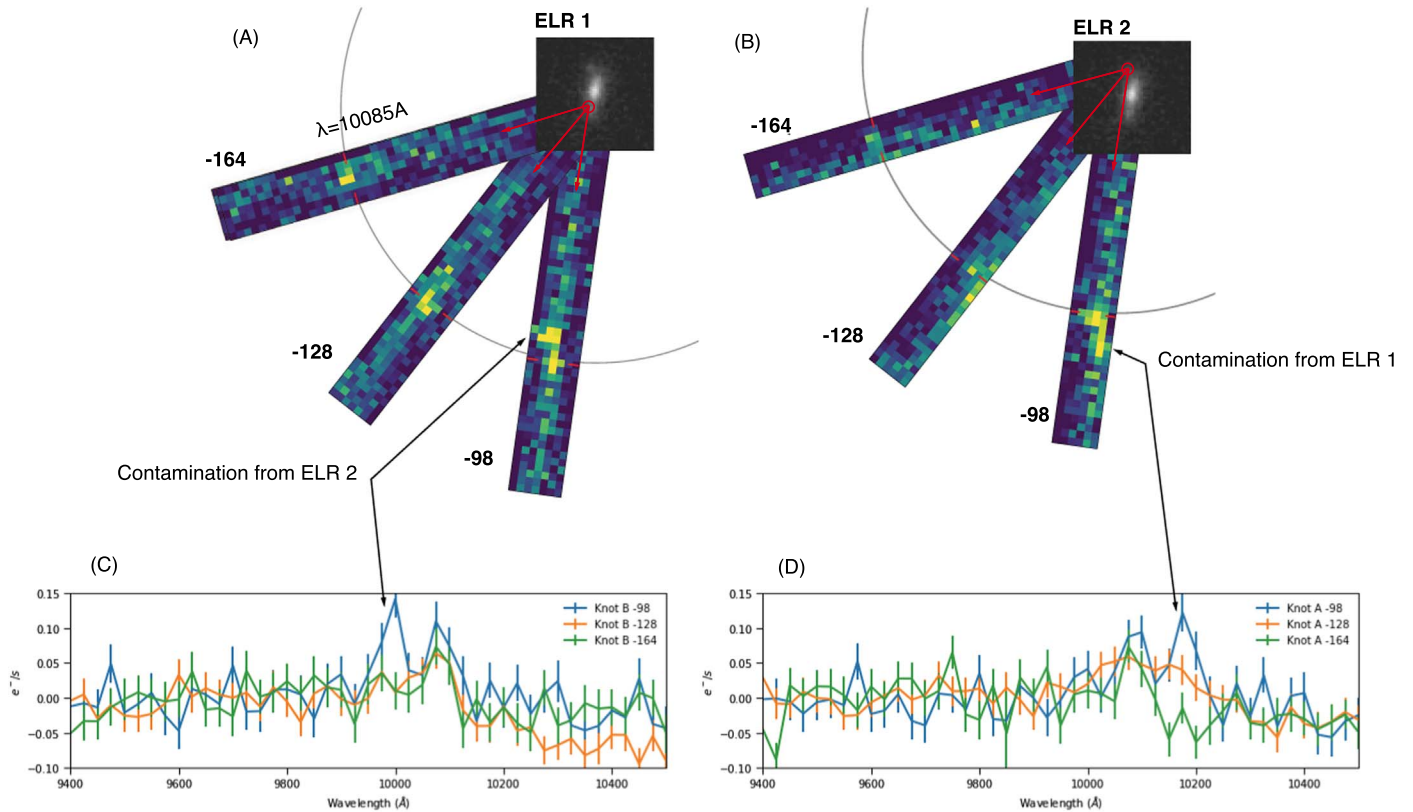
**Figure 2.** Illustration of the EM2D method and how multiple orientations allow us to both pinpoint the origin of an emission line as well as determine an accurate estimate of the observed wavelength of the emission line.

thousand candidates per mosaic. Next, the spectral trace was computed for each of the emission line candidates using the five available PAs and the G102 grism calibration from Pirzkal et al. (2016). This was repeated for each of the four FIGS fields. The exact methodology and figures detailing the initial 2D extraction and resulting 1D spectra can be found in Sections 3.3.1 and 3.3.2 of Pirzkal et al. (2017).

For all of the candidates, the intersect between two spectral traces was computed and the observed wavelengths of the emission lines were obtained from these intersects. If the inferred wavelength of the two line candidates differed by more than 48 Å (equivalent to 2 WFC3 pixels), the intersect was discarded. When using 2, 3, 4, 5 (and  $n$ ) PAs, the maximum of possible pairs number is 1, 3, 6, 10, and, generally speaking,  $(n^2 - n)/2$ . A single region with a strong emission line could therefore appear as 10 separate intersects in a list. A robust list of emission line regions was selected using a density-based spatial clustering (DBSCAN) algorithm (Ester et al. 1996). This algorithm groups together points closely packed together within a defined parameter space and rejects outliers in low-density regions of this space. In this case, the R.A., decl., and wavelength ( $\lambda$ ) comprise the three-dimensional parameter space, and a cluster size of  $0''.129$  in R.A. and decl. and 25 Å in wavelength was used to define intersects. The three parameters of each cluster were then averaged to produce the

emission-line region candidate list. Each of these candidates has a grade equal to the number of intersects at this location and with this wavelength. Emission lines at different wavelengths, i.e., [O III]  $\lambda\lambda 4959, 5007$  and H  $\alpha$   $\lambda 6563$  + [N II]  $\lambda\lambda 6548, 6584$  were allowed to originate from distinct star-forming regions, where any object can have more than one star-forming region. In fact, no restrictions were imposed on the number or location of star-forming regions, nor were they required to be located within known objects. The advantage of this is that it permits for the discovery of naked emission lines, i.e., objects with clear emission lines that are not detected in the continuum.

Although at least two intersects can improve the accuracy of both the wavelength of an emission line and its location within a galaxy, we required that an emission line be detected in at least three different PAs in order to be considered for further analysis. The more stringent requirement assures that detections are robust against contamination, errant pixels, cosmic rays, etc. After this cut, a total of 1338 emission line candidates were identified in the four FIGS fields. Spectra of the corresponding star-forming regions were extracted in a manner similar to that described in Sections 3.3.1 and 3.3.2 of Pirzkal et al. (2017), but restricting the extraction width to twice the size of the detected emission lines in the 2D dispersed images. Spectra from different PAs were combined together using a weighted average. Due to the very low S/N levels probed using the EM2D method, this initial list of emission line candidates



**Figure 3.** The EM2D method used in this paper. We rely on the detection of emission line candidates in dispersed images in different PAs and use our knowledge of the dispersion to determine the location of the emission-line regions. As this figure illustrates, objects with multiple emission line regions (ELR1 and ELR2) can be identified, which is something not readily done using normal extraction. The latter can in fact lead to erroneous line detection when multiple emission lines are produced within different emission line regions. Panels A and B show the location of the emission line for each ELR in the original observations, which is at a wavelength of 10085 Å. The emission lines are marked on the 2D spectra using red lines and the dispersion directions are shown by red arrows pointing away from the source. Panels C and D show the extracted spectra and illustrate how different ELRs in a galaxy can contaminate each other. This is readily visible in the 2D spectra as well as the 1D spectra for PA = −98. On the other hand, PA = −164 allows for a clean extraction of the spectra, while PA = −128 results in some smearing of the emission line.

required individual visual inspection to remove false positives, and to verify the quality of an emission line at the wavelength computed using the EM2D method. The first pass at line fitting allowed for fitting a single Gaussian profile as well as multiple Gaussian profiles to account for the [O II]  $\lambda\lambda 3726, 3729$ , [O III]  $\lambda\lambda 4959, 5007$ , H $\beta$   $\lambda 4861$ , H $\alpha$   $\lambda 6563 + [\text{N II}] \lambda\lambda 6548, 6584$ , and [S II]  $\lambda\lambda 6717, 6731$  multiplets. This final cut resulted in a final list of 338 emission line candidates within 302 separate emission-line regions (82, 53, 83, 83 in the GN1, GN2, GS1, and GS2 fields, respectively). This final list was then compared and correlated with the FIGS object catalogs (Section 3.2.2 in Pirzkal et al. 2017) in order to locate the host galaxies. These emission line candidates were produced by a total of 234 distinct objects (58, 47, 63, 66 in the GN1, GN2, GS1, and GS2 fields, respectively).

The method presented here builds on the 2D emission line scheme for PEARS (PEARS-2D) first presented in Straughn et al. (2009) and Pirzkal et al. (2013). However, the main differences are that the PEARS-2D method required only a minimum of two PAs to accept an emission line detection, each with a peak at least  $1.1\sigma$  above the background. Emission lines were detected in a continuum-subtracted image where the continuum was estimated by smoothing the data. The extraction relied purely on the reduction package *aXe* (Pirzkal et al. 2001; Kümmel et al. 2009). The improved methodology and reduction used here, and described in more detail in Pirzkal

et al. (2017), significantly improve both background subtraction, and contamination by using SBE.

### 3.2. Emission-line Identification

Once the spectra are extracted, the next step is to identify the detected emission lines in each of the 338 candidate objects. The method of identification assumes the observed lines are Ly $\alpha$ , [O II]  $\lambda\lambda 3726, 3729$ , H $\beta$   $\lambda 4861$ , [O III]  $\lambda\lambda 4959, 5007$ , or H $\alpha$   $\lambda 6563 + [\text{N II}] \lambda\lambda 6548, 6584$ , as these are the brightest, most prominent rest-frame optical lines. Identification of emission lines is based on a classification scheme for the host galaxies. Hosts are classified as Type I if the spectroscopic redshift of the host galaxy was already publicly available from sources in the literature. These redshifts were used as a starting point to identify the emission line and determine the final redshift of the source. Type II hosts are those for which we could identify multiple lines such as the H $\alpha$   $\lambda 6563 + [\text{N II}] \lambda\lambda 6548, 6584 + [\text{S II}] \lambda\lambda 6717, 6731$  or the [O III]  $\lambda\lambda 4959, 5007 + \text{H}\beta$   $\lambda 4861$  lines. The H $\alpha$   $\lambda 6563 + [\text{N II}] \lambda\lambda 6548, 6584$  and [S II]  $\lambda\lambda 6717, 6731$  lines are resolved in the G102 grism (see Figure 7). Similarly, all three [O III]  $\lambda\lambda 4959, 5007$  and H $\beta$   $\lambda 4861$  lines are resolved. Type II hosts also include objects with multiple lines such as [O II]  $\lambda\lambda 3726, 3729$  together with [O III]  $\lambda\lambda 4959, 5007$ , or [O III]  $\lambda\lambda 4959, 5007$  together with H $\alpha$   $\lambda 6563 + [\text{N II}] \lambda\lambda 6548, 6584$ . The G102 FIGS data were supplemented with existing archival G141 data (using data from proposals

11600, 12099, 12177, and 12461) extending wavelength coverage to  $1.6 \mu\text{m}$ ), which were processed and extracted in the same manner as the FIGS data. However, the G141 data were *only* used to detect other bright emission lines, such as spectroscopically confirming the identification of [O II]  $\lambda\lambda 3726, 3729$  in the G102 data. We made limited use of the G141 data, because the G141 observations are significantly more shallow and suffer from heterogeneous coverage compared to the five-epoch G102 FIGS data. This makes it problematic to compute the completeness levels for the G141 data. Type III galaxies are those cases where neither publicly available redshifts, nor identification of multiple emission lines could be made. Instead, spectrophotometric redshift estimates of the host galaxy were used. To compute these, photometric redshifts were first determined using the program BEAGLE (Chevallard & Charlot 2016) and the available FIGS photometric catalogs. Next, a posterior distribution of the redshifts for each object was created to help identify an observed single emission line and compute a more accurate spectroscopic redshift for the source (See, Xu et al. 2007; Xia et al. 2011).

In some cases, candidate objects could be identified using more than one approach. The order of priority for emission line identification is: (1) ground-based spectroscopy; (2) grism spectroscopy; and (3) photometric redshifts (ground- and/or space-based). When a spectroscopic redshift was in good agreement with our grism observations, we designated the host galaxy as Type I. When multiple emission lines in the slitless grism data improved upon a known spectroscopic redshift, we designated the host galaxy as Type II. In most cases, multiple emission lines were identified in a spectrum. In a few rare circumstances, emission line identifications were uncertain. This occurred specifically when a single emission line was detected with the G102 slitless grism data, and was inconsistent with the photometric redshift of the host galaxy for all of the lines summarized above. In these instances, such objects were removed from the sample. The final sample then contains a total of 302 star-forming regions in 234 distinct galaxies, 159 of which were spectroscopically confirmed as Type I or Type II. A more detailed breakdown of our sample is shown in Table 1 and Table 2, while Figure 4 shows the distribution of the host galaxy apparent magnitudes as a function of redshift. As we show in Figure 5 we identified emission line regions in approximately 20% of the entire FIGS sample, down to a continuum magnitude of  $m_{F105W} = 25$ .

### 3.2.1. Redshift Accuracy

As noted in Section 3.1.1, the de facto standard of wavelength identification is to assume that the source of the emission line is the broadband center of the object. However, in cases where the emission line region and/or underlying host galaxy is resolved (or partially resolved), there will exist an offset between the actual source of the emission line and the broadband center of the object. This offset results in an error in the value of the observed wavelength of the emission line. This error is proportional to the distance between the emission line region and the galaxy center. Because EM2D is robust enough to detect emission line regions within a host galaxy, it can be used to significantly improve the accuracy of redshifts attained with slitless grism spectroscopy, approaching the level of accuracy one can obtain from higher-resolution slit spectroscopy.

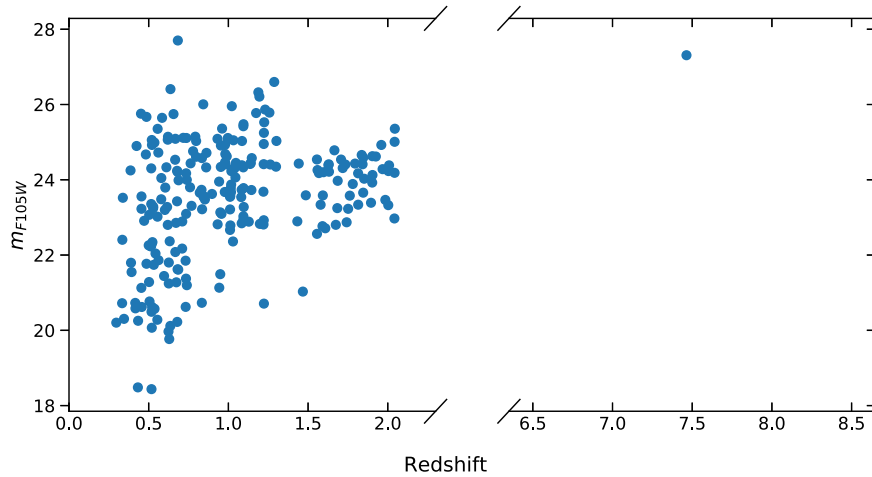
**Table 1**  
FIGS EM2D Emission-line Regions and Galaxies

Field	Type	Number of Galaxies	Number of SF Regions
GN1	I	31	41
GN1	II	15	29
GN1	III	12	12
GN1	Total	58	82
GN2	I	16	19
GN2	II	14	16
GN2	III	17	18
GN2	Total	47	53
GS1	I	23	34
GS1	II	22	28
GS1	III	18	21
GS1	Total	63	83
GS2	I	19	23
GS2	II	19	24
GS2	III	28	37
GS2	Total	66	84
All	I	89	117
All	II	70	97
All	III	75	88
All	Total	234	302

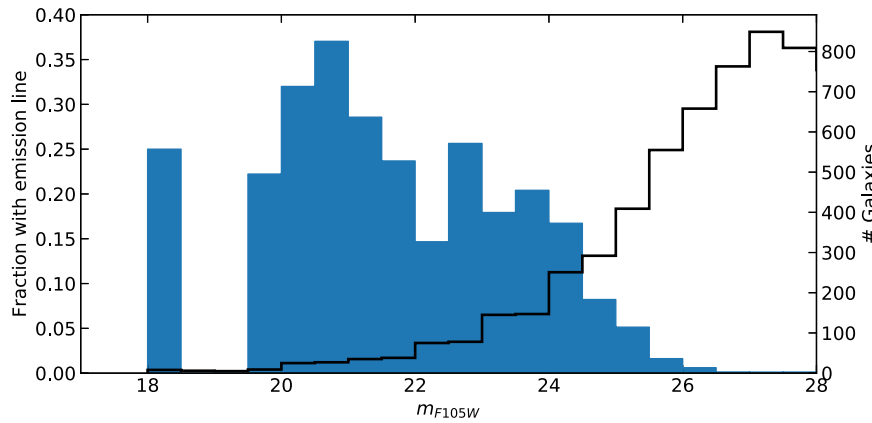
**Note.** Breakdown of the number of distinct host galaxies and distinct emission-line regions detected and identified using the EM2D method. We also show the breakdown for Type I (spectroscopic), Type II (grism spectroscopic), and Type III (spectrophotometric redshift) line and therefore redshift identification.

The limiting factor in the accuracy of the EM2D method is the ability to accurately measure the exact position of the emission lines in the continuum-subtracted slitless observations (See Section 3.1.2). Using a traditional centroiding method, emission lines with  $S/N = 100$  can be determined down to  $\sim 0.1$  pixel accuracy. The error in the location of the star-forming regions in galaxies is equivalent to a native WFC3 IR pixel, or  $0''.129$ . The EM2D method can discern emission line candidates that differ by at most  $25 \text{ \AA}$  so that the expected error in our wavelength calibration is better than  $|\delta z|/(1+z) \approx 0.002$ . This accuracy can be improved by increasing the number of position angles used.

We can empirically quantify the improvement made to the redshift determination of these sources using the FIGS data themselves, and calculating what observed wavelength we would have measured for every emission line in our sample when observed at each PA on the sky. In each case, the location of the emission lines in the host galaxy is used, and the resulting offset projected in the dispersion direction  $\Delta x$ . The resulting error in wavelength is then  $\Delta \lambda \approx \Delta x \times 25 \text{ \AA}$ . This allows for the direct estimation of the effect of spatially offset stellar emission lines when using low-resolution slitless observations. For the purpose of this test, only objects with a single and robustly detected emission line region were used, namely those that have been detected in  $n \geq 8$  combinations of PAs (See Section 3.2). Figure 6 demonstrates how large a fraction of the redshift derived for the FIGS sources would be off by  $|\delta z|/(1+z) > 0.01$ , assuming that the source of the emission line is also the center of the source. Taking into



**Figure 4.** Host galaxies' apparent magnitude (F105W) as a function of redshift in the FIGS EM2D sample.



**Figure 5.** Fraction of FIGS galaxies observed in more than two position angles with an emission line region identified using the EM2D method (filled histogram). We also show the distribution of FIGS galaxies observed in more than two position angles.

**Table 2**  
FIGS EM2D Emission Lines

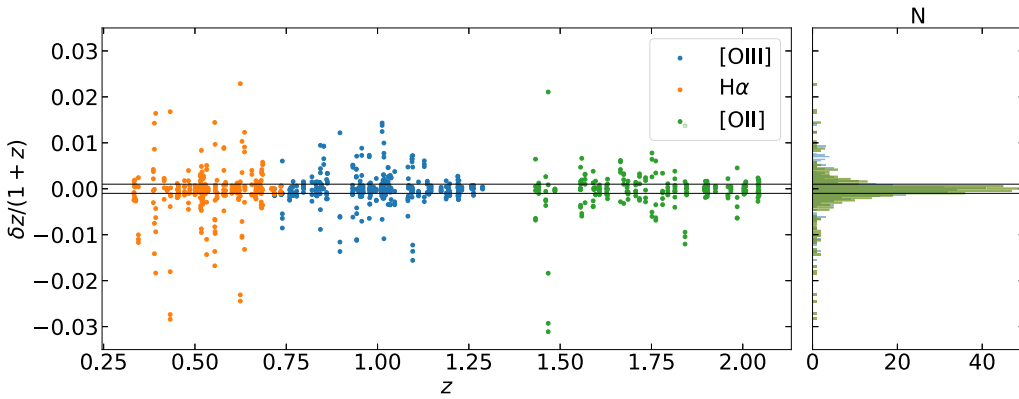
Line	GN1	GN2	GS1	GS2	Total
$H\alpha$ $\lambda$ 6563+[N II] $\lambda$ $\lambda$ 6548,6584 + [S II] $\lambda$ $\lambda$ 6717,6731	44	11	39	44	138
[O III] $\lambda$ $\lambda$ 4959,5007 + $H\beta$ $\lambda$ 4861	31	25	38	22	116
[O II] $\lambda$ $\lambda$ 3726,3729	7	16	10	26	59
$Ly\alpha$	0	0	0	1	1

**Note.** Number of individual  $H\alpha$   $\lambda$ 6563+[N II] $\lambda$  $\lambda$ 6548,6584, [S II] $\lambda$  $\lambda$ 6717,6731, [O III] $\lambda$  $\lambda$ 4959,5007 +  $H\beta$  $\lambda$  4861, and [O II] $\lambda$  $\lambda$ 3726,3729 lines measured in the EM2D FIGS galaxy sample.

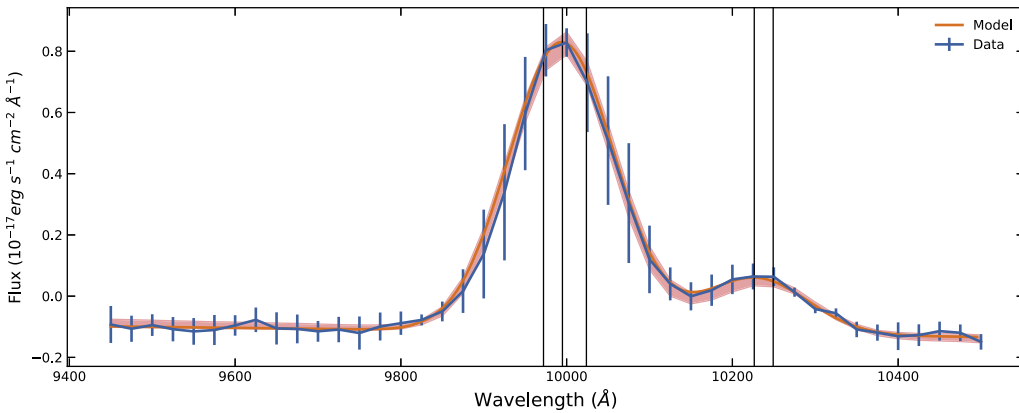
account that the resolution of the EM2D method is 1 native WFC3 G102 pixel, or  $0''.129$  in the image and  $25 \text{ \AA}$  in the spectra, 3% of the redshift estimates of the FIGS sources would be off by  $|\delta z|/(1+z) > 0.01$ , 8% of the redshift estimates would be off by  $|\delta z|/(1+z) > 0.005$ , and 12% would be off by  $|\delta z|/(1+z) > 0.001$ . Thus, 2D methods such as EM2D can significantly improve the spectroscopic redshifts of ELGs by fully taking into account the location of the star-forming regions within these galaxies.

### 3.3. Flux Measurements

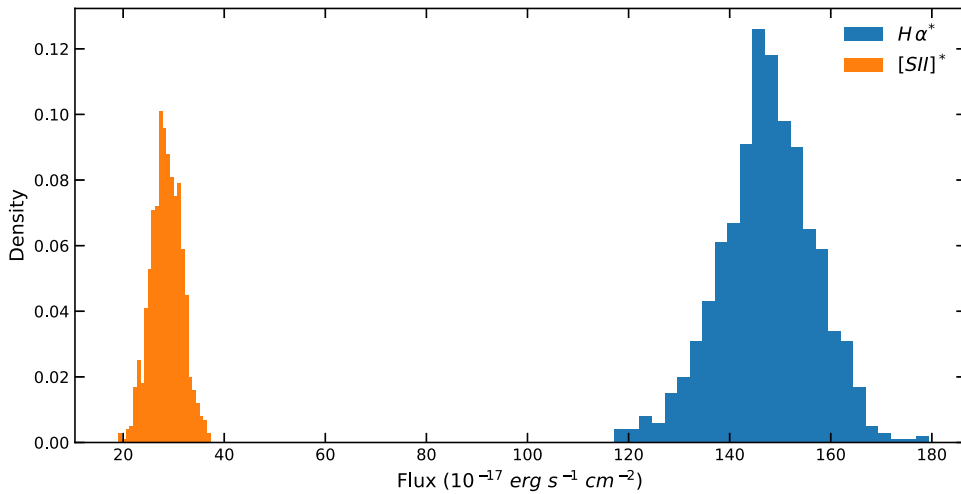
The improved line identification and redshift constraints from EM2D were used to systematically fit the  $H\alpha$   $\lambda$ 6563 +[N II] $\lambda$  $\lambda$ 6548,6584 triplet, the [S II] $\lambda$  $\lambda$ 6717,6731 doublet, and the [O III] $\lambda$  $\lambda$ 4959,5007,  $H\beta$  $\lambda$  4861, and [O II] $\lambda$  $\lambda$ 3726,3729 lines in the G102 spectra. A Gaussian profile was fit to each emission line, and a continuum level was estimated using the regions surrounding the expected positions of these emission lines. An initial fitting of each emission line was then followed by a Markov-Chain Monte Carlo (MCMC) based fitting method. This method is based upon the methodology and algorithms first described in Pirzkal et al. (2013). The MCMC fitting was used to define proper confidence intervals for the redshift of the source, the width of the emission lines (assumed to be the same for all the lines in a given spectrum), and the line fluxes. In the case of [O II] $\lambda$  $\lambda$ 3726,3729 the line flux ratio was defined to be within a range of  $0.25 < \lambda 3726/\lambda 3729 < 1.45$ , following Osterbrock & Ferland (2006). The [O III] $\lambda$  5007, [O III] $\lambda$  4959, and  $H\beta$  $\lambda$  4861 lines were fitted using three Gaussian profiles with a flux ratio of [O III] $\lambda$  5007 to [O III] $\lambda$  4959  $\approx 2.984$ , following Storey & Zeppen (2000). Finally, the  $H\alpha$   $\lambda$ 6563+[N II] $\lambda$  $\lambda$ 6548,6584 and the [S II] $\lambda$  6717/



**Figure 6.** The error in the redshift estimates of the FIGS emission line galaxies ( $\delta z/(1+z)$ ), if they had been observed using single PA observations, as a function of  $z$  for the  $H\alpha$ , [O III] and [O II] emission-line hosts. The solid black lines indicate  $\delta z/(1+z) = \pm 0.001$ . The panel on the right shows a histogram of the error in redshift.



**Figure 7.** Example of an MCMC based fit of the  $H\alpha$   $\lambda 6563$ + $[N II]\lambda\lambda 6548,6584$  and  $[S II]\lambda\lambda 6717,6731$  lines. This is source GS2 3186. The blue line with an error bar shows the data. The green line shows the formal best fit. The thin vertical lines show the location of the  $H\alpha$   $\lambda 6563$ + $[N II]\lambda\lambda 6548,6584$  and  $[S II]\lambda\lambda 6717,6731$  lines. As is often the case in slitless observations, the large line widths are caused by the spatial structure of the source and the instrumental PSF, not kinematics.



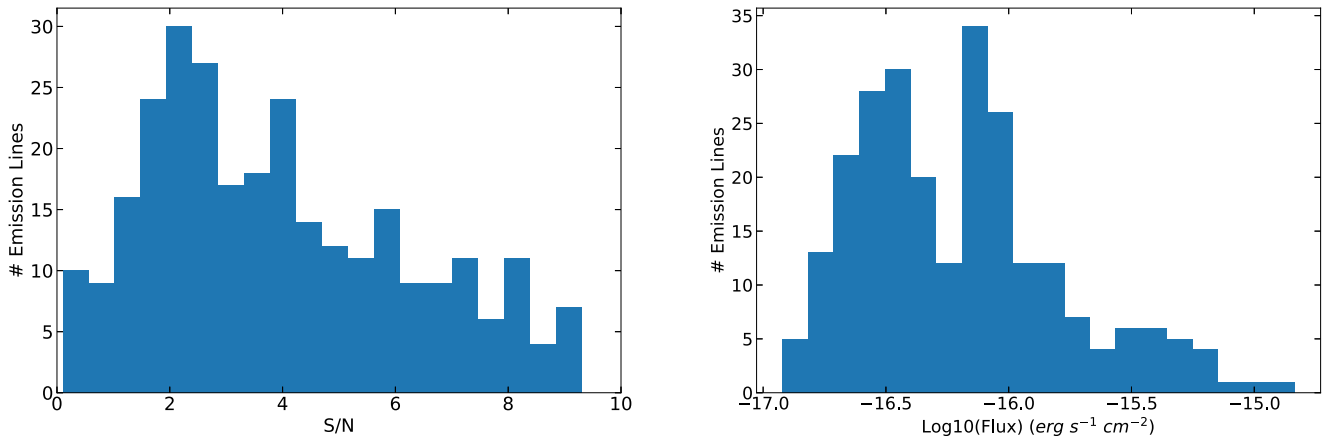
**Figure 8.** Histograms of the posterior distributions of the sum of the fluxes of the  $H\alpha$   $\lambda 6563$ + $[N II]\lambda\lambda 6548,6584$  lines, and of the  $[S II]\lambda\lambda 6717,6731$  lines in the spectrum shown in Figure 7.

$[S II]\lambda 6731 < 1.42$ ,  $[N II]\lambda 6583/[N II]\lambda 6548 = 3$ , and  $0.01 < H\alpha \lambda 6464/[N II]\lambda 6583 < 1$ , following Osterbrock & Ferland (2006). Figure 7 shows the best fit of the blended  $H\alpha$   $\lambda 6563$ + $[N II]\lambda\lambda 6548,6584$  lines and the  $[S II]\lambda\lambda 6717,6731$  doublet. The posterior probability distributions of the fluxes of the  $H\alpha$   $\lambda 6563$ + $[N II]\lambda\lambda 6548,6584$  lines and the  $[S II]\lambda\lambda 6717,6731$  doublet are shown in Figure 8. The distribution of line fluxes that we measured is shown in Figure 9, where we show the

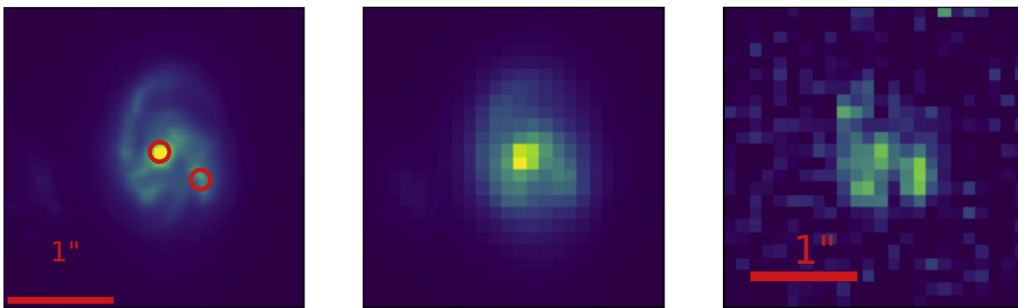
distributions of measured S/N and line flux of the FIGS EM2D emission lines.

The low resolution of the G102 grism does not allow for the lines to be fully resolved separately. Therefore, the line fluxes we derive for the  $H\alpha$   $\lambda 6563$  and the  $[N II]\lambda\lambda 6548,6584$  lines are highly degenerate, and we report the combined  $H\alpha$   $\lambda 6563$ + $[N II]\lambda\lambda 6548,6584$  line fluxes. This is also the case for the the  $[S II]\lambda\lambda 6717,6731$  lines. Similarly, the  $H\beta\lambda 4861$  and [O III]





**Figure 9.** Left panel: distribution of the S/N of the measured line fluxes of [O II] $\lambda\lambda$ 3726,3729, [O III] $\lambda\lambda$ 4959,5007, and H $\alpha$   $\lambda$ 6563+[N II] $\lambda\lambda$ 6548,6584 in the emission-line regions identified in FIGS galaxies. Right panel: distributions of fluxes for emission lines with an S/N greater than 2.



**Figure 10.** Left panel: galaxy GS1 1295 at  $z = 0.42$  and the two H $\alpha$   $\lambda$ 6563+[N II] $\lambda\lambda$ 6548,6584 emission-line regions (red circles) that we identified in this object using the EM2D method, as seen in the ACS F850LP filter. Middle panel: the same galaxy as seen in the WFC3 F105W IR filter. Right panel: a full 2D forward modeling of the H $\alpha$   $\lambda$ 6563+[N II] $\lambda\lambda$ 6548,6584 emission in this object, facilitated by the accurate EM2D estimate of the observed wavelength of the emission feature. There is significant structure and extended line emission in this objects, and the two brightest ones were detected using the EM2D method. The other knots, seen in the right panel, are too faint to be detected by EM2D.

$\lambda\lambda$ 4959,5007 were fit simultaneously, and in this case the wider separation of these lines allowed us to estimate the flux of these three lines independently. Finally, the [O II]  $\lambda\lambda$ 3726,3729 lines were fitted using two Gaussian profiles (although it is the sum of the flux from the doublet that is relevant for physical quantities such as the star formation rates) but as it is the case for the H $\alpha$   $\lambda$ 6563 and the [N II]  $\lambda\lambda$ 6548,6584 lines, the derived line fluxes are highly degenerate.

#### 3.4. Locations of Emission-line Regions

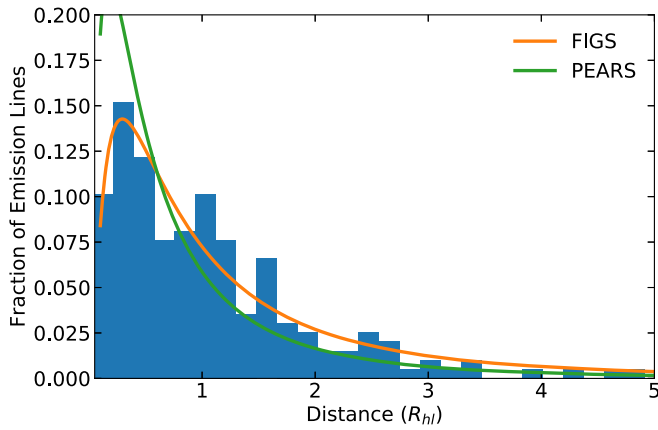
As noted above, the EM2D method allows us to identify the position line emission within specific galaxies. In Pirzkal et al. (2013), we showed that most emission lines are generated at a significant distance from the center of the galaxy over the redshift range of  $0 < z < 1.5$ . Using the near-IR FIGS data, we are now able to extend this work to the redshift range of  $0.3 < z < 2.5$ . Figure 10 shows the star-forming regions identified in a  $z = 0.42$  galaxy.

In Figure 11 we plot the distribution of emission line positions, normalized to the half-light radius ( $R_{hl}$ ) of the host galaxies. The latter was computed by measuring the  $R_{hl}$  of the galaxies in all available broadband *HST* imaging, and by computing the rest-frame  $R_{hl}$  by interpolating the values measured in different bands (ACS F435W, F606W, F775W, F814W, F850LP, and WFC3 F105W, F125W, F140W, and F160W where available). As we show in this figure, the distribution peaks at  $\approx 0.3R_{hl}$  and  $\approx 44\%$  of emission lines

originate more than one  $R_{hl}$  away from galaxy center, where the peak of the distribution is a consequence of the limited spatial resolution of the FIGS survey. This observed distribution is consistent with the one we derived in Pirzkal et al. (2013), although with a slightly larger proportion of knots being at larger distances. Exponential fits to the PEARS and FIGS distributions are also shown in Figure 11.

#### 3.5. Extended/Diffuse Emission

We note that a majority of the galaxies in the FIGS sample have unresolved, or only marginally resolved, star-forming regions. A small number of objects show signs of physically extended and/or diffuse emission. The extended sizes of these emission-line regions result in the apparent appearance of broadened lines in the extracted FIGS spectra. With the resolution of the G102 grism of  $\approx 36 \text{ \AA}$ , our ability to measure velocity dispersion is thus limited to about  $1000 \text{ km s}^{-1}$ , which is much greater than what one would expect from the dark matter halos in these galaxies. Some extended emission line regions were first flagged during the manual line classification process described in Section 3.1.2. Here, we found 53 extended H $\alpha$   $\lambda$ 6563+[N II] $\lambda\lambda$ 6548,6584 regions (in 41 distinct sources), and 9 extended [O II] $\lambda\lambda$ 3726,3729 regions. Defining a region to be spatially extended if the measured emission lines FWHM are wider than  $50 \text{ \AA}$ , we find that approximately 45% of the H $\alpha$   $\lambda$ 6563+[N II] $\lambda\lambda$ 6548,6584 regions are extended, while only 17% of the [O III] $\lambda$  5007 and [O II] $\lambda\lambda$ 3726,3729 regions are spatially extended. We can expect some of these objects to



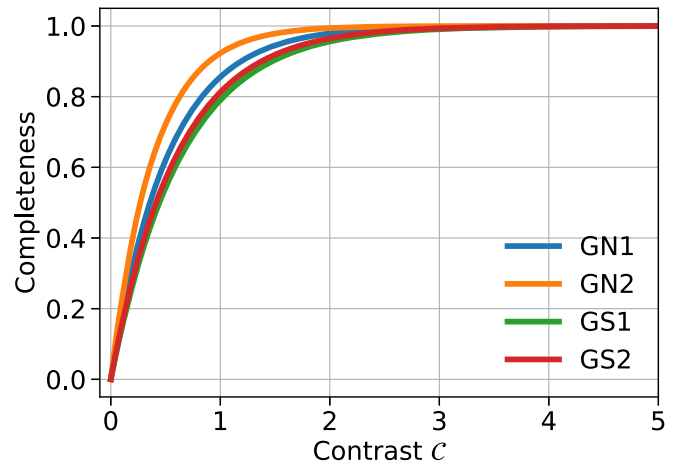
**Figure 11.** Distance of the FIGS emission line regions from the centroid of the host galaxies in units of galaxy half-light radius ( $R_{hl}$ ). We show an exponential fit to the FIGS histogram in orange. A fit of the PEARS distribution from Pirzkal et al. (2013) is shown in green.

have extended active galactic nucleus (AGN) emission lines. In order to estimate the fraction of the FIGS EM2D sources we expect to be caused by an AGN, we cross-correlated our emission line catalog with the X-ray catalogs of Luo et al. (2017), Xue et al. (2016), and Villforth et al. (2010). Assuming a 1 arcsec matching radius between the two, we find that  $\approx 4\%$  of the FIGS emission lines are potentially AGN-driven. This fraction increases to 10% if we allow for a 5 arcsec matching radius. Of the 41 objects showing spectroscopically resolved  $H\alpha$   $\lambda 6563 + [N II] \lambda \lambda 6548, 6584$  emission, we find that 3 (7) sources are detected in the X-ray when using a matching radius of 1 (5) arcsecond. Those AGN-dominated candidates are GN1 1497, GS1 2614, and GS2 1653 using a 1 arcsec matching radius, and GN1 1497, GS1 2614, GS1 2363, GS1 2518, GS1 4308, GS1 1299, and GS2 1653, using a 5 arcsec matching radius. We therefore find 7% (17%) of our extended  $H\alpha$   $\lambda 6563 + [N II] \lambda \lambda 6548, 6584$  sources to possibly be AGN-dominated.

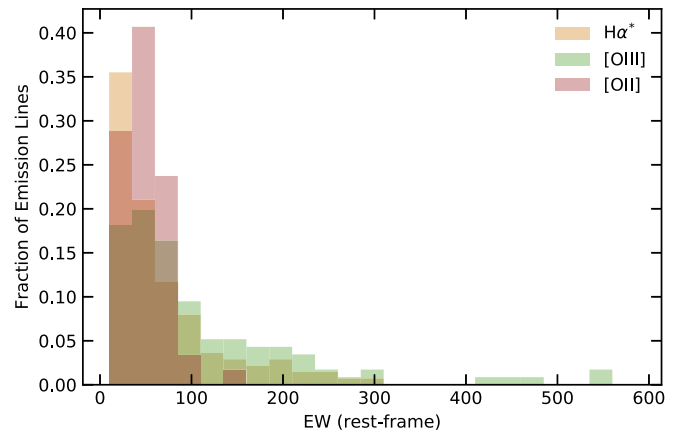
In cases of larger galaxies, extended emission can lead to multiple detections using the EM2D method, and can cause artificially high  $n$  values to be assigned to a single star formation knot as several knots are found to be very close together, and merged during the clustering step described above. We found 15 sources that show signs of extended emission based on this criterion ( $n > 10$ ) alone. The galaxy shown in Figure 10 is one of these. Once the observed wavelength of these emission lines is accurately derived, one can easily use this information to forward model the continuum-subtracted 2D spectra and reconstruct full 2D emission line maps. The result of the  $H\alpha$   $\lambda 6563 + [N II] \lambda \lambda 6548, 6584$  2D map reconstruction for the same galaxy is shown in Figure 10. Our implementation of this 2D reconstruction will be fully described in the next paper in this series, where we describe the physical properties of emission-line-selected FIGS galaxies in details.

### 3.6. Additional Emission Lines

The flux of additional emission lines was also measured. For each spectrum, we fitted Gaussian profiles to additional emission lines, assuming that these were at the observed wavelengths determined using the redshifts determined during the fit of the  $H\alpha$   $\lambda 6563 + [N II] \lambda \lambda 6548, 6584$  triplet, the  $[S II] \lambda \lambda 6717, 6731$  doublet, the  $[O III] \lambda \lambda 4959, 5007$ , the  $H\beta \lambda 4861$ , and the  $[O II] \lambda \lambda 3726, 3729$  lines. We used the same MCMC based flux measurement technique that we described above



**Figure 12.** Completeness function  $f(C)$  for the EM2D FIGS emission lines as a function of  $C = \frac{F(\text{line})}{F(\text{cont.})}$ .

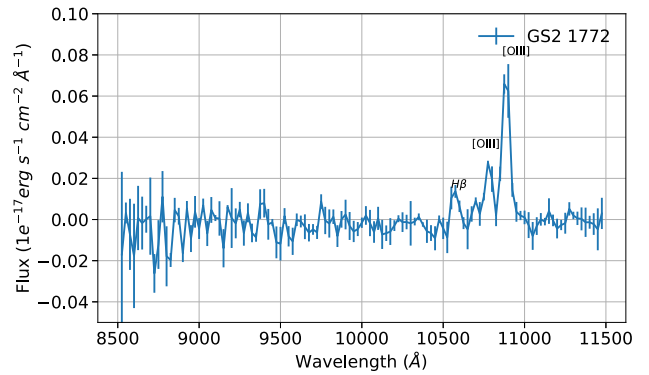
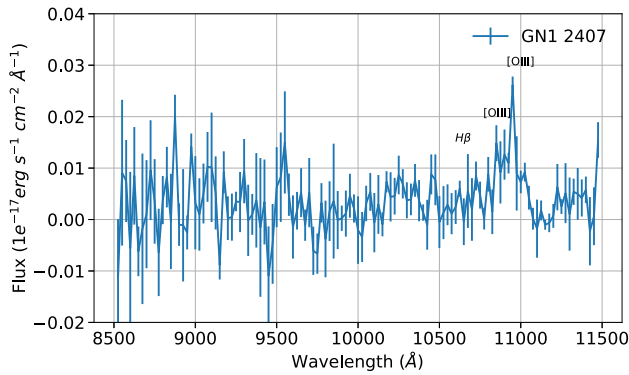


**Figure 13.** Distribution of rest-frame EWs for the  $[O II] \lambda \lambda 3726, 3729$  ( $[O II]^*$ ),  $[O III] \lambda 5007$ , and  $H\alpha$   $\lambda 6563 + [N II] \lambda \lambda 6548, 6584$  ( $H\alpha^*$ ) emission lines.

while fixing the width of the emission lines to the value computed for the brighter emission lines. We performed flux measurements for the  $He I \lambda 5877$ ,  $Mg II \lambda 2799$ ,  $O I \lambda 7774$ ,  $[Ar III] \lambda \lambda 7136$ ,  $[Ar III] \lambda \lambda 7753$  lines as well as the  $H\gamma \lambda 4342$ ,  $[O III] \lambda 4363$ ,  $[Ne III] \lambda \lambda 3869 / [Ne III] \lambda \lambda 3890$ ,  $[Ne V] \lambda \lambda 3347 / [Ne V] \lambda \lambda 3427$ ,  $[O I] \lambda \lambda 6302 / [S III] \lambda \lambda 6312$  potentially blended lines. We estimate that out of a possible 770 additional line detections, at most 4% of the FIGS spectra show tentative evidence of emission at these additional wavelengths ( $> 2\sigma$  detection in objects that are not overly extended with  $FWHM < 40 \text{ \AA}$ ). The measured line fluxes of these additional emission lines are between  $5 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$  and  $2.5 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ .

### 3.7. Completeness and Survey Area

The EM2D method is a relatively complicated method, and it is important to quantify its ability to detect emission lines. We determined the completeness function separately for each of the four FIGS fields, and for the combined flux of the  $H\alpha$   $\lambda 6563 + [N II] \lambda \lambda 6548, 6584$  lines, the flux of the  $[O III] \lambda 5007$  line, and for the combined flux of the  $[O II] \lambda \lambda 3726, 3729$  emission lines. Emission-line regions were added randomly within individual objects in the field, and used to create dispersed simulations of emission line regions with emission lines of



**Figure 14.** Spectra of two faint galaxies with large EW [O III] lines. The host galaxies GN1 2407 and GS2 1772 have measured broadband F105W magnitudes of  $m_{AB} = 27.29$  and  $m_{AB} = 26.25$  and line fluxes of  $(1.46 \pm 0.22) \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$  and  $(4.34 \pm 0.32) \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ , respectively. The measured rest-frame equivalent widths of the [O III] 5007 Å lines are  $254 \text{ Å} \pm 63 \text{ Å}$  and  $482 \text{ Å} \pm 69 \text{ Å}$ , respectively.

varying observed fluxes and at varying observed wavelengths. These simulated data were then processed using the EM2D method like the original FIGS data. These simulations showed that our ability to detect emission lines in the dispersed images was strongly dependent on the contrast ratio between the simulated line flux and the dispersed broadband flux of the host galaxy at the location of the emission-line region. We define the contrast ratio  $C = \frac{F(\text{line})}{F(\text{cont.})}$  as the ratio between the flux

produced in the emission line region, assumed to be from one strong emission line such as [O II], [O III] or  $\text{H}\alpha$   $\lambda\lambda 6563 + [\text{N II}] \lambda\lambda 6548, 6584$ , and the local broadband flux of the host galaxy.

The fraction of emission lines that we recovered, using the same criteria as listed in Section 3.1.2, was tabulated as a function of  $C$ . This process was carried out 200 emission lines at a time, and a total of 40,000 emission lines were generated using this process. Figure 12 shows a plot of the completeness function for the FIGS survey as a function of  $C$ . We also used emission-line simulations to determine the effective area of each of the FIGS fields. Since the G102 grism observations are rotated and offset with respect to each other, the exact area of overlap is a function of the different PAs at which a FIGS field was observed. To quantify this process, we generated a sample of bright emission line, well within the wavelength range of the G102 grism, both in and outside each of each of the FIGS fields of view. We then kept track of the locations on the sky from which emission lines were detected using the EM2D method using the same selection criteria used for the actual FIGS data. This process allowed us to estimate that the effective areas of the four FIGS grism fields are 4.66, 4.69, 4.62, and 4.21  $\text{arcmin}^2$  for the GN1, GN2, GS1, and GS2 fields, respectively. Note that the smaller effective size of the GS2 field is a direct result of a very bright source being at the center of the field.

## 4. Results

### 4.1. “Naked” Emission Lines

EM2D identifies emission lines independently of their host galaxies. We note that in every single case, we were able to easily identify an EM2D emission region with its host galaxy using nothing more than the SeXtractor segmentation map information (Bertin & Arnouts 1996). We found no occurrence of a confirmed emission line region that was not formally within the segmentation map area of a galaxy. As we show in Figure 13, we did, however, identify emission line regions with

very large EWs that are located within faint host galaxies. Examples of such sources are the two galaxies GN1 2407 and GS2 1772 shown in Figure 14. These objects have bright [O III] emission lines with observed fluxes of  $1.5 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$  and  $4.39 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ , respectively. The continuum sources, with  $m_{AB} = 27.29$  and  $m_{AB} = 26.25$  in the WFC3 F105W filter, are estimated to be at the redshifts of  $z = 1.18$  and  $z = 1.17$ , respectively. We estimate these emission lines to have rest-frame equivalent widths (EWs) of 254 Å and 482 Å, respectively. At the lower redshifts of  $0.112 < z < 0.36$  and  $z < 0.05$ , Cardamone et al. (2009) and Yang et al. (2017) have respectively identified strong SFGs known as Blueberries and Green Peas. Based solely on an [O III] EW that is larger than 300 Å, we identify 5 strongly star-forming regions in the EM2D FIGS galaxies at  $0.77 < z < 1.29$  out of a sample of 116 [O III] emission-line regions.

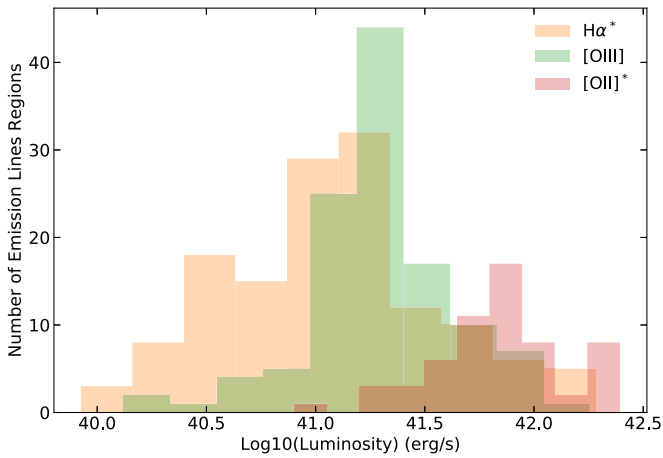
In terms of ELGs, we find five [O III] galaxies with at least one EW  $> 300 \text{ Å}$  star-forming region out of 96 galaxies, 5% of our sample. For comparison, we find that out of  $\approx 500,000$  SFGs at  $z < 0.6$  in SDSS data release DR10 with emission-line properties extracted by the Portsmouth group reported in the SDSS EmissionLinesPort table, 7% of the galaxies have [O III] line rest-frame EW  $> 300 \text{ Å}$ .

### 4.2. High-redshift Sources

The EM2D method successfully identified one  $\text{Ly}\alpha$  emitter. This object (GS2 1406), while already known (Larson et al. 2018) was independently detected using the EM2D method. The  $\text{Ly}\alpha$  emission was blindly detected in five PAs, and associated with object GS2 1406. Based on the EM2D method and assuming that the emission line is  $\text{Ly}\alpha$ , we estimate the redshift of this source to be  $7.464^{+0.004}_{-0.006}$ . We measure a  $\text{Ly}\alpha$  line flux of  $1.58^{+0.6}_{-0.2} \times 10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1}$ , and a very high EW of  $172 \text{ Å} \pm 46 \text{ Å}$ , which is consistent with the results shown in Larson et al. (2018). A second known high-redshift source at  $z = 7.51$  (GN1 1292, Tilvi et al. 2016) only shows emission lines in two position angles and is therefore not included in our FIGS EM2D sample, which required an emission line to be detected in at least three PAs (See Section 3.1.2).

### 4.3. Line Luminosities and Luminosity Functions

Figure 15 shows the distributions of  $\text{H}\alpha$   $\lambda 6563 + [\text{N II}] \lambda\lambda 6548, 6584$  lines, [O III]  $\lambda 5007$  and the [O II]  $\lambda\lambda 3726, 3729$



**Figure 15.** Distribution of the luminosities of individual emission line regions. The luminosities we show are for the [O II] $\lambda\lambda 3726,3729$  doublet ([O II]\*), the sum of the H $\alpha$   $\lambda 6563$ + [N II] $\lambda\lambda 6548,6584$  lines (H $\alpha^*$ ), and the [O III] $\lambda$  5007 line.

emission line luminosities. Combining our line catalog and the completeness, we derive the luminosity function of the H $\alpha$   $\lambda 6563$ + [N II] $\lambda\lambda 6548,6584$  [O III] $\lambda$  5007 and [O II] $\lambda\lambda 3726,3729$  emitters. The luminosity function was determined using the  $1/V_{\max}$  method, which we can express as:

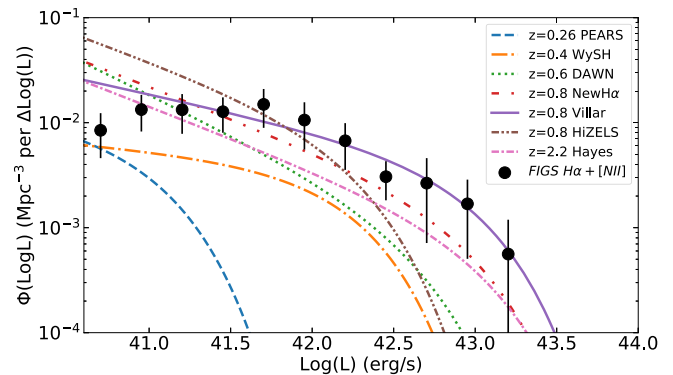
$$\Phi(\log L_i) = \frac{1}{\Delta \log L} \sum_j \frac{1}{V_j}, \quad (1)$$

where  $\Delta \log L$  is the logarithmic bin width of the luminosity function, and  $V_j$  is the maximum volume within which emission line  $j$  at a redshift of  $z_j$  would still be included in our sample. The volume  $V_j$ , when accounting for the completeness of our sample (Section 3.7), can be computed as

$$V_j = \frac{\Omega}{4\pi} \int_{z_{j,\min}}^{z_{j,\max}} R(\lambda) f(C) \frac{dV_c(z)}{dz} dz. \quad (2)$$

Here,  $\Omega$  is the solid angle of our survey (in sr),  $V_c$  is the comoving volume element at redshift  $z$ ,  $f(C)$  is the completeness function that we described above, and  $R(\lambda)$  is the normalized sensitivity function of the G102 grism. The latter takes a value of zeros outside of the bandpass of the G102 grisms, so that the values of  $z_{j,\min}$  and  $z_{j,\max}$  can therefore be taken to be 0 and  $\infty$ , respectively.

The distribution of observed, non-dust-corrected line region luminosities is shown in Figure 15. We performed a systematic search for 2D emission lines using ACS on *HST* as part of the previous program PEARS. The use of the G102 grism by the FIGS survey, however, extends the redshift ranges to an epoch that is closer to the peak of star formation with redshift ranges of  $0.3 < z < 0.72$ ,  $0.70 < z < 1.26$ , and  $1.28 < z < 2.0$  for H $\alpha$   $\lambda 6563$ + [N II] $\lambda\lambda 6548,6584$ , [O III] $\lambda 5007$ , and [O II] $\lambda\lambda 3726,3729$ , respectively. The FIGS complements the PEARS survey and the two surveys overlap in the redshift ranges that they probe:  $0.3 < z < 0.45$ ,  $0.7 < z < 0.9$ , and  $1.28 < z < 1.54$  for the H $\alpha$   $\lambda 6563$ + [N II] $\lambda\lambda 6548,6584$ , [O III] $\lambda 5007$ , and [O II] $\lambda\lambda 3726,3729$  lines, respectively. The number of available FIGS sources within these redshift ranges is small (10, 18, and 5, respectively), but we can still compare the average volume densities derived from both surveys. We find that the mean volume densities agree well, as we estimate these to be  $0.0015 \pm 0.00034$ ,  $0.00055 \pm 0.000055$ ,  $0.00012 \pm 0.000035 \text{ Mpc}^{-3}$  for the FIGS survey versus  $0.0016 \pm$



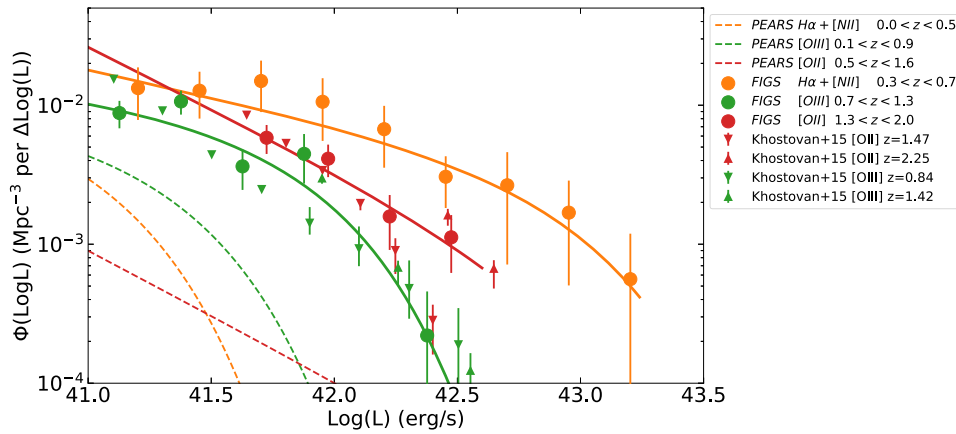
**Figure 16.** FIGS H $\alpha$   $\lambda 6563$ + [N II] $\lambda\lambda 6548,6584$  emission-line luminosity function at  $z = 0.75$ , completeness-corrected and extinction-corrected. We also show the H $\alpha$  luminosity function from the DAWN survey at  $z = 0.6$  from A. Gonzalez et al. (2018, in preparation) and references therein.

$0.00028$ ,  $0.00035 \pm 0.000069$ ,  $0.000099 \pm 0.000024 \text{ Mpc}^{-3}$  for the PEARS surveys, for the H $\alpha$   $\lambda 6563$ + [N II] $\lambda\lambda 6548,6584$ , [O III] $\lambda$  5007, and [O II] $\lambda\lambda 3726,3729$  lines, respectively.

We can compare our H $\alpha$  observations to previous ground-based studies of H $\alpha$  emission lines such as Pirzkal et al. (PEARS  $z = 0.26$ , 2013), Dale et al. (WySH  $z = 0.40$ , 2010), A. Gonzalez (DAWN  $z = 0.62$ , 2018, in preparation), Villar et al. ( $z = 0.84$ , 2008), Ly et al. (NEWH $\alpha$   $z = 0.84$ , 2011), Sobral et al. (HiZELS  $z = 0.84$ , 2013), and Hayes et al. ( $z = 2.2$ , 2010).

In order to do so, we must recompute the luminosity function for the FIGS H $\alpha$  galaxies by taking into account the total integrated luminosity for each individual galaxy. We can estimate the total integrated H $\alpha$   $\lambda 6563$ + [N II] $\lambda\lambda 6548,6584$  flux by simply measuring it from the FIGS spectra of these EM2D-selected objects, which we have described in Pirzkal et al. (2017). This measured line flux is a better estimate of the integrated flux for the entire object in cases of diffuse emission, or when multiple emission-line regions are present. We applied a completeness correction factor based on the value of the brightest emission line region. We also added the luminosity-dependent dust attenuation law from Hopkins et al. (2001). The result is shown in Figure 16. As this figure shows, this result is most consistent with the results of Villar et al. (2008), despite the widely different selection method and survey areas and less consistent with the results of Sobral et al. (2013), as we detect a larger number of high-luminosity sources.

The luminosity functions for the H $\alpha$   $\lambda 6563$ + [N II] $\lambda\lambda 6548,6584$ , [O III] $\lambda$  5007, and [O II] $\lambda\lambda 3726,3729$  ELGs are shown in Figure 17. This figure was derived by estimating the total integrated H $\alpha$   $\lambda 6563$ + [N II] $\lambda\lambda 6548,6584$  [O III], and [O II] observed flux from each host galaxy and is meant to show the observed densities of these targets, accounting only for instrumental completeness and not applying dust correction. For comparison, we also plot the luminosity functions from Pirzkal et al. (2013), derived for ELGs selected in a similar manner but at lower redshifts. While the two surveys show similar object densities over the overlapping redshift ranges of the PEARS and FIGS surveys, the FIGS survey detected significantly brighter line emission. This is consistent with stronger star formation in galaxies at an epoch that is closer to the peak of star formation history at  $1.5 < z < 2.5$ . We also show measurements from Khostovan et al. (2015) and show good agreement with these observations too.



**Figure 17.** Luminosity functions derived for the FIGS survey for  $H\alpha$   $\lambda 6563 + [N II] \lambda \lambda 6548, 6584$ ,  $[O III] \lambda 5007$ , and  $[O II] \lambda \lambda 3726, 3729$ . We also plot the luminosity functions from Pirzkal et al. (2013), derived at lower redshifts for the PEARS project (dash lines), and from Khostovan et al. (2015; triangle symbols)

## 5. Conclusion

In order to study star formation within a wide range of galaxies, one first needs to establish a proper sample of objects, with as little selection bias as possible. Being able to detect multiple emission-line regions within galaxies, or detect extended and diffuse emission-line regions, was the goal of this paper.

We showed how we used our EM2D technique to generate a catalog of ELGs over a wide range of redshifts by applying this method to the very deep WFC3 G102 slitless spectroscopic data obtained from the FIGS survey. There are several advantages to using the EM2D method. First, emission lines are detected independently of their host galaxies. This allows for the detection of emission lines with large EWs. Second, the exact location (to within about 0.5 WFC3 pixel, or  $0''.06$ ) of the source of the emission line can be determined, yielding both more accurate wavelength estimates and better spectroscopic redshifts. Third, strong multiple emission lines, such as  $[O II] \lambda \lambda 3726, 3729$ ,  $[O III] \lambda \lambda 4959, 5007 + H\beta \lambda 4861$ , and  $H\alpha \lambda 6563 + [N II] \lambda \lambda 6548, 6584$  are detected independently, providing a check on the self-consistency of the emission-line region detections. Fourth, multiple emission-line regions, or diffuse extended emission-line regions, can be identified. Building on this selection process, we have shown how 2D emission line maps can be created. A detailed modeling of the physical parameters of these galaxies will be presented in the next paper of the series. We found that a large fraction of FIGS galaxies show signs of star formation occurring in multiple regions. We also identified a significant fraction of these objects to have extended or diffuse star formation, as opposed to bulge-dominated emission. We did not identify naked emission lines without continuum, but identified several very high EW sources, including one  $Ly\alpha$  source at  $z = 7.5$ . Overall, approximately 20% of our  $[O III]$  sample has a star-forming region with  $EW > 300 \text{ \AA}$ . Finally, we showed that to derive accurate spectroscopic redshifts using slitless observations, one should rely on observations taken at multiple position angles and on methods such as EM2D. This should prove relevant to the planning of observations using future missions using slitless spectroscopic modes such as *JWST* (NIRISS and NIRCAM) and WFIRST.

This work is based on observations made with the NASA/ESA *Hubble Space Telescope*, obtained [from the Data Archive] at the Space Telescope Science Institute, which is

operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with program #13779.

Support for program #13779 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555.

E.C.L., J.C., and S.C. acknowledge support from the European Research Council (ERC) via an Advanced Grant under grant agreement no. 321323- NEOGAL. The authors would like to thank R.T. Gatto for discussions that provided invaluable assistance in developing the themes presented in this manuscript. A.C. acknowledges the grants ASI n.I/023/12/0 “Attivit relative alla fase B2/C per la missione Euclid” and PRIN MIUR 2015 “Cosmology and Fundamental Physics: illuminating the Dark Universe with Euclid.”

## ORCID iDs

Barry Rothberg <https://orcid.org/0000-0003-2283-2185>  
 Sangeeta Malhotra <https://orcid.org/0000-0002-9226-5350>  
 James Rhoads <https://orcid.org/0000-0002-1501-454X>  
 Norman Grogin <https://orcid.org/0000-0001-9440-8872>  
 Stephane Charlot <https://orcid.org/0000-0003-3458-2275>  
 Steven L. Finkelstein <https://orcid.org/0000-0001-8519-1130>  
 Anton M. Koekemoer <https://orcid.org/0000-0002-6610-2048>  
 Parviz Ghavamian <https://orcid.org/0000-0002-9886-0839>  
 François Hammer <https://orcid.org/0000-0002-2165-5044>  
 Rebecca L. Larson <https://orcid.org/0000-0003-2366-8858>  
 Lise Christensen <https://orcid.org/0000-0001-8415-7547>  
 Andrea Cimatti <https://orcid.org/0000-0002-4409-5633>  
 Nimish P. Hathi <https://orcid.org/0000-0001-6145-5090>  
 Bhavin Joshi <https://orcid.org/0000-0002-7593-8584>  
 Robert W. O’Connell <https://orcid.org/0000-0002-8190-7573>  
 John Pharo <https://orcid.org/0000-0003-2249-2539>  
 Darach Watson <https://orcid.org/0000-0002-4465-8264>  
 Rogier A. Windhorst <https://orcid.org/0000-0001-8156-6281>  
 Nadia L Zakamska <https://orcid.org/0000-0001-6100-6869>

## References

- Aller, L. H. 1942, *ApJ*, **95**, 52
- Atek, H., Malkan, M., McCarthy, P., et al. 2010, *ApJ*, **723**, 104
- Avila, R. J., Hack, W., Cara, M., et al. 2015, in ASP Conf. Ser. 495, *Astronomical Data Analysis Software and Systems XXIV*, ed. A. R. Taylor & E. Rosolowsky (San Francisco, CA: ASP), 281
- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
- Boroson, T. A., Salzer, J. J., & Trotter, A. 1993, *ApJ*, **412**, 524
- Cardamone, C., Schawinski, K., Sarzi, M., et al. 2009, *MNRAS*, **399**, 1191
- Chevallard, J., & Charlot, S. 2016, *MNRAS*, **462**, 1415
- Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, *AJ*, **112**, 839
- Dale, D. A., Barlow, R. J., Cohen, S. A., et al. 2010, *ApJL*, **712**, L189
- Djorgovski, S., Spinrad, H., McCarthy, P., & Strauss, M. A. 1985, *ApJL*, **299**, L1
- Ester, M., Kriegel, H. P., Sander, J., & Xu, X. 1996, in Proc. Second Int. Conf. Knowledge Discovery and Data Mining AAAI Press, *A Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise* (New York: ACM), 226
- Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJL*, **600**, L93
- Hayes, M., Schaerer, D., & Östlin, G. 2010, *A&A*, **509**, L5
- Hopkins, A. M. 2004, *ApJ*, **615**, 209
- Hopkins, A. M., Connolly, A. J., Haarsma, D. B., & Cram, L. E. 2001, *AJ*, **122**, 288
- Kennicutt, R. C., Jr. 1998, *ARA&A*, **36**, 189
- Kewley, L. J., & Dopita, M. A. 2002, *ApJS*, **142**, 35
- Khostovan, A. A., Sobral, D., Mobasher, B., et al. 2015, *MNRAS*, **452**, 3948
- Kümmel, M., Walsh, J. R., Pirzkal, N., Kuntschner, H., & Pasquali, A. 2009, *PASP*, **121**, 59
- Larson, R. L., Finkelstein, S. L., Pirzkal, N., et al. 2018, *ApJ*, **858**, 94
- Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, *ApJS*, **228**, 2
- Ly, C., Lee, J. C., Dale, D. A., et al. 2011, *ApJ*, **726**, 109
- Madau, P., & Dickinson, M. 2014, *ARA&A*, **52**, 415
- Madau, P., Pozzetti, L., & Dickinson, M. 1998, *ApJ*, **498**, 106
- Markarian, B. E. 1967, *Afz*, **3**, 55
- Mayall, N. U. 1936, *PASP*, **48**, 14
- McCarthy, P. J., Yan, L., Freudling, W., et al. 1999, *ApJ*, **520**, 548
- McGaugh, S. S. 1991, *ApJ*, **380**, 140
- Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, *ApJS*, **225**, 27
- Nagao, T., Maiolino, R., & Marconi, A. 2006, *A&A*, **459**, 85
- Oke, J. B., & Gunn, J. E. 1983, *ApJ*, **266**, 713
- Osterbrock, D. E., & Ferland, G. J. 2006, in *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*, ed. D. E. Osterbrock & G. J. Ferland (2nd ed.; Sausalito, CA: Univ. Science Books), 2006
- Pasquali, A., Pirzkal, N., Walsh, J. R., et al. 2003, in *Astronomy, Cosmology and Fundamental Physics*, ed. P. A. Shaver, L. DiLella, & A. Giménez (Berlin: Springer), 471
- Pirzkal, N., Burgasser, A. J., Malhotra, S., et al. 2009, *ApJ*, **695**, 1591
- Pirzkal, N., Malhotra, S., Ryan, R. E., et al. 2017, *ApJ*, **846**, 84
- Pirzkal, N., Pasquali, A., & Demleitner, M. 2001, *STECF*, **29**, 5
- Pirzkal, N., Rothberg, B., Ly, C., et al. 2013, *ApJ*, **772**, 48
- Pirzkal, N., Ryan, R., & Brammer, G. 2016, *Trace and Wavelength Calibrations of the WFC3 G102 and G141 IR Grisms*, Inst. Sci. Rep., **WFC3 2016-15**
- Pirzkal, N., Xu, C., Malhotra, S., et al. 2004, *ApJS*, **154**, 501
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, *A&A*, **571**, A16
- Rubin, V. C., Ford, W. K., Jr., & D'Odorico, S. 1970, *ApJ*, **160**, 801
- Rubin, V. C., Krishna Kumar, C., & Ford, W. K., Jr. 1972, *ApJ*, **177**, 31
- Smith, M. G. 1975, *ApJ*, **202**, 591
- Sobral, D., Smail, I., Best, P. N., et al. 2013, *MNRAS*, **428**, 1128
- Storey, P. J., & Zeppen, C. J. 2000, *MNRAS*, **312**, 813
- Straughn, A. N., Meurer, G. R., Pirzkal, N., et al. 2008, *AJ*, **135**, 1624
- Straughn, A. N., Pirzkal, N., Meurer, G. R., et al. 2009, *AJ*, **138**, 1022
- Tilvi, V., Pirzkal, N., Malhotra, S., et al. 2016, *ApJL*, **827**, L14
- Treu, T., Schmidt, K. B., Brammer, G. B., et al. 2015, *ApJ*, **812**, 114
- Villar, V., Gallego, J., Pérez-González, P. G., et al. 2008, *ApJ*, **677**, 169
- Villforth, C., Koekemoer, A. M., & Grogin, N. A. 2010, *ApJ*, **723**, 737
- Xia, L., Malhotra, S., Rhoads, J., et al. 2011, *AJ*, **141**, 64
- Xu, C., Pirzkal, N., Malhotra, S., et al. 2007, *AJ*, **134**, 169
- Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2016, *ApJS*, **224**, 15
- Yang, H., Malhotra, S., Rhoads, J. E., & Wang, J. 2017, *ApJ*, **847**, 38