

An $H\alpha$ view of galaxy buildup in the first 2 Gyr: Luminosity functions at $z \sim 4\text{--}6.5$ from NIRCam/grism spectroscopy

Alba Covelopaz^{1,*}, Emma Giovino¹, Pascal A. Oesch^{1,2,3}, Romain A. Meyer¹, Andrea Weibel¹, Gabriel Brammer^{2,3}, Yoshinobu Fudamoto⁴, Josephine Kerutt⁵, Jamie Lin⁶, Jasleen Matharu^{2,3}, Rohan P. Naidu^{7,**}, Anna Velichko^{1,8}, Victoria Bollo⁹, Rychard Bouwens¹⁰, John Chisholm¹¹, Garth D. Illingworth¹², Ivan Kramarenko¹³, Daniel Magee¹², Michael Maseda¹⁴, Jorrry Matthee¹³, Erica Nelson¹⁵, Naveen Reddy¹⁶, Daniel Schaerer¹, Mauro Stefanon^{17,18}, and Mengyuan Xiao¹

¹ Department of Astronomy, University of Geneva, Chemin Pegasi 51, 1290 Versoix, Switzerland

² Cosmic Dawn Center (DAWN), Copenhagen, Denmark

³ Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark

⁴ Center for Frontier Science, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan

⁵ Kapteyn Astronomical Institute, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands

⁶ Department of Physics and Astronomy, Tufts University, 574 Boston Avenue, Medford, MA 02155, USA

⁷ MIT Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Ave., Cambridge, MA 02139, USA

⁸ Institute of Astronomy, Kharkiv National University, 4 Svobody Sq., Kharkiv 61022, Ukraine

⁹ European Southern Observatory, Karl-Schwarzschildstrasse 2, D-85748 Garching bei München, Germany

¹⁰ Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands

¹¹ Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712-1205, USA

¹² Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA

¹³ Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400 Klosterneuburg, Austria

¹⁴ Department of Astronomy, University of Wisconsin–Madison, 475 N. Charter St., Madison, WI 53706, USA

¹⁵ Department for Astrophysical and Planetary Science, University of Colorado, Boulder, CO 80309, USA

¹⁶ Department of Physics and Astronomy, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA

¹⁷ Departament d’Astronomia i Astrofísica, Universitat de València, C. Dr. Moliner 50, E-46100 Burjassot, València, Spain

¹⁸ Unidad Asociada CSIC “Grupo de Astrofísica Extragaláctica y Cosmología” (Instituto de Física de Cantabria – Universitat de València), Valencia, Spain

Received 24 September 2024 / Accepted 14 January 2025

ABSTRACT

The $H\alpha$ nebular emission line is an optimal tracer for recent star formation in galaxies. With the advent of *JWST*, this line has recently become observable at $z > 3$ for the first time. We present a catalog of 1050 $H\alpha$ emitters at $3.7 < z < 6.7$ in the GOODS fields obtained from a blind search in *JWST* NIRCam/grism data. We made use of the FRESCO survey’s 124 arcmin² of observations in GOODS-North and GOODS-South with the F444W filter, probing $H\alpha$ at $4.9 < z < 6.7$, and the CONGRESS survey’s 62 arcmin² of observations in GOODS-North with F356W, probing $H\alpha$ at $3.8 < z < 5.1$. We found an overdensity with 98 sources at $z \sim 4.4$ in GOODS-N, and confirmed previously reported overdensities at $z \sim 5.2$ in GOODS-N and at $z \sim 5.4$ and $z \sim 5.9$ in GOODS-S. We computed the observed $H\alpha$ luminosity functions (LFs) in three bins centered at $z \sim 4.45$, 5.30, and 6.15, which are the first such measurements at $z > 3$ obtained based purely on spectroscopic data, robustly tracing galaxy star formation rates (SFRs) beyond the peak of the cosmic star formation history. We compared our results with theoretical predictions from three different simulations and found good agreement at $z \sim 4\text{--}6$. The UV LFs of this spectroscopically confirmed sample are in good agreement with pre-*JWST* measurements obtained with photometrically selected objects. Finally, we derived SFR functions and integrated them to compute the evolution of the cosmic SFR densities across $z \sim 4\text{--}6$, finding values in good agreement with recent UV estimates from Lyman-break galaxies, which imply a continuous decrease in SFR density by a factor of three over $z \sim 4$ to $z \sim 6$. Our work shows the power of NIRCam grism observations to efficiently provide new tests for early galaxy formation models based on emission line statistics.

Key words. galaxies: evolution – galaxies: formation – galaxies: high-redshift – galaxies: luminosity function, mass function – galaxies: star formation

1. Introduction

Uncovering the assembly histories of galaxies throughout cosmic time is a primary objective of modern astrophysics. In the early Universe, the star formation rate density (SFRD) increases with time up to a peak at $z \sim 2\text{--}3$, at the so-called cosmic noon (Madau & Dickinson 2014). Probing

the SFRD at high redshifts is essential to shedding light on the process of early galaxy assembly. However, until recently the star formation rate (SFR) estimates at $z > 3$ almost exclusively relied on rest-frame UV continuum estimates, which involved an uncertain dust correction (e.g., Wyder et al. 2005; Schiminovich et al. 2005; Dahlen et al. 2007; Reddy & Steidel 2009; Robotham & Driver 2011; Cucciati et al. 2012; Bouwens et al. 2012a,b; Schenker et al. 2013; Bouwens et al. 2015).

* Corresponding author; alba.covelopaz@unige.ch

** NASA Hubble Fellow.

At lower redshifts, the $H\alpha$ emission line is among the most commonly used estimators for galaxy star formation (e.g., Erb et al. 2006; Geach et al. 2008; Förster Schreiber et al. 2009; Hayes et al. 2010; Sobral et al. 2013). This line is the result of hydrogen gas recombination in the presence of ionizing photons emitted by the most massive stars in a galaxy. These stars live for a short period of <10 Myr, making $H\alpha$ an optimal tracer for recent star formation in a galaxy (e.g., Kennicutt 1998). Moreover, $H\alpha$ is less affected by dust attenuation than the UV continuum. Multiple studies have used $H\alpha$ measurements to determine star formation densities at various redshifts (e.g., Stroe & Sobral 2015 at $0.19 < z < 0.23$, Sobral et al. 2013 at $0.4 < z < 2.3$, Terao et al. 2022 at $2.1 < z < 2.5$, Bollo et al. 2023 at $3.86 < z < 4.94$). With ground-based near-IR spectrographs such as MOS-FIRE and KMOS, spectroscopic detections of $H\alpha$ have become standard at $1.4 < z < 2.6$ (e.g., Kashino et al. 2013; Steidel et al. 2014; Kriek et al. 2015).

With the recent advent of *JWST* (Gardner et al. 2006; Rigby et al. 2023), the $H\alpha$ line is now accessible for high-resolution spectroscopy and imaging at $z > 2.5$, and some studies have already shown its capabilities (e.g., Nelson et al. 2024; Matharu et al. 2024; Pirie et al. 2024). This calls for a blind luminosity-limited search of $H\alpha$ emitters in order to provide a stringent estimate of the star formation rates in a well-defined and complete sample. In this paper, we present the results of a blind search of $H\alpha$ emitters at $3.8 < z < 6.6$ in the GOODS fields. For this, we used the *JWST* First Reionization Epoch Spectroscopically Complete Observations (FRESCO; Cycle 1, Oesch et al. 2023) and the Complete NIRCcam Grism Redshift Survey (CONGRESS; Cycle 2, Egami et al. 2023) programs. FRESCO obtained NIRCcam imaging and slitless spectroscopy with the F444W filter in an area of 62 arcmin^2 in GOODS-North and 62 arcmin^2 in GOODS-South. CONGRESS used the same configuration as FRESCO and covered the same 62 arcmin^2 in GOODS-North with the F356W filter. We use this new catalog of spectroscopically confirmed $H\alpha$ emitters to obtain the $H\alpha$ luminosity functions (LFs) at $z > 2.5$ derived from spectroscopic data for the first time and derive new cosmic SFRD measurements at redshifts $z \sim 4.45, 5.3$, and 6.15 .

This paper is structured as follows. In Sect. 2 we describe our data reduction to obtain the full catalog of $H\alpha$ emitters. In Sect. 3 we detail the completeness characterization of this catalog. In Sect. 4 we obtain the $H\alpha$ LFs, star formation rate functions, and SFRDs at $z \sim 4.45, 5.3$, and 6.15 . In Sect. 5 we discuss our findings and compare them to previous studies at redshifts $z = 0-10$. Throughout this work, we adopt a concordance flat Λ CDM cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$. Magnitudes are listed in the AB system (Oke & Gunn 1983).

2. Data

2.1. Observations

We used NIRCcam/grism spectra with the F444W filter from FRESCO (Oesch et al. 2023) and the F356W filter from CONGRESS (Sun, Egami et al., in prep., Egami et al. 2023). Both surveys observed 62 arcmin^2 of the GOODS-North field, and their combined spectra cover the wavelength range $3.1-5.0 \mu\text{m}$. Moreover, FRESCO observed 62 arcmin^2 of the GOODS-South field, covering $3.8-5.0 \mu\text{m}$. The combination of FRESCO and CONGRESS data allowed us to perform a blind search of $H\alpha$ emitters in the redshift range of $3.7 < z < 6.7$ in GOODS-North and $4.9 < z < 6.7$ in GOODS-South.

For our photometric catalogs in the two GOODS fields, we used all the publicly available Hubble Space Telescope (HST; e.g., Giavalisco et al. 2004; Grogin et al. 2011; Koekemoer et al. 2011) and *JWST* (e.g., Eisenstein et al. 2023a,b; Williams et al. 2023) imaging in the MAST archive¹. The basic procedure applied to derive the photometric catalogs is described in Weibel et al. (2024). To ensure a homogeneous depth for the detection image across the FRESCO survey footprint, we generated custom mosaics of the F444W and F210M images using imaging from FRESCO only and excluded any available data from other surveys in these filters. These reductions were combined in an inverse-variance weighted stack to obtain a detection image with a typical 5σ depth of 28.5 mag in $0.16''$ radius circular apertures. We ran SourceExtractor (Bertin & Arnouts 1996) in dual mode to measure fluxes through circular apertures with a radius of $0.16''$ in all available filters, including data from all available surveys. Prior to measuring the fluxes, all images were matched to the point spread function (PSF) resolution in the F444W filter. We scaled the aperture fluxes to the fluxes measured through Kron apertures on a PSF-matched version of the detection image, and then to total fluxes by accounting for the encircled energy of the Kron ellipse on the F444W PSF.

2.2. Grism extractions

The NIRCcam/grism data were reduced and processed with the publicly available *grizli* code², using the standard CRDS grism dispersion files from pmap 1123 made available in September 2023 and slightly modified sensitivity functions from *grizli*³ (see also, e.g., Oesch et al. 2023; Matharu et al. 2024; Meyer et al. 2024). Given the limited wavelength range of the grism spectra, sources with $H\alpha$ emission lines often do not show any other strong line in the F444W filter (except for sources with additional F356W grism spectra in GOODS-N). Additionally, given the single dispersion angle of the FRESCO data, the association of a given emission line with an individual galaxy along the dispersion direction is not always unique. It is therefore not straightforward to blindly search for $H\alpha$ emitters. Instead, we had to rely on initial redshift guesses, such as an existing spectroscopic redshift or a photometric redshift estimate, to search for $H\alpha$ emission lines in galaxies. Therefore, we extracted grism spectra only for sources, for which $H\alpha$ lines are expected to lie in the grism wavelength range with a non-zero probability. In particular, our extractions are based on several criteria: (1) sources with eazy photometric redshifts at $z_{\text{phot}} = 4.85-6.7$ for the F444W FRESCO observations or $z_{\text{phot}} = 3.7-5.2$ for the F356W CONGRESS field and with photometric redshift uncertainties of less than 20% and with $m_{444} < 28 \text{ AB mag}$. (2) We additionally extracted all sources with ancillary spectroscopic redshifts in the same ranges stated above.

We then run a line search with *grizli* over the 2.5 to 97.5 percentiles of the redshift probability density functions as measured with eazy. For sources with spectroscopic redshifts, we search within $0.02 \times (1 + z_{\text{spec}})$ of the spectroscopic redshift, which corresponds to velocity offsets of up to $\sim 6000 \text{ km s}^{-1}$.

¹ This includes partial or full coverage by the ACS F435W, F606W, F775W, F814W, and F850LP filters; the WFC3 F105W, F125W, F140W, and F160W filters; and the NIRCcam F090W, F115W, F150W, F182M, F200W, F210M, F277W, F335M, F356W, F410M, and F444W filters in both fields. The GOODS-S field has additional partial coverage by the NIRCcam F162M, F250M, F300M, F430M, F460M, and F480M as well as the MIRI F560W and F770W filters.

² <https://github.com/gbrammer/grizli>

³ Available at <https://s3.amazonaws.com/grizli-v2/JWSTGrism/NircamSensitivity/index.html>

2.3. Selection of $H\alpha$ emitters

After the line search, we selected the sources with $H\alpha$ S/N > 3 from the *grizli*-extracted catalog. Upon visual inspection, we found several multi-component sources presenting either very nearby neighbors or clumpy disks that were split into different components during the *SourceExtractor* run. As a result, several sources in our catalog were clumps from the same galaxy, which were assigned the same $H\alpha$ emission line. To avoid accounting for the same galaxy more than once, we adopted a similar procedure to [Matthee et al. \(2023\)](#): we grouped sources with identical grism redshifts and an angular separation <0.6 arcsec, which corresponds to a physical distance of $\lesssim 3.3$ – 4.4 kpc at the studied redshift range ($3.8 < z < 6.6$). We then re-extracted the spectra of the multi-component objects as one source with *grizli* and re-computed the photometry and object positions as flux-weighted centroids. This led to a sample of 1081 sources in CONGRESS and 858 in FRESCO, where $\sim 10\%$ of these are multi-component galaxies.

Each source of this sample was then visually inspected by four team members using the tool *specvizitor*⁴. Each individual inspector assigned a quality flag to every source, following the classification in [Meyer et al. \(2024\)](#): a quality flag $q = 3$ was assigned to sources with a clear $H\alpha$ line and a matching morphology between the direct image and the 2D spectrum and line map; $q = 2$ to sources with a lower S/N, but a clear morphology match; $q = 1$ to unclear sources; $q = 0$ to objects lacking an $H\alpha$ line; and $q = -1$ to contamination by neighboring objects.

The FRESCO $H\alpha$ emitters in GOODS-North with $z > 5.25$ also show [O III] 5008 Å emission in CONGRESS data. This enables us to further check the quality of the 179 FRESCO sources that have a valid CONGRESS spectrum within the $H\alpha$ range, and we find that 36 sources have an improved quality flag after inspecting CONGRESS, while only five have been downgraded. The final quality flag of each source is the average of the four inspector grades. Figure 1 shows the extractions of four sources with different quality flags.

For the final $H\alpha$ catalog we selected sources with an overall quality flag $q \geq 1.75$, meaning that a majority of the team believed these lines to be likely real. The result is a catalog with 1050 $H\alpha$ emitters: 554 from CONGRESS (in GOODS-N) and 503 from FRESCO (318 in GOODS-N and 185 in GOODS-S), with seven sources in both surveys. Despite our original S/N cut of $S/N > 3$ for the parent sample, all the sources flagged with $q \geq 1.75$, turned out to have $S/N > 5$ in our final catalog.

2.4. Comparison to JADES NIRSpec catalogs

Of the sources in our $H\alpha$ catalog, 123 are also present in JADES DR3 ([D'Eugenio et al. 2024](#)), which contains prism and grating redshifts for targeted sources in both GOODS fields. After comparing our grism redshifts to JADES prism redshifts, we found these to agree very well, within a scatter of $0.001 \times (1 + z)$ and no outliers. We also compare our $H\alpha$ fluxes to those of JADES DR3 in Appendix B. The grism and NIRSpec line flux measurements are broadly consistent albeit with a significant dispersion of up to 0.26 dex. While the grism data measures the full (spatially resolved) emission line without flux loss (in principle), the NIRSpec line flux measurements need to be corrected for slit losses. It is likely that this correction as well as different relative calibrations introduce this dispersion in flux measurements. Overall, this comparison is very reassuring. A

more detailed comparison between the NIRSpec catalogs and the NIRCам grism line fluxes is deferred to our data release paper.

2.5. Removal of broad-line AGN

During the visual inspection process, we found 13 of our sources to show a significant broad component in their $H\alpha$ lines ($\text{FWHM} > 800 \text{ km s}^{-1}$) consistent with AGN activity. These sources are reported in Appendix C. Eight of them are in FRESCO and match those previously reported as Little Red Dots by [Matthee et al. \(2024\)](#). The other five are found at lower redshifts in the CONGRESS data and are new discoveries. These sources are removed from our sample of LF measurements as we are only interested in star-forming galaxies in this work. We note, however, that our results do not change significantly with or without these sources.

2.6. UV luminosities

In order to compare our results with a different star formation rate tracer, we derive UV luminosities (and absolute magnitudes) using the fluxes in our photometric catalog. For each source, we determine which filters lie in the rest-frame UV wavelength range, i.e., 1400–2700 Å. We fit these fluxes to $f_\nu \propto \lambda^{\beta-2}$ by doing a linear fit in log–log space, and then interpolate the flux value at 1500 Å. This $f_{1500, \text{UV}}$ is then converted to L_{UV} .

2.7. SED fitting and dust corrections

To account for dust attenuation, we fit the photometry of our sources using the *Bagpipes* SED fitting code ([Carnall et al. 2018](#)). We adopted a star formation history consisting of a delayed tau model and a recent burst, which we allowed to occur between 0 and 10 Myr before the time of observation. We used a ([Calzetti et al. 2000](#)) dust attenuation curve. The redshift was fixed to the grism spectroscopic redshift; however, we do not include the grism emission line fluxes in the fit. We allowed the visual attenuation to vary from $A_V = 0$ –5 mag, and the metallicity from $Z = 0.1 Z_\odot$ – $Z_\odot$, using uniform priors. Based on the resulting visual attenuation, A_V , we further derive the dust attenuation at the $H\alpha$ wavelength based on the [Calzetti et al. \(2000\)](#) curve, assuming that the nebular reddening is the same as the continuum reddening.

Of the sources in our $H\alpha$ catalog, 86 have Balmer decrements measured by JADES NIRSpec spectroscopy as released in DR3 ([D'Eugenio et al. 2024](#)). JADES obtained $H\alpha$ and $H\beta$ fluxes for targeted sources in both GOODS fields using NIRSpec G140M/G235M/G395M grating and prism spectroscopy. This allowed us to compare the $H\alpha$ dust attenuation, $A_{H\alpha}$, computed by *Bagpipes* SED fitting to that measured directly from the Balmer decrement, using the same dust attenuation law from [Calzetti et al. \(2000\)](#). We find that both methods provide $A_{H\alpha}$ results with large uncertainties; nevertheless, the median $H\alpha$ dust attenuation for the sample is consistent with both methods: $A_{H\alpha, \text{med}} = 0.47$ mag. Throughout this paper, we thus use the $H\alpha$ dust attenuation values from *Bagpipes* derived for each source.

3. Completeness

To accurately compute luminosity number densities, we need to account for the completeness of our sample, which depends on the continuum as well as the line fluxes and the position of the

⁴ <https://github.com/ivkram/specvizitor>

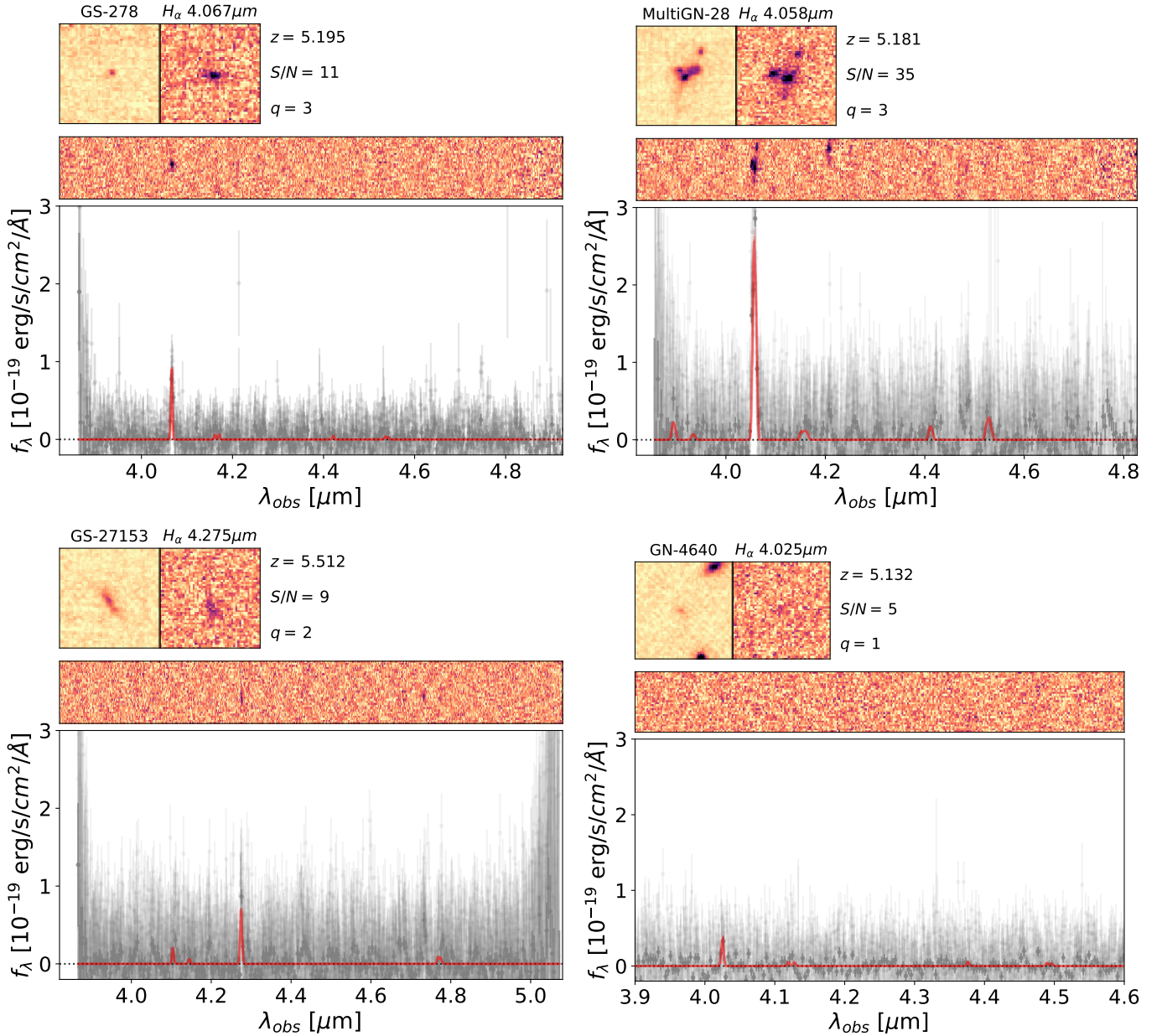


Fig. 1. Direct F444W stamps, H α maps, 2D spectra, and 1D spectra of four sources in the FRESCO H α catalog with different quality flags. Top left: single-component galaxy with $q = 3$. Top right: multi-component galaxy with $q = 3$. Bottom left: single-component with $q = 2$. Bottom right: single-component with $q = 1$. For the analysis of this paper, we only include sources with $q \geq 1.75$.

sources in the field. Following each step of our catalog creation, we identify four sources of incompleteness: photometric detection, 2D spectrum extraction with *grizli*, line detection with *grizli*, and visual inspection.

3.1. Photometric detection completeness

We assess the detection completeness of the photometric catalogs following the method described in Meyer et al. (2024), which uses the GaLaxy survey Completeness AlgoRithm 2 (GLACiAR2; Leethochawalit et al. 2022). GLACiAR2 performs injection-recovery simulations by creating synthetic galaxy profiles at a range of input magnitudes and morphological parameters, injecting them at random locations in the imaging mosaics and running SourceExtractor in the same way as for the real catalog to obtain the fraction of recovered sources as a func-

tion of the observed magnitude in the detection image, mag_{det} . As in Meyer et al. (2024), the synthetic galaxies that we generated follow a log-normal distribution centered at $R_{\text{eff}} = 0.8$ kpc, which was selected to fit the morphology of our photometric catalog in F444W and F210M. The completeness is well approximated by a sigmoid function. Thus, we fit a sigmoid function to the recovered sources as a function of mag_{det} , separately for each field, which yielded the photometric detection completeness functions, C_{det} , already reported in Meyer et al. (2024) separately for GN and GS:

$$C_{\text{det;GN}} = \frac{0.95}{1 + \exp[1.61 \times (\text{mag}_{\text{det}} - 28.71)]} \quad \text{and} \quad (1)$$

$$C_{\text{det;GS}} = \frac{0.93}{1 + \exp[1.66 \times (\text{mag}_{\text{det}} - 28.55)]}. \quad (2)$$

3.2. Extraction completeness

During the line search process with `grizli` described in Sect. 2.2, we only considered sources with photometric redshift uncertainties of less than 20%. This constitutes a source of incompleteness, as we did not extract sources with higher uncertainties, and some of these could still be $H\alpha$ emitters. As expected, most of the sources with photometric redshifts in the target range that were not extracted were among the faintest targets in the photometric catalog. To account for this, we considered all the sources with an easy photometric redshift whose redshift probability function overlaps significantly with the grism redshift sensitivity; i.e., we consider all the sources whose lower 16th percentile fulfills $z_{\text{phot},16} \geq 3.72$ (4.86) for CONGRESS (FRESCO), and $z_{\text{phot},84} \leq 5.15$ (6.69) for the upper 84th percentile. We binned these sources according to their detection magnitude and calculated the percentage of sources in each bin that were extracted. We then fitted this magnitude-dependent completeness again with a sigmoid function. The result is shown separately for CONGRESS and FRESCO in Fig. 2, and the sigmoid functions that model this extraction completeness in each field are

$$C_{\text{ext;CONGRESS}} = \frac{0.96}{1 + \exp[4.17 \times (\text{mag}_{\text{det}} - 27.82)]} \quad \text{and} \quad (3)$$

$$C_{\text{ext;FRESCO}} = \frac{0.94}{1 + \exp[2.70 \times (\text{mag}_{\text{det}} - 28.09)]}. \quad (4)$$

Both of these functions cross the 50% limit at about 28 mag.

3.3. Line detection completeness

Once the 2D spectrum of each source is extracted, we use `grizli` to identify emission lines and extract their properties. Our final catalog only contains sources with an $H\alpha$ flux $S/N > 5$, as we cannot reliably confirm that lower S/N detections are genuine $H\alpha$ emission lines rather than contamination. This leads to 1388 galaxies (662 in CONGRESS and 726 in FRESCO) that we are discarding, some of which might be genuine $H\alpha$ emitters.

To assess this line detection incompleteness, we simulate a catalog of ~ 4000 mock $H\alpha$ emitters for each field and survey in a uniform grid of X and Y pixel position, emission line fluxes, and redshift. At each position, we extract the full grism spectrum (i.e., including possible contamination) and add synthetic emission lines with point source morphology. We generate synthetic emission lines for each mock source: one line of $H\alpha$ (6564.59 Å), a doublet of [N II] (6549.84 Å, 6585.23 Å), and a doublet of [S II] (6718.32 Å, 6732.71 Å). The flux ratios for these five lines were set to 1 : 0.022 : 0.065 : 0.070 : 0.051. These were generated for each source in a grid of wavelength positions and $\log(H\alpha)$ fluxes, covering the full range of observable wavelengths (3.1–4.0 μm for CONGRESS; 3.9–5.0 μm for FRESCO) and observed $H\alpha$ fluxes (8×10^{-19} – $8.5 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$).

Finally, we used `grizli` to extract the mock $H\alpha$ emission lines for each source and each grid combination as for real sources. If the extracted $H\alpha$ line had a $S/N > 5$, the completeness for that grid point was assigned 1; otherwise, the completeness was 0. We combined the resulting completeness grids of all the mock sources in each field and survey by using an average of all ~ 4000 grids. This led to three separate completeness functions, each depending on two variables: the wavelength (i.e., redshift) and $H\alpha$ flux of the emission line. Although the completeness is also position-dependent, we marginalized over this

dependence by placing sources uniformly across the FRESCO field and averaging the completeness array of all sources. These line-detection completeness functions are presented in Fig. 3. These show an asymmetric wavelength dependence due to the filter/grism throughput curve. In terms of $H\alpha$ -flux level, the 50% limit is generally crossed at $\log(f_{H\alpha}/\text{erg s}^{-1} \text{ cm}^{-2}) \sim -17.7$, which corresponds to the FRESCO 5σ sensitivity limit of $\sim 2 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ as expected from the ETC (see Oesch et al. 2023).

3.4. Visual inspection completeness

Due to possible contamination from other sources along the dispersion direction, grism spectra need to be visually inspected. Experienced inspectors decide whether a line is likely to be a real $H\alpha$ emission line. However, some genuine emission lines might be missed during this process, due to a low S/N or the presence of contamination. To account for this visual inspection incompleteness, we selected a subsample of 320 mock $H\alpha$ emitters with $S/N > 5$, separately for CONGRESS and FRESCO. For each survey, we selected mock sources at random pixel positions and redshifts in a grid of $\log(S/N)$. We did not cover the whole S/N range of our observations, as at the high- S/N end we found that emission lines are always properly identified during visual inspection, and thus the bright sample is complete in this sense.

We visually inspected these mock sources and gave them quality flags in the same manner as with the real data. Then, we computed the fraction of mock sources with $q \geq 1.75$ in each flux bin, which correspond to the sources that would be selected into our sample if they were real. We fitted this S/N -dependent completeness with a sigmoid function. The result is shown separately for CONGRESS and FRESCO in Fig. 4, and the sigmoid functions that model this visual inspection completeness in each field are

$$C_{\text{ins;CONGRESS}} = \frac{0.90}{1 + \exp[15.46 \times (\log_{10}(f_{H\alpha}) - 0.89)]} \quad \text{and} \quad (5)$$

$$C_{\text{ins;FRESCO}} = \frac{0.94}{1 + \exp[13.530.87 \times (\log_{10}(f_{H\alpha}) - 0.87)]}. \quad (6)$$

Both of these functions cross the 50% limit at $S/N \sim 7.5$.

3.5. Total completeness

The total completeness is a multiplication of the four completeness functions obtained in this section: photometric detection, extraction, line detection, and visual inspection. We compute the completeness for each source, which depends on the detection magnitude, $H\alpha$ flux, line wavelength (redshift), and $H\alpha$ S/N . The majority of the galaxies in our catalog have a total completeness above 65%, with only a quarter of them below this limit. In Fig. 5, we show the $H\alpha$ fluxes of galaxies as a function of the detection magnitude and we highlight the sources with an estimated completeness lower than 65%. As can be seen, completeness corrections start to become significant for sources below a detection magnitude of 27 mag and for fluxes fainter than $3 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$, corresponding to $\log L_{H\alpha}/\text{erg s}^{-1} \lesssim 41.75$; $M_{\text{UV}} \gtrsim -20$. Bins of luminosity with average completeness less than 65% are not included in the following derivations of LFs.

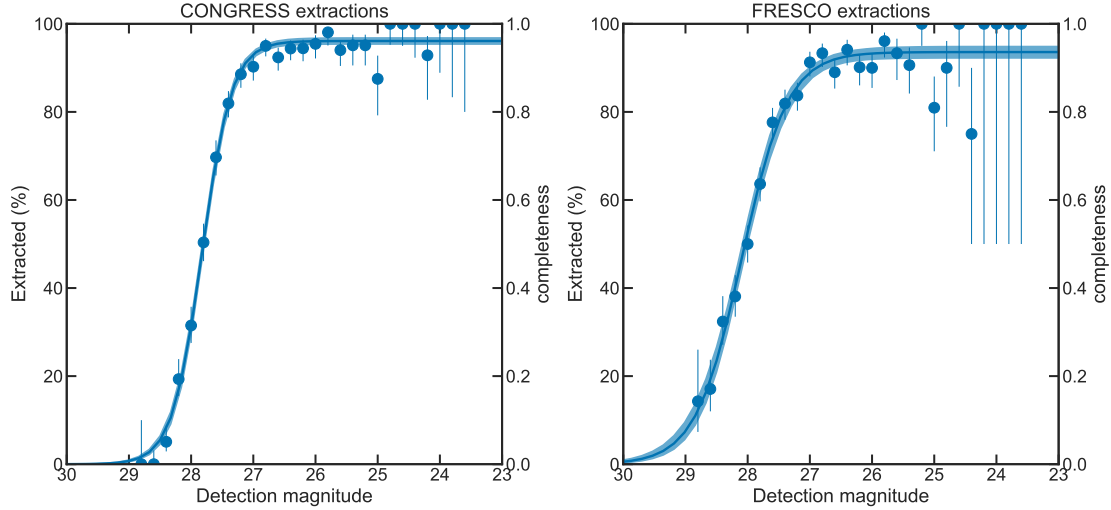


Fig. 2. Percentage of sources that were extracted with *grizli*, in detection magnitude bins. The error bars represent the binomial confidence intervals. The extraction completeness is modeled as a sigmoid function, separately for CONGRESS (left) and FRESCO (right). The resulting sigmoid functions are shown in Eqs. (3) and (4). We reach maximum completeness at mag ~ 27 .

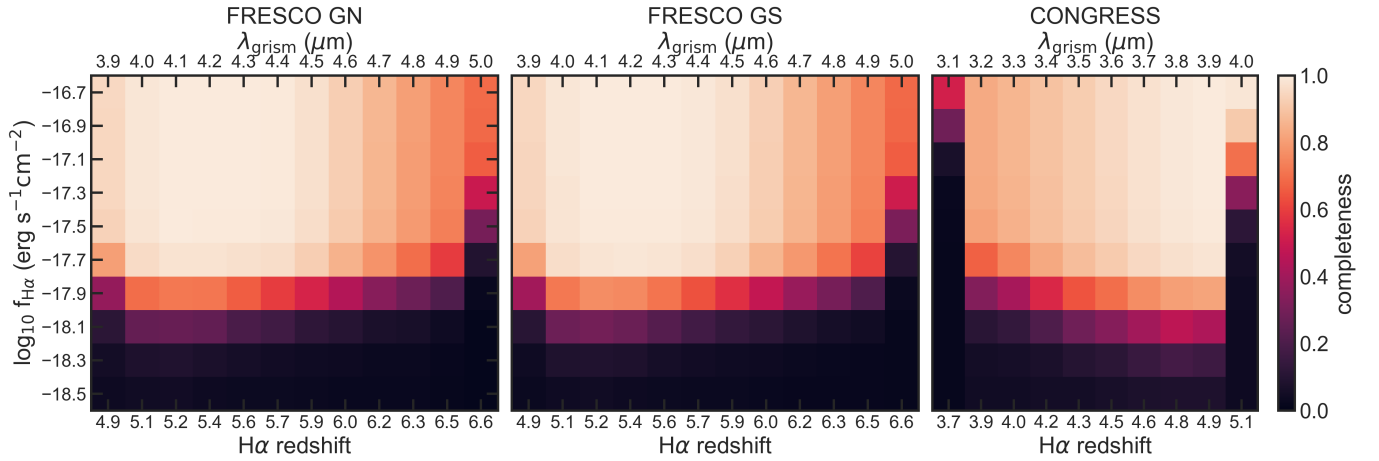


Fig. 3. Completeness of our line detection criteria as a function of grism wavelength (probing $H\alpha$ redshift) and $H\alpha$ flux, for each field and survey. Top: FRESCO GOODS-N. Middle: FRESCO GOODS-S. Bottom: CONGRESS (GOODS-N). These show an asymmetric wavelength dependence due to the sensitivity curves and cross the 50% limit at $\log(f_{H\alpha}/\text{erg s}^{-1} \text{cm}^{-2}) \sim -17.7$.

4. Results

4.1. Redshift distribution and overdensities

We show the redshift distribution of our $H\alpha$ emitters in Fig. 6, separately for each survey and field. From CONGRESS data, we find a significant overdensity in GOODS-N at $z \sim 4.4$, with 98 out of all 554 $H\alpha$ emitters from the CONGRESS field lying at $z = 4.43 \pm 0.03$. In the FRESCO fields, we find the same overdensities that were reported in Helton et al. (2024): at $z \sim 5.4$ and $z \sim 5.9$ in GOODS-S and a very large overdensity at $z \sim 5.2$ in GOODS-N (see also Herard-Demanche et al. 2025).

As mentioned in Sect. 2.3, among the $H\alpha$ emitters in our catalog, we found 13 sources consistent with AGN activity: five in CONGRESS and eight in FRESCO (reported in Appendix C). This constitutes a $1.09^{+0.38}_{-0.28}\%$ of such sources in our catalog. For CONGRESS, four out of the five AGN candidates belong to the overdensity at $z \sim 4.4$, suggesting a possible connection between AGN and larger scale environment at the 3σ level. We defer a more detailed discussion of this to a future paper.

4.2. Observed $H\alpha$ luminosity function

We now compute the $H\alpha$ luminosity function at $3.8 < z < 6.6$. We split our full $H\alpha$ catalog into three redshift bins: one for all CONGRESS data, with $3.8 < z < 5.1$; and two for FRESCO data, splitting our redshift range in half: $4.9 < z < 5.7$, and $5.7 < z < 6.6$. We remove the broad-line $H\alpha$ emitters from the luminosity function calculation, as their $H\alpha$ flux does not only probe star formation (see Sect. C). This resulted in 527 CONGRESS sources at $z \sim 4.45$, 379 FRESCO sources at $z \sim 5.3$, and 94 FRESCO sources in the last bin centered at $z \sim 6.15$.

For each redshift bin, we derive the observed $H\alpha$ luminosity function following the V/V_{max} method in Schmidt (1968). For each $\log L_{H\alpha}$ bin, we calculate the number density as

$$\phi d(\log L_{H\alpha}) = \sum_i \frac{V}{V_{\text{max},i} C_i}, \quad (7)$$

where the summation i runs over all galaxies in the given luminosity bin, V is the full FRESCO survey volume over the redshift

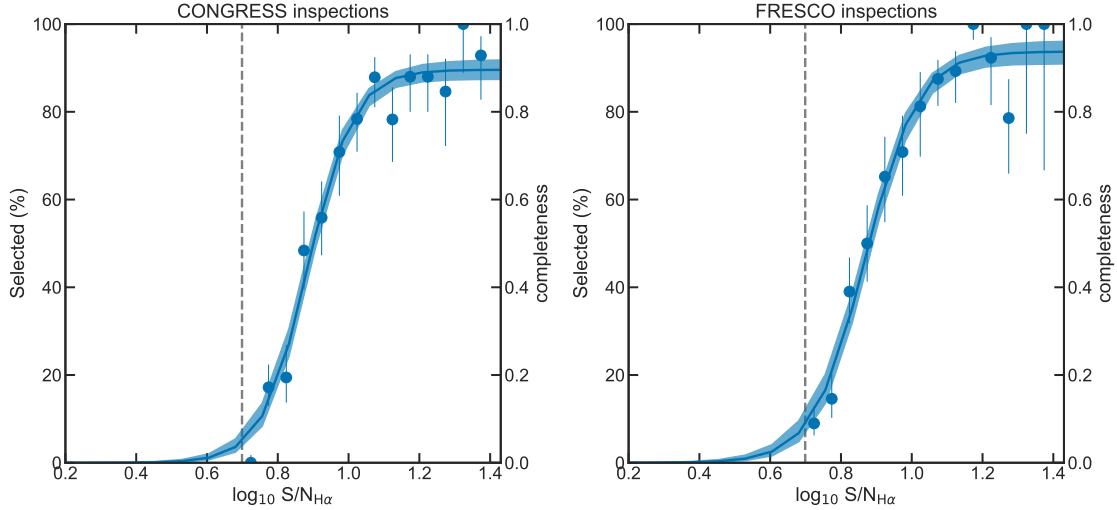


Fig. 4. Percentage of mock sources that were selected through visual inspection as $q \geq 1.75$, in $\log(H\alpha)$ -flux bins. The visual inspection completeness is modeled as a sigmoid function, separately for CONGRESS (left) and FRESCO (right). The vertical dashed line marks the $S/N > 5$ limit. The resulting sigmoid functions are shown in Eqs. (5) and (6).

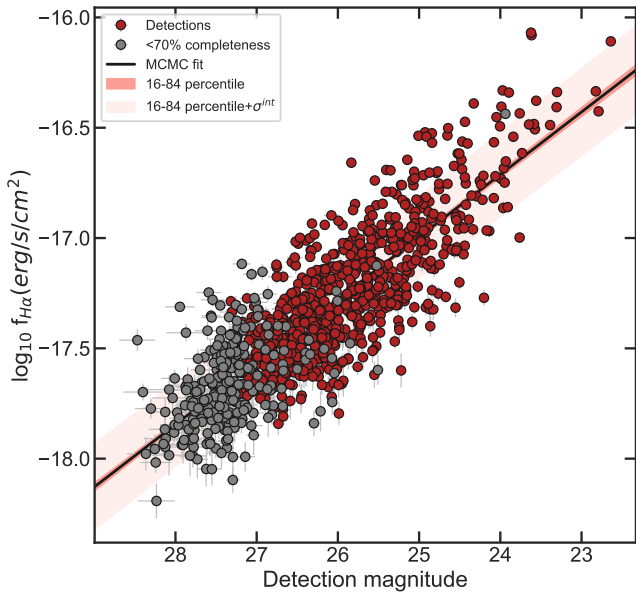


Fig. 5. Relation between our catalog's $H\alpha$ flux and the detection magnitude (stack of F444W and F210M). The solid, black line represents a Bayesian linear fit, while the shadowed area represents the intrinsic 1σ dispersion around the mean relation. Grey, filled circles are the sources with completeness below 65%, while red circles correspond to galaxies with higher completeness. As can be seen, our catalog of $H\alpha$ emitters is starting to become incomplete at $\sim 3 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$, where we are likely missing emitters that are not included in our parent catalog or are not extracted due to our requirements on photo-zs.

bin, $V_{\max,i}$ is the maximum volume in which one could theoretically detect source i in that $\log L_{H\alpha}$ bin at redshift z_i with $S/N > 5$, and C_i is the completeness of the source i . The uncertainty of the number density includes Poisson noise, the completeness uncertainty, as well as cosmic variance following Trenti & Stiavelli (2008), obtained with the cosmic variance calculator⁵. The results for a $\log L_{H\alpha}$ bin width of 0.25 dex are

⁵ Available at <https://www.ph.unimelb.edu.au/~mtrenti/cvc/CosmicVariance.html>

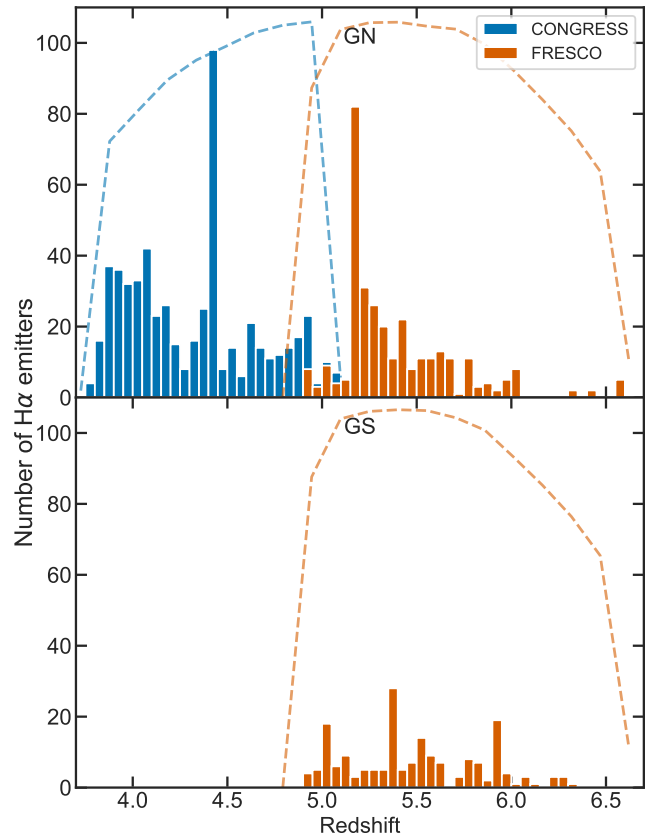


Fig. 6. Redshift distribution of the $H\alpha$ emitters found in CONGRESS (GOODS-N) and FRESCO (GOODS-N and GOODS-S). Dashed lines show the $S/N > 5$ line detection completeness as a function of redshift for each survey and field. We find clear overdensities in GOODS-N at $z \sim 4.4$ and $z \sim 5.2$, and in GOODS-S at $z \sim 5.4$ and $z \sim 5.9$.

shown in Table 1 and Fig. 7. We do not consider the number density values with an average completeness below 65% for the fit, although we do show them in the figure as transparent points. We also include the dust-corrected $H\alpha$ luminosity function bins in Fig. 7, which will be discussed in Sect. 4.5.

Table 1. H α luminosity functions.

$\log L_{\text{H}\alpha}$ erg s $^{-1}$	N	$\langle c \rangle$	ϕ_{comp} 10 $^{-3}$ Mpc $^{-3}$
Congress ($z \sim 4.45$)			
42.00	149	0.70	$3.75^{+0.60}_{-0.59}$
42.25	108	0.73	$2.59^{+0.44}_{-0.43}$
42.50	38	0.75	$0.90^{+0.21}_{-0.19}$
42.75	20	0.73	$0.48^{+0.15}_{-0.13}$
43.00	4	0.73	$0.096^{+0.077}_{-0.048}$
43.25	2	0.75	$0.047^{+0.062}_{-0.031}$
FRESCO ($z \sim 5.3$)			
42.00	135	0.75	$3.04^{+0.48}_{-0.47}$
42.25	87	0.80	$1.82^{+0.32}_{-0.30}$
42.50	47	0.82	$0.97^{+0.20}_{-0.19}$
42.75	18	0.82	$0.366^{+0.118}_{-0.098}$
43.00	7	0.83	$0.142^{+0.079}_{-0.056}$
FRESCO ($z \sim 6.15$)			
42.00	36	0.68	$1.18^{+0.23}_{-0.34}$
42.25	23	0.69	$0.72^{+0.17}_{-0.22}$
42.50	13	0.70	$0.38^{+0.15}_{-0.12}$
42.75	4	0.75	$0.120^{+0.075}_{-0.067}$

We fit a Schechter function to the observed H α luminosity functions:

$$\phi = \frac{dn}{d \log L} = \ln(10) \phi^* 10^{(\log L - \log L^*)(\alpha+1)} \exp \left[10^{(\log L - \log L^*)} \right], \quad (8)$$

parametrized by ϕ^* as the characteristic number density, L^* as the characteristic luminosity, and α as the faint-end slope. To estimate these parameters, we employed a Bayesian inference approach using Markov Chain Monte Carlo (MCMC) sampling with the pymc package. We set weakly informative priors for the parameters chosen as normal distributions centered on the values found by Bollo et al. (2023) at $z \sim 4.4$. We started the sampling process by finding the parameter values that maximize the posterior distribution. Subsequently, we performed MCMC sampling with 5000 iterations across four chains to obtain posterior distributions for the parameters. The posteriors were used to derive the median values and 16th–84th percentiles for each parameter, providing a comprehensive statistical characterization of the fitted Schechter function. The results of these fits are listed in Table 2 and also shown in Fig. 7.

4.3. Simulated H α luminosity function

In order to compare the observed H α luminosity functions with theoretical predictions, we added several results from simulations to Fig. 7. In particular, we use the results of the First Light And Reionization Epoch Simulations (FLARES) (Lovell et al. 2021; Vijayan et al. 2021) and SPHINX²⁰ (Rosdahl et al. 2022),

specifically the SPHINX Public Data Release v1 (SPDRv1) catalog of mock observations described by Katz et al. (2023).

While there are significant differences between the predicted LFs from these three simulations, we find decent agreement overall with our observational estimates. In particular, we find very good agreement with the predictions from FLARES, especially at the bright end. Nevertheless, in the highest redshift bin, the FLARES LF shifts towards the higher end of our fit. Interestingly, all three simulations consistently predict a higher number density of sources below our detection limit (at $\log L_{\text{H}\alpha}/\text{erg s}^{-1} \sim 41.5$) than the extrapolation of our derived LFs. This might indicate that our faint end slopes are somewhat too shallow and deeper observations are needed to test this. However, we note that the overall shape of our observed LFs, including the faint-end slopes, are in very good agreement with the predictions from the SC SAM model, except that they are offset by up to ~ 0.3 – 0.5 dex in normalization. On the other hand, we find that SPHINX underpredicts the LF, where it overlaps with our sample, which only occurs over a very small luminosity range due to the smaller volume of SPHINX. The different results between the simulations probably stem from different approaches to the modeling of line emission, mostly separating SPHINX from FLARES and SC SAM. SPHINX can model the interstellar medium and self-consistently compute nebular emission; however, this is done at the expense of box size, which means that bright galaxies such as those observed by FRESCO and CONGRESS are rare, hence the large uncertainties in the bright end of the LF. In contrast, FLARES and SC SAM model emission lines through the use of Cloudy (Ferland et al. 1998; Chatzikos et al. 2023). Similar work in comparing high redshift line emission LFs with simulations was done by Meyer et al. (2024), who also found good agreement with FLARES and SC SAM, and that SPHINX underpredicts their values. These comparisons show that the summary statistics of emission line measurements like the LFs are a powerful new avenue to test early galaxy formation models.

4.4. UV luminosity function

Our new catalog of H α emitters in the GOODS fields allowed us to revisit the UV LF measurements that have been performed over the last decade with photometrically selected Lyman-break galaxies, but now for the first time with spectroscopically confirmed sources. Using the same methods as in Sect. 4.2, we divided our sources into the same three redshift bins and used a M_{UV} bin size of 0.625 mag, akin to the $\log L_{\text{H}\alpha}$ bin separation. The results are shown in Fig. 8, plotted together with the UV LF Schechter fits from Bouwens et al. (2021). We find that, for M_{UV} bins where the average completeness is above 65%, our results based on spectroscopic redshifts are an excellent match, within uncertainty, to those obtained by Bouwens et al. (2021) using photometrically selected objects. This provides further confidence in the accuracy of previous UV LF measurements at these epochs.

4.5. Star formation rate functions

The UV and H α LFs are based on observed quantities. However, what we really would like to constrain is the buildup of star formation in early galaxies at $z > 3$. We can now do this based on the observed H α emission lines, after a correction for dust extinction. To convert the H α luminosities to SFRs, we used the relation in Kennicutt (1998), scaled by a factor of 1.8 to consider a Chabrier (2003) initial mass function (IMF):

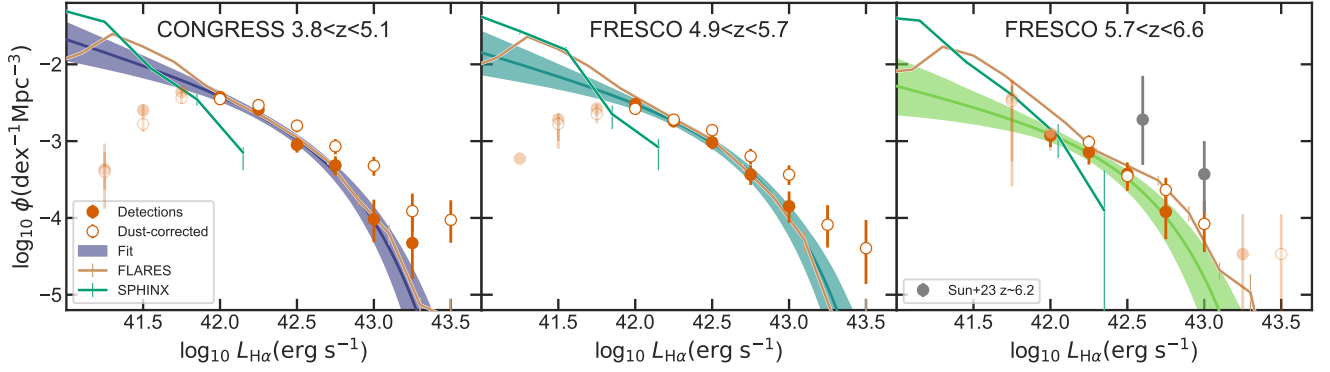


Fig. 7. Observed $H\alpha$ luminosity function at $z = 3.8\text{--}6.6$, split in three redshift bins. Solid circles show the observed number densities, which are transparent when their average completeness is below 65% (these values are not considered for the fit). Solid lines within shaded regions show the fitted Schechter functions and the 68% confidence intervals. Also shown are the results of FLARES (Lovell et al. 2021; Vijayan et al. 2021) and SPHINX (Katz et al. 2023) at $z \sim 4$, $z \sim 5$, and $z \sim 6$; and the results of Sun et al. (2023) at $z \sim 6.2$.

Table 2. Schechter parameters of the observed $H\alpha$ luminosity functions.

z	$\log L_{H\alpha}^*$ erg s^{-1}	$\phi_{H\alpha}^*$ 10^{-3} Mpc^{-3}	$\alpha_{H\alpha}$
4.45	$42.60^{+0.17}_{-0.16}$	$0.89^{+0.69}_{-0.43}$	$-1.64^{+0.27}_{-0.21}$
5.30	$42.65^{+0.18}_{-0.17}$	$0.69^{+0.55}_{-0.35}$	$-1.58^{+0.28}_{-0.25}$
6.15	$42.46^{+0.24}_{-0.20}$	$0.45^{+0.40}_{-0.25}$	$-1.49^{+0.36}_{-0.33}$

$$\text{SFR}_{H\alpha} (\text{M}_{\odot} \text{yr}^{-1}) = 10^{-41.36} L_{H\alpha, \text{int}} (\text{erg s}^{-1}). \quad (9)$$

We computed the dust-corrected $H\alpha$ luminosity using intrinsic $H\alpha$ luminosities, obtained with

$$L_{H\alpha, \text{int}} = L_{H\alpha, \text{obs}} \times 10^{0.4 A_{H\alpha}}, \quad (10)$$

where $A_{H\alpha}$ is the dust attenuation at 6564.6 Å derived in Sect. 2.7. We then computed the SFR functions in the same way as the $H\alpha$ LFs (see Eq. (7)). This led to the $H\alpha$ -derived star formation rate functions reported in Table 3 and shown in Fig. 9. We then fitted Schechter functions to these values, using the same MCMC techniques as with the $H\alpha$ luminosity function. The results of these fits are shown in Table 4.

At all redshifts, we find characteristic SFRs of $\sim 100 \text{ M}_{\odot} \text{yr}^{-1}$. This is significantly higher than the characteristic SFRs previously derived by Smit et al. (2016) for the SFR functions at these redshifts based on $H\alpha$ lines inferred from Spitzer/IRAC broadband photometry. We also show these estimates in Fig. 9. As can be seen, we find a significantly higher number density of galaxies with high SFRs ($\log \text{SFR}/\text{M}_{\odot} \text{yr}^{-1} > 2$) compared to the previous estimates from Smit et al. (2016). This difference is likely not the result of wrongly estimated emission line fluxes from the photometric data; it is more likely a consequence of different techniques: while we derive the SFR function directly from individual dust-corrected $H\alpha$ luminosities, Smit et al. (2016) used their $H\alpha$ -based SFRs to derive an average dust correction to convert the UV LFs to SFR functions. We suspect that this procedure likely underestimates the number of high-SFR sources that are present due to the significant amount of scatter in the $L_{UV} - L_{H\alpha}$

relation. Additionally, the UV-based estimates probe star formation on longer timescales than $H\alpha$, which leads to smaller dispersion. This is supported by the fact that our estimates are in very good agreement with the ones from Bollo et al. (2023) who also derived SFR functions at $z \sim 4.5$ based on $H\alpha$ lines inferred from Spitzer/IRAC broadband photometry, but who used a similar procedure as we adopt here to derive the SFR functions for individual sources.

We can further compare the derived SFR functions with simulations. In particular, we have computed the SFR functions directly from the SPHINX and FLARES data. Once again, we find relatively good agreement with FLARES over most of the SFR range. Albeit, the FLARES SFR functions show a dip at $\text{SFR} = 10 \text{ M}_{\odot} \text{yr}^{-1}$ at all redshifts, where their number densities are lower than what we observe by up to 0.5 dex. For the SPHINX simulation, the situation is different. While it predicts fainter $H\alpha$ lines than we observe, the SFR functions are higher than what we find by more than 0.5 dex. This overestimation of the SFR function is similar to the one seen in the stellar mass function (see Weibel et al. 2024). Taken together, this likely indicates that a large amount of dust is present in the simulated galaxies, which obscures the observed quantities such as the $H\alpha$ (or the UV fluxes; see also Katz et al. 2023), leading to an under-predicted $H\alpha$ LF even though the intrinsic number density of galaxies at a given SFR is higher than observed.

4.6. Star formation rate densities

Equipped with the SFR functions derived in the previous section, we can now compute the SFRDs from our $H\alpha$ measurements. We do this by integrating the SFR function, separately for each of the three redshift bins. For consistency with previous work, we set the lower integration limit to $\text{SFR}_l = 0.27 \text{ M}_{\odot} \text{yr}^{-1}$, corresponding to a UV luminosity limit of $M_{UV, l} = -17$. This limit was also used in Bouwens et al. (2015), that we use to compare the UV and $H\alpha$ -based SFRD estimates. As the integrands are Schechter functions, we can solve these integrals analytically using:

$$\begin{aligned} \rho_{\text{SFR}, H\alpha} &= \int_{\log \text{SFR}_l}^{\infty} \phi(\log \text{SFR}) \text{SFR} d(\log \text{SFR}) \\ &= \phi_{\text{SFR}}^* \text{SFR}^* \Gamma(\alpha + 2, \text{SFR}_l / \text{SFR}^*), \end{aligned} \quad (11)$$

where Γ is the incomplete gamma function:

$$\Gamma(a, x) = \int_x^{\infty} t^{a-1} e^{-t} dt. \quad (12)$$

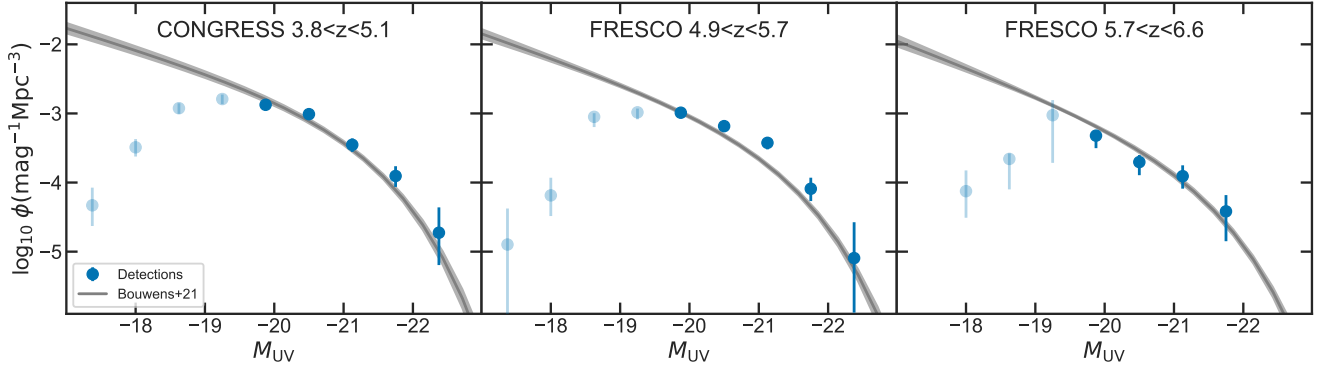


Fig. 8. GOODS fields UV luminosity function at $z = 3.8\text{--}6.6$, split into three different redshift bins. Solid circles show the observed number densities, which are transparent when their average completeness is below 65%. Grey lines and shaded regions are the UV luminosity functions with 68% confidence intervals obtained by Bouwens et al. (2021) for redshifts 4.45, 5.3, and 6.15, respectively.

Table 3. Star formation rate functions from $H\alpha$ emitters.

$\log \text{SFR}$ $\text{M}_{\odot} \text{yr}^{-1}$	N	ϕ_{SFR} 10^{-3}Mpc^{-3}
Congress ($z \sim 4.45$)		
0.64	137	$3.53^{+0.58}_{-0.56}$
0.89	121	$2.94^{+0.49}_{-0.48}$
1.14	66	$1.59^{+0.31}_{-0.29}$
1.39	36	$0.86^{+0.20}_{-0.18}$
1.64	20	$0.48^{+0.15}_{-0.12}$
1.89	5	$0.123^{+0.085}_{-0.056}$
2.14	4	$0.094^{+0.076}_{-0.047}$
FRESCO ($z \sim 5.3$)		
0.64	112	$2.64^{+0.43}_{-0.42}$
0.89	88	$1.91^{+0.33}_{-0.32}$
1.14	67	$1.39^{+0.26}_{-0.25}$
1.39	31	$0.63^{+0.16}_{-0.14}$
1.64	18	$0.366^{+0.118}_{-0.098}$
1.89	4	$0.082^{+0.065}_{-0.040}$
2.14	2	$0.040^{+0.053}_{-0.027}$
FRESCO ($z \sim 6.15$)		
0.89	32	$0.97^{+0.22}_{-0.27}$
1.14	11	$0.35^{+0.14}_{-0.12}$
1.39	8	$0.231^{+0.102}_{-0.097}$
1.64	3	$0.084^{+0.083}_{-0.048}$

For each redshift bin, we integrated each of the 5000 posterior Schechter function parameters obtained during the MCMC sampling and used the median of these results as the SFRD value, and the 16th–84th percentiles as the uncertainties.

This yields the SFRDs reported in Table 4 for each redshift bin: at $z \sim 4.45$ we find $0.042 \text{ M}_{\odot} \text{yr}^{-1} \text{Mpc}^{-3}$ and this declines by a factor of 3 towards $z \sim 6.15$. As will be discussed in the

following section, this is in good agreement with dust-corrected, UV-based measurements (e.g., Bouwens et al. 2015).

5. Discussion

5.1. $H\alpha$ luminosity function evolution

Figure 10 shows the redshift evolution of the $H\alpha$ LF across cosmic history from $z \sim 0.2$ up to $z \sim 6.1$ based on literature estimates as well as our new results. We observe an evolution of these luminosity functions, where $z \sim 6$ galaxies are ~ 1 dex less numerous than those at cosmic noon. This coincides with the epoch of galaxy buildup; hence it is expected that galaxy numbers increase with decreasing redshift. We also observe that the bright end of the $H\alpha$ luminosity function is low ($\log(L_{H\alpha}/\text{erg s}^{-1}) \sim 42.5$) at local redshifts ($0.2 < z < 0.4$), and it increases with redshift up to a maximum of $\log(L_{H\alpha}/\text{erg s}^{-1}) \sim 44$ at $z \sim 2\text{--}3$, to then decrease back to $\log(L_{H\alpha}/\text{erg s}^{-1}) \sim 42.5$. Given that $H\alpha$ is a tracer of galaxy star formation rate, this behavior agrees with the evolution shown by Madau & Dickinson (2014), where there is a peak of star formation at $z \sim 2\text{--}3$. This is also noticeable in Fig. 11, where we compare our best-fit Schechter parameters to those of previous studies at $z \sim 0.2\text{--}4.4$. The characteristic luminosity $\log L^*$ follows the increasing trend with redshift up to $z \sim 3$, to then remain fairly constant for all $4 < z < 6$. Even though we note that there are degeneracies in Schechter function fits, the trends in Fig. 11 indicate that the key driver for the evolution of the $H\alpha$ LFs at $z < 3$ is the characteristic luminosity, which then turns to an evolution in density at $z > 4$, i.e., the normalization parameter $\log \phi^*$ shows a slight decrease with increasing redshift, in agreement with previous results (Bollo et al. 2023).

At the same time, we do not find any significant evolution of the faint-end slope, α , with redshift, unlike for the UV luminosity functions reported by Bouwens et al. (2021), where this parameter tends to flatten with decreasing redshift, being as steep as $\alpha \sim -2$ at $z \sim 7$. However, the uncertainty in our faint-end slope estimates is still very large. Deeper grism exposures will be required to constrain this better in the future, such as the ones from the ALT survey (Naidu et al. 2024).

5.2. Star formation rate density evolution

Figure 12 shows the SFRD values at various redshifts obtained from integrating the $H\alpha$ luminosity functions down to $\text{SFR}_l = 0.27 \text{ M}_{\odot} \text{yr}^{-1}$. For comparison, we also included values obtained

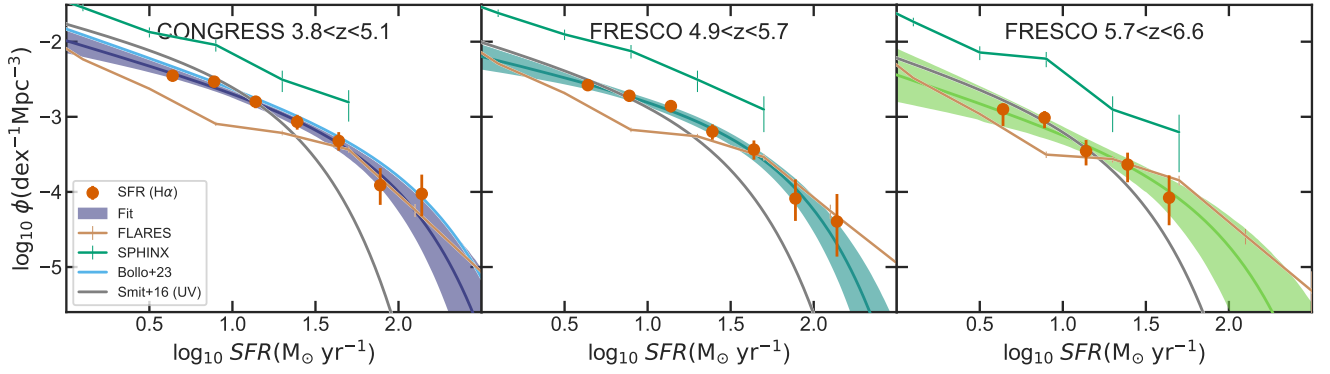


Fig. 9. Star formation rate functions at $z = 3.8$ – 6.6 split into three redshift bins. Solid circles show the number densities derived from dust-corrected $H\alpha$ luminosities. Solid lines within shaded regions show the fitted Schechter functions and the 68% confidence intervals for $H\alpha$ values. Also shown are the simulations results of FLARES (Lovell et al. 2021; Vijayan et al. 2021) and SPHINX (Katz et al. 2023) at $z \sim 4$, $z \sim 5$, and $z \sim 6$; and the literature $H\alpha$ -derived results from Bollo et al. (2023) at redshifts $3.86 < z < 4.94$ and UV-derived from Smit et al. (2016) at $z \sim 3.8$, 4.9 , and 5.9 (scaled to match our $H\alpha$ conversion factor).

Table 4. Schechter parameters of the star formation rate functions.

z	$\log \text{SFR}^*$ $\text{M}_\odot \text{yr}^{-1}$	ϕ_{SFR}^* 10^{-3}Mpc^{-3}	α_{SFR}	$\rho_{\text{SFR:H}\alpha}$ $\text{M}_\odot \text{yr}^{-1} \text{Mpc}^{-3}$
4.45	$2.03^{+0.32}_{-0.21}$	$0.33^{+0.34}_{-0.22}$	$-1.64^{+0.20}_{-0.19}$	$0.041^{+0.005}_{-0.006}$
5.30	$1.89^{+0.19}_{-0.17}$	$0.40^{+0.30}_{-0.19}$	$-1.52^{+0.20}_{-0.18}$	$0.029^{+0.003}_{-0.004}$
6.15	$1.98^{+0.38}_{-0.34}$	$0.09^{+0.16}_{-0.06}$	$-1.73^{+0.30}_{-0.26}$	$0.012^{+0.003}_{-0.003}$

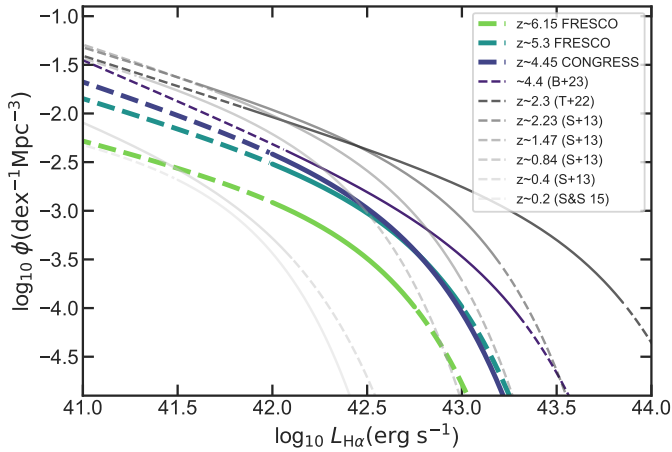


Fig. 10. Evolution of the $H\alpha$ LFs across cosmic history. We compare our best-fit $H\alpha$ LFs at $z \sim 4.45$, $z \sim 5.3$, and $z \sim 6.15$, to previous results: Stroe & Sobral (2015) at $z \sim 0.2$; Sobral et al. (2013) at $z \sim 0.4$, 0.84 , 1.47 , and 2.23 ; Terao et al. (2022) at $z \sim 2.3$; and Bollo et al. (2023) at $z \sim 4.4$. Solid lines represent the regions that were probed observationally in each study. For those studies that only reported the dust-corrected luminosity functions, we inferred the uncorrected functions by reversing the dust correction applied for a fair comparison. We observe a trend of number densities increasing with decreasing redshift up to cosmic noon, to then reverse this trend and decrease until the present universe.

from UV measurements by Bouwens et al. (2015). This provides us with a view of the star formation history of the universe at redshifts $0 \leq z \leq 8$. We find that both $H\alpha$ and UV measurements tend to lie ~ 0.2 dex above the parametrization of Madau & Dickinson (2014), which was derived from UV and

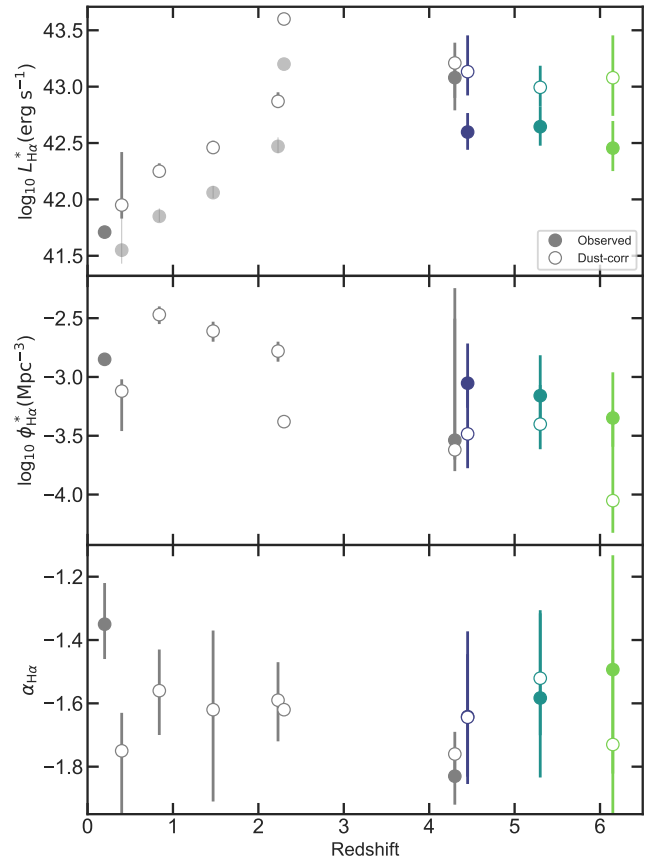


Fig. 11. Cosmic evolution of the $H\alpha$ LF Schechter parameters. Our best-fit values are compared to previous results at lower redshifts: Stroe & Sobral (2015) at $z \sim 0.2$; Sobral et al. (2013) at $z \sim 0.4$, 0.84 , 1.47 , and 2.23 ; Terao et al. (2022) at $z \sim 2.3$; and Bollo et al. (2023) at $z \sim 4.4$. Filled circles represent the parameters of observed $H\alpha$ LFs, while open circles are for dust-corrected fits. For those studies that only reported the Schechter parameters of dust-corrected measurements, we inferred the $\log L^*$ parameter by reversing the dust correction applied, and show those values as fainter solid circles.

IR measurements prior to Bouwens et al. (2015) and integrating down to $\text{SFR}_l = 0.03 L^*$. This disagreement arises from the difference between integration limits.

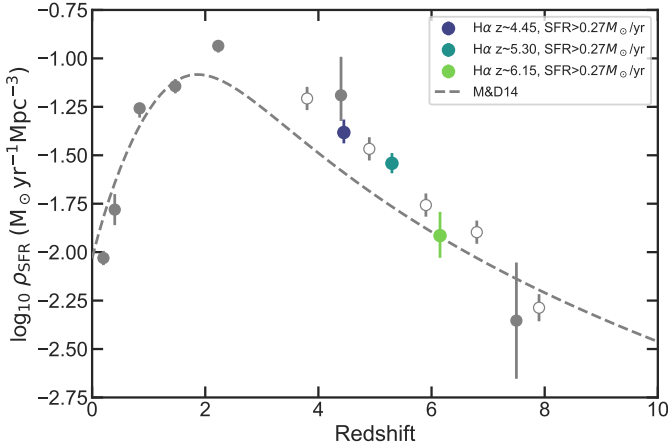


Fig. 12. Cosmic evolution of the SFRD. Our values compared to previous results: Stroe & Sobral (2015) at $z \sim 0.2$; Sobral et al. (2013) at $z \sim 0.4, 0.84, 1.47$, and 2.23 ; Terao et al. (2022) at $z \sim 2.3$; Bollo et al. (2023) at $z \sim 4.4$; and Rinaldi et al. (2023) at $z \sim 7.5$. Filled circles represent values obtained from $H\alpha$ luminosity functions, while open circles are from Bouwens et al. (2015) UV luminosity functions. The dotted line is the fit by Madau & Dickinson (2014). All literature results were scaled to match our $H\alpha$ conversion factor.

It is important to note that both $H\alpha$ -derived and UV-derived SFRD measurements agree remarkably well with each other. Given that $H\alpha$ traces star formation at much shorter timescales than UV (~ 10 Myr vs. ~ 100 Myr), it is common that both tracers provide different SFR values for the same galaxy due to burstiness (e.g., Cole et al. 2023; Calabrò et al. 2024; Clarke et al. 2024). Additionally, they are subject to different dust attenuation. Nevertheless, both SFRD measurements agreeing at the studied redshift range shows that these effects average out upon integration.

6. Summary and conclusions

In this work, we presented a new catalog of $H\alpha$ emitters in the GOODS fields obtained using NIRCcam/grism spectroscopy from *JWST*'s FRESKO and CONGRESS surveys. These 1050 sources entail a completeness-characterized census of line emitters at $3.8 < z < 6.6$ in GOODS-North and at $4.9 < z < 6.6$ in GOODS-South. With this catalog, we find a newly discovered overdensity of 98 $H\alpha$ emitters in GOODS-North at $z \sim 4.4$ and we confirm the overdensities reported by Helton et al. (2024). We also find five new broad $H\alpha$ emitters and confirm those reported as LRDs in Matthee et al. (2024).

We computed the UV LFs and confirmed the photometry-based results from Bouwens et al. (2021) that were based on Lyman break selections. We computed the first $H\alpha$ LFs at $z > 3$ based purely on spectroscopic data. We also derived the $H\alpha$ LFs from different simulations, finding that our observed estimates are consistent with the predicted values. By comparing our observed results to previous $H\alpha$ LFs over a range of redshifts, we find that both the bright and faint ends evolve from the local universe by increasing with increasing redshift up to a maximum at $z \sim 2-3$, to then decrease back. This picture is consistent with the SFRD evolution described by Madau & Dickinson (2014), which we confirmed by obtaining three new values at redshifts 4.45, 5.3, and 6.15 from the integrated dust-corrected $H\alpha$ luminosity functions. Our SFRD values agree with those

derived from UV measurements, despite $H\alpha$ and UV tracing star formation at different timescales.

While fitting Schechter functions to the observed $H\alpha$ luminosity functions, we find a large uncertainty in the faint-end slope, α , which is also degenerate with the characteristic number density, ϕ^* . This suggests that it is necessary to conduct deeper observations using NIRSspec to probe the faint end of the $H\alpha$ luminosity function (as is being done with, e.g., the ALT survey). At the same time, COSMOS-3D (Kakiichi et al. 2024) and parallel programs will increase the observed area and find the most massive galaxies, probing the brighter end of the luminosity functions. Our work demonstrates the power of *JWST* NIRCcam/grism to efficiently obtain complete emission-line-selected galaxy samples in a limited observation time (only ~ 2 h on source for FRESKO). In turn, these galaxy samples provide completely new constraints on galaxy formation models.

Data availability

The full $H\alpha$ catalog can be found at <https://github.com/astroalba/fresco>

Acknowledgements. This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. These observations are associated with program Nos. 1895 and 3577. The authors sincerely thank the CONGRESS team (PIs: Egami & Sun) for developing their observing program with a zero-exclusive-access period. We thank Aswin Vijayan and Harley Katz for their help in analyzing the simulation data from FLARES and SPHINX. This work has received funding from the Swiss State Secretariat for Education, Research, and Innovation (SERI) under contract number MB22.00072, as well as from the Swiss National Science Foundation (SNSF) through project grant 200020_207349. The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant DNRF140. Support for program #1895 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 5-03127. Support for this work for RPN was provided by NASA through the NASA Hubble Fellowship grant HST-HF2-51515.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. MS acknowledges support from the European Research Commission Consolidator Grant 101088789 (SFEER), from the CIDEGENT/2021/059 grant by Generalitat Valenciana, and from project PID2023-149420NB-I00 funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU.

References

- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
- Bollo, V., González, V., Stefanon, M., et al. 2023, *ApJ*, **946**, 117
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012a, *ApJ*, **754**, 83
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012b, *ApJ*, **752**, L5
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, *ApJ*, **803**, 34
- Bouwens, R. J., Oesch, P. A., Stefanon, M., et al. 2021, *AJ*, **162**, 47
- Calabrò, A., Pentericci, L., Santini, P., et al. 2024, *A&A*, **690**, A290
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, **533**, 682
- Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, *MNRAS*, **480**, 4379
- Chabrier, G. 2003, *PASP*, **115**, 763
- Chatzikos, M., Bianchi, S., Camilloni, F., et al. 2023, *Rev. Mexicana Astron. Astrofis.*, **59**, 327
- Clarke, L., Shapley, A. E., Sanders, R. L., et al. 2024, *ApJ*, **977**, 133
- Cole, J. W., Papovich, C., Finkelstein, S. L., et al. 2023, *ApJS*, submitted [arXiv:2312.10152]
- Cucciati, O., Tresse, L., Ilbert, O., et al. 2012, *A&A*, **539**, A31
- Dahlen, T., Mobasher, B., Dickinson, M., et al. 2007, *ApJ*, **654**, 172
- D'Eugenio, F., Cameron, A. J., Scholtz, J., et al. 2024, *ApJS*, submitted [arXiv:2404.06531]
- Egami, E., Sun, F., Alberts, S., et al. 2023, Complete NIRCcam Grism Redshift Survey (CONGRESS), JWST Proposal. Cycle 2, ID. #3577

- Eisenstein, D. J., Johnson, B. D., Robertson, B., et al. 2023a, *ApJS*, submitted [arXiv:2310.12340]
- Eisenstein, D. J., Willott, C., Alberts, S., et al. 2023b, *ApJS*, submitted [arXiv:2306.02465]
- Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006, *ApJ*, 647, 128
- Ferland, G. J., Korista, K. T., Verner, D. A., et al. 1998, *PASP*, 110, 761
- Förster Schreiber, N. M., Genzel, R., Bouché, N., et al. 2009, *ApJ*, 706, 1364
- Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, *Space Sci. Rev.*, 123, 485
- Geach, J. E., Smail, I., Best, P. N., et al. 2008, *MNRAS*, 388, 1473
- Gialalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJ*, 600, L93
- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, 197, 2011
- Hayes, M., Schaerer, D., & Östlin, G. 2010, *A&A*, 509, L5
- Helton, J. M., Sun, F., Woodrum, C., et al. 2024, *ApJ*, 974, 41
- Herard-Demanche, T., Bouwens, R. J., Oesch, P. A., et al. 2025, *MNRAS*, accepted [arXiv:2309.04525]
- Kakiichi, K., Egami, E., Fan, X., et al. 2024, COSMOS-3D: A Legacy Spectroscopic/Imaging Survey of the Early Universe, JWSTProposal. Cycle 3, ID. #5893
- Kashino, D., Silverman, J. D., Rodighiero, G., et al. 2013, *ApJ*, 777, L8
- Katz, H., Rosdahl, J., Kimm, T., et al. 2023, *Open J. Astrophys.*, 6, 44
- Kennicutt, R. C., Jr. 1998, *ApJ*, 498, 541
- 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
- Leethochawalit, N., Trenti, M., Morishita, T., Roberts-Borsani, G., & Treu, T. 2022, *MNRAS*, 509, 5836
- Lovell, C. C., Vijayan, A. P., Thomas, P. A., et al. 2021, *MNRAS*, 500, 2127
- Madau, P., & Dickinson, M. 2014, *ARA&A*, 52, 415
- Matharu, J., Nelson, E. J., Brammer, G., et al. 2024, *A&A*, 690, A64
- Matthee, J., Mackenzie, R., Simcoe, R. A., et al. 2023, *ApJ*, 950, 67
- Matthee, J., Naidu, R. P., Brammer, G., et al. 2024, *ApJ*, 963, 129
- Meyer, R. A., Oesch, P. A., Giovinazzo, E., et al. 2024, *MNRAS*, 535, 1067
- Naidu, R. P., Matthee, J., Kramarenko, I., et al. 2024, *Open J. Astrophys.*, submitted [arXiv:2410.01874]
- Nelson, E., Brammer, G., Giménez-Arteaga, C., et al. 2024, *ApJ*, 976, L27
- Oesch, P. A., Brammer, G., Naidu, R. P., et al. 2023, *MNRAS*, 525, 2864
- Oke, J. B., & Gunn, J. E. 1983, *ApJ*, 266, 713
- Pirie, C. A., Best, P. N., Duncan, K. J., et al. 2024, *MNRAS*, submitted [arXiv:2410.11808]
- Reddy, N. A., & Steidel, C. C. 2009, *ApJ*, 692, 778
- Rigby, J., Perrin, M., McElwain, M., et al. 2023, *PASP*, 135, 048001
- Rinaldi, P., Caputi, K. I., Costantin, L., et al. 2023, *ApJ*, 952, 143
- Robotham, A. S. G., & Driver, S. P. 2011, *MNRAS*, 413, 2570
- Rosdahl, J., Blaizot, J., Katz, H., et al. 2022, *MNRAS*, 515, 2386
- Schenker, M. A., Robertson, B. E., Ellis, R. S., et al. 2013, *ApJ*, 768, 196
- Schiminovich, D., Ilbert, O., Arnouts, S., et al. 2005, *ApJ*, 619, L47
- Schmidt, M. 1968, *ApJ*, 151, 393
- Smit, R., Bouwens, R. J., Labbé, I., et al. 2016, *ApJ*, 833, 254
- Sobral, D., Smail, I., Best, P. N., et al. 2013, *MNRAS*, 428, 1128
- Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, *ApJ*, 795, 165
- Stroe, A., & Sobral, D. 2015, *MNRAS*, 453, 242
- Sun, F., Egami, E., Pirzkal, N., et al. 2023, *ApJ*, 953, 53
- Terao, Y., Spitler, L. R., Motohara, K., & Chen, N. 2022, *ApJ*, 941, 70
- Trenti, M., & Stiavelli, M. 2008, *ApJ*, 676, 767
- Vijayan, A. P., Lovell, C. C., Wilkins, S. M., et al. 2021, *MNRAS*, 501, 3289
- Weibel, A., Oesch, P. A., Barrufet, L., et al. 2024, *MNRAS*, 533, 1808
- Williams, C. C., Tacchella, S., Maseda, M. V., et al. 2023, *ApJS*, 268, 64
- Wyder, T. K., Treyer, M. A., Milliard, B., et al. 2005, *ApJ*, 619, L15

Appendix A: Repeated sources

Within our catalog of 1050 $H\alpha$ emitters, we find seven sources to show $H\alpha$ emission in both FRESCO and CONGRESS surveys, two of which are multi-component. This overlap is only technically possible at $4.86 < z < 5.15$, which in both surveys correspond to the edge of the grism, and thus the sensitivity at this redshift range is much lower than at the center of the grism. In fact, our seven overlapping sources are in the subrange $4.91 < z < 5.03$, where there are only 17 confirmed sources in FRESCO and 21 in CONGRESS. Moreover, the overlapping sources are the brighter emitters in this redshift range, and only the fainter ones are not retrieved in the other survey, respectively.

The seven sources that we retrieve from both surveys show the same grism redshift, within a scatter of $0.05 \times (1 + z)$, and the same flux without an offset, within 0.06σ .

Appendix B: Flux calibration

In Fig. B.1 we compare our $H\alpha$ measurements to those obtained by JADES DR3 (D'Eugenio et al. 2024) for the 123 sources present in both catalogs. Our measurements were obtained with NIRCam/grism spectroscopy, while JADES DR3 used NIRSpec G140M/G235M/G395M grating and prism spectroscopy. We find a median offset on the $H\alpha$ fluxes of $0.13 - 0.15$ dex (calculated as $\log(f_{\text{grism}}/f_{\text{JADES}})$), which likely arises from an undercorrection on the NIRSpec fluxes. The scatter between the two measurements is likely due to calibration differences between NIRSpec and NIRCam/grism. We find that both calibrations are consistent within $0.26(0.20)\sigma$.

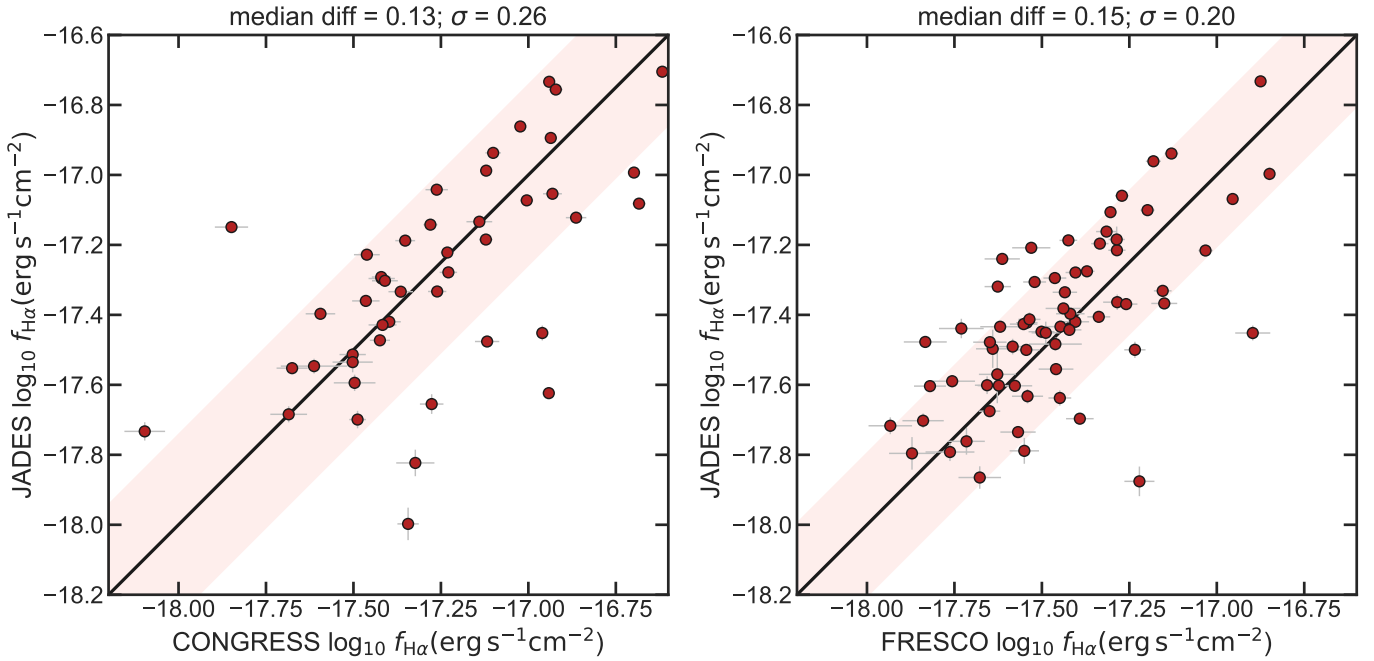


Fig. B.1. Comparison between our $H\alpha$ fluxes, obtained with NIRCam/grism, and the $H\alpha$ fluxes in JADES DR3 (D'Eugenio et al. 2024), obtained with NIRSpec prism. The solid lines are the 1:1 relations, while the shaded regions are the 1σ deviations. The median offsets written on top are calculated as $\log(f_{\text{grism}}/f_{\text{JADES}})$.

Appendix C: Broad $H\alpha$ emitters

Upon visual inspection, we found 13 of our sources to show broad $H\alpha$ lines likely consistent with AGN activity: five in CONGRESS and eight in FRESCO. The eight FRESCO sources are those reported by Matthee et al. (2024), where they have been classified as Little Red Dots, while the five sources in CONGRESS have not been reported before, and thus we performed an analysis to determine the FWHM of their broad emission. Following the procedure in Matthee et al. (2024), we fit their emission-line profiles with a combination of narrow and broad components of $H\alpha$ and $[\text{N II}]$. We find the values reported in Table C.1. Regarding the morphology, four out of the five broad $H\alpha$ emitters in CONGRESS are point-like sources, as shown on the stamps in Fig. C.1.

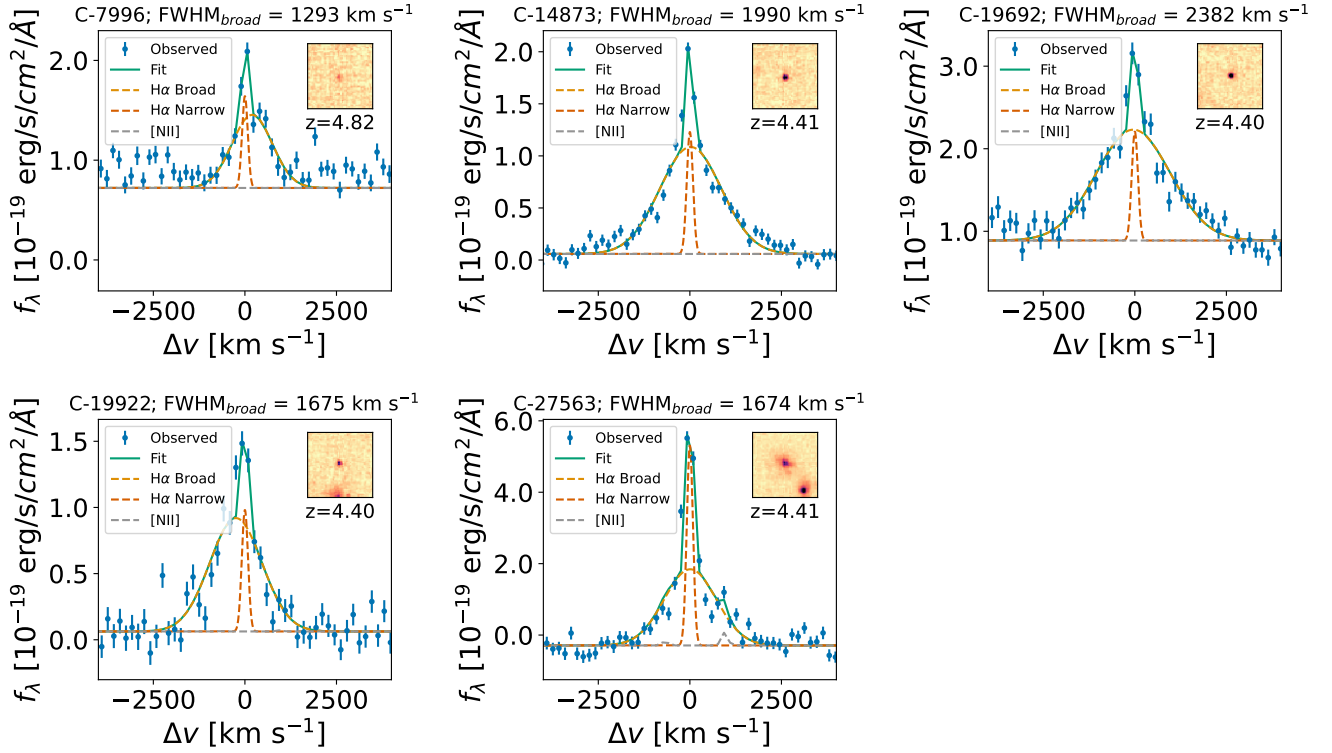


Fig. C.1. 1D spectra and F444W stamps of the five broad H α emitters in CONGRESS identified in this work. Filled circles represent the observed flux values of the spectra, solid lines are the multicomponent Gaussian fits (continuum+H α +[N II]), and dashed lines are the H α broad, narrow, and [N II] components of these fits. The spectra are centered on the redshift of the narrow component.

Table C.1. Coordinates, redshifts, and broad-component FWHM of the 13 broad H α emitters found in this work

ID	RA deg	DEC deg	z_{spec}	FWHM km s $^{-1}$
CONGRESS-7996	189.187148	62.272917	4.821	1293 $\pm$ 178
CONGRESS-14873	189.161854	62.251068	4.409	1990 $\pm$ 131
CONGRESS-19692	189.286517	62.238126	4.404	2383 $\pm$ 210
CONGRESS-19922	189.305664	62.236941	4.402	1675 $\pm$ 102
CONGRESS-27563	189.276132	62.214160	4.406	1674 $\pm$ 143
FRESCO-GN-3549	189.179308	62.292535	5.359	1443 $\pm$ 138
FRESCO-GN-5896	189.285526	62.280772	5.087	1503 $\pm$ 185
FRESCO-GN-7827	189.072083	62.273428	5.145	1473 $\pm$ 107
FRESCO-GN-9123	189.057063	62.268938	5.245	2206 $\pm$ 171
FRESCO-GN-10745	189.344287	62.263365	5.242	2949 $\pm$ 180
FRESCO-GN-16262	189.281029	62.247311	5.539	2556 $\pm$ 100
FRESCO-GN-26821	189.300147	62.212065	5.225	1982 $\pm$ 179
FRESCO-MultiGS-09	53.1385956	-27.790274	5.483	1396 $\pm$ 99