The Astrophysical Journal, 935:94 (13pp), 2022 August 20 https://doi.org/10.3847/1538-4357/ac7e44

High Equivalent Width of Hα+[N II] Emission in $z \sim 8$ Lyman-break Galaxies from IRAC 5.8 μm Observations: Evidence for Efficient Lyman-continuum Photon Production in the Epoch of Reionization

Mauro Stefanon[1,2,3], Rychard J. Bouwens[1], Garth D. Illingworth[4], Ivo Labbé[5], Pascal A. Oesch[6,7], and Valentino Gonzalez[8,9]
[1] Leiden Observatory, Leiden University, NL-2300 RA Leiden, The Netherlands; mauro.stefanon@uv.es
[2] Departament d'Astronomia i Astrofísica, Universitat de València, C. Dr. Moliner 50, E-46100 Burjassot, València, Spain
[3] Unidad Asociada CSIC "Grupo de Astrofísica Extragaláctica y Cosmología" (Instituto de Física de Cantabria—Universitat de València), Spain
[4] UCO/Lick Observatory, University of California, Santa Cruz, 1156 High St, Santa Cruz, CA 95064, USA
[5] Centre for Astrophysics and SuperComputing, Swinburne, University of Technology, Hawthorn, Victoria, 3122, Australia
[6] Departement d'Astronomie, Université de Genève, 51 Ch. des Maillettes, CH-1290 Versoix, Switzerland
[7] Cosmic Dawn Center (DAWN), Niels Bohr Institute, University of Copenhagen, Jagtvej 128, Kø benhavn N, DK-2200, Denmark
[8] Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago 7591245, Chile
[9] Centro de Astrofísica y Tecnologias Afines (CATA), Camino del Observatorio 1515, Las Condes, Santiago 7591245, Chile
Received 2022 April 5; revised 2022 June 29; accepted 2022 July 4; published 2022 August 19

Abstract

We measure, for the first time, the median equivalent width (EW) of Hα+[N II] in star-forming galaxies at $z \sim 8$. Our estimate leverages the unique photometric depth of the Spitzer/IRAC 5.8 μm band mosaics (probing ≈5500–7100 Å at $z \sim 8$) of the GOODS Reionization Era Wide Area Treasury from Spitzer (GREATS) program. We median-stacked the stamps of 102 Lyman-break galaxies in the 3.6, 4.5, 5.8, and 8.0 μm bands, after carefully removing potential contamination from neighboring sources. We infer an extreme rest-frame $\mathrm{EW_0(H\alpha+[N\,II])} = 2328^{+1326}_{-1127}$ Å from the measured red $[3.6]-[5.8] = 0.82 \pm 0.27$ mag, consistent with young ($\lesssim 10^7$ yr) average stellar population ages at $z \sim 8$. This implies an ionizing photon production efficiency of $\log(\xi_{\mathrm{ion,0}}/\mathrm{erg\,Hz^{-1}}) = 25.97^{+0.18}_{-0.28}$. Such a high value for photoproduction, similar to the highest values found at $z \lesssim 4$, indicates that only modest escape fractions $f_{\mathrm{esc}} \lesssim 0.3$ (at 2σ) are sufficient for galaxies brighter than $M_{\mathrm{UV}} < -18$ mag to reionize the neutral hydrogen at $z \sim 8$. This requirement is relaxed even more to $f_{\mathrm{esc}} \leqslant 0.1$ when considering galaxies brighter than $M_{\mathrm{UV}} \approx -13$ mag, consistent with recent luminosity functions and as typically assumed in studies addressing reionization. These exceptional results clearly indicate that galaxies can be the dominant source of reionizing photons, and provide us with an exciting glimpse into what we might soon learn about the early universe, and particularly about the reionization epoch, from forthcoming JWST/MIRI and NIRCam programs.

Unified Astronomy Thesaurus concepts: Reionization (1383); Lyman-break galaxies (979); High-redshift galaxies (734); H alpha photometry (691); Early universe (435)

1. Introduction

The characterization of emission lines is a fundamental tool with which to study the physical processes governing the formation and evolution of galaxies. Hα constitutes one of the most reliable estimators of galaxies' star formation rates (SFRs, e.g., Moustakas 2006; Madau & Dickinson 2014) over short timescales ($\lesssim$10 Myr, e.g., Kennicutt 1998; Kennicutt & Evans 2012) because it tightly correlates with the production of ionizing photons by OB stars, it does not depend on the metallicity, and it is less affected by dust attenuation than rest-UV lines. Moreover, because the rest-frame optical light correlates with the stellar mass ($M_\star$, e.g., Stefanon et al. 2017), the equivalent width (EW) of Hα provides a first estimate (modulo a factor $M_\star/L_{\mathrm{optical}}$) of the specific star formation rate (sSFR, e.g., Fumagalli et al. 2012; Mármol-Queraltó et al. 2016; Faisst et al. 2016).

Optical and near-IR spectroscopy have allowed astronomers to probe Hα up to $z \sim 2.5$ (e.g., Fumagalli et al. 2012; Sobral et al. 2016; Reddy et al. 2018; Nanayakkara et al. 2020). Progress at $z \sim 4$–5 has been enabled by analyzing broadband photometric data from Spitzer/IRAC and interpreting the observed blue $[3.6]-[4.5] < 0$ mag colors as the result of Hα emission contributing to the flux density in the 3.6 μm band (e.g., Stark et al. 2013; Smit et al. 2016; Bouwens et al. 2016c; Rasappu et al. 2016; Faisst et al. 2016, 2019a; Caputi et al. 2017; Lam et al. 2019; Harikane et al. 2018a; Maseda et al. 2020).

Constraining Hα at $z \gtrsim 6$ has proven to be quite challenging for a number or reasons. At these redshifts, the flux densities in both the IRAC 3.6 and 4.5 μm bands are enhanced by nebular line emission ([O III]+Hβ and Hα+[N II], respectively, e.g., Stefanon et al. 2021a), making it difficult to ascertain whether the observed colors are due to the combination of nebular line and continuum emission or just to the continuum. This situation is exacerbated by the general lack of spectroscopic redshifts, essential for identifying which specific nebular lines could be contributing to the flux density in each band. Finally, further uncertainties are introduced by the still unconstrained line ratios at these early epochs (see, e.g., Brinchmann et al. 2008; Steidel et al. 2014; Kewley et al. 2015; Faisst et al. 2016; Harikane et al. 2018b; Stefanon et al. 2022 for discussions on

line ratios potentially evolving with cosmic time), and by the significantly shallower (3–6× depth) data currently available at 5–10 μm (e.g., Stefanon et al. 2021b). These challenges have largely prevented us from securing an emission-line-free continuum estimate at rest-frame optical wavelengths.

Fortunately, a favourable window exists again for $7.0 \lesssim z \lesssim 8.7$. In this redshift range, Hα crosses into the IRAC 5.8 μm band, [O III]+Hβ contribute exclusively to the 4.5 μm band flux density, and the 3.6 μm band is free from significant line emission (e.g., Stefanon et al. 2022). We note, however, that potentially significant emission by [O II] in the 3.6 μm band could mimic a Balmer break, systematically affecting the estimate of the rest-frame optical continuum. In these cases, observations at wavelengths bluer than those in which the Balmer/4000 Å break is expected, or at rest-frame optical wavelengths free from line emission (e.g., the 8.0 μm band at $z \sim 8$) would significantly reduce the impact of this degeneracy on the estimate of line intensities. Thus the potential exists to isolate the key lines to individual bands, and particularly to separate Hα from significant contamination by other lines.

Notably, this redshift range covers ≈80% of the epoch of reionization (EoR) (see, e.g., Planck Collaboration et al. 2020; Robertson 2021). Estimating Hα at these epochs, therefore, is particularly valuable for constraining the Lyman-continuum (LyC) ionizing emissivity (e.g., Leitherer & Heckman 1995) and the rate of production of H-ionizing photons ($\xi_{\rm ion}$, e.g., Bouwens et al. 2016a).

However, the use of observations in the 5.8 and 8.0 μm bands is not a panacea. Both bands suffer from lower sensitivities since 5.8 and 8.0 μm band images could only be acquired during the Spitzer cryogenic mission, whereas those in the 3.6 and 4.5 μm bands continued also during the warm mission. As a result, the sensitivities available in the 5.8 and 8.0 μm bands are generally >5–10× shallower than those available in the 3.6 and 4.5 μm bands. For this reason, most of the studies characterizing the physical properties of EoR galaxies have so far focused on estimating the intensity of [O III]+Hβ from the measured [3.6] – [4.5] color (Smit et al. 2014; Castellano et al. 2017; De Barros et al. 2019; Stefanon et al. 2019, 2022; Bowler et al. 2020; Strait et al. 2020, 2021; Endsley et al. 2021b). Indeed, no detections of Hα from the color excess in the 5.8 μm band have been published so far. The few estimates existing in the literature are indirect, and stem from converting the EW([O III]+Hβ) assuming standard Case B recombination coefficients and metal-line ratios either from best-fit spectral energy distribution (SED) analyses or extracted from tabulated values (e.g., Smit et al. 2014; Stefanon et al. 2022; Endsley et al. 2021b). Recently, attempts to detect emission in the 5.8 μm band for individual $z \sim 7$–8 Lyman-break galaxies (LBGs) were performed by Asada & Ohta (2022), leveraging the lensing magnification of the foreground galaxy clusters Abell 2744, Abell 1063, Abell 370, and MACS-J0717 from the Hubble Frontier Fields program (HFF, Lotz et al. 2017), but were not successful. The lack of direct detections has resulted in a chronic absence of direct constraints on the Hα intensity at these pivotal redshifts.

One possible approach to compensate for the current lack of deep 5–10 μm data consists in combining the imaging available for samples of galaxies and extracting their average properties. Our recently released IRAC mosaics from the GOODS Reionization Era Wide Area Treasury from Spitzer program (GREATS, PI: I. Labbé, Stefanon et al. 2021b) include all the relevant IRAC observations acquired in the four bands over the CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) GOODS-N and GOODS-S fields (Giavalisco et al. 2004) across the almost two decades of Spitzer operations. Notably for this study, the GREATS 5.8 μm band mosaic provides ≈4–140× deeper coverage than what exists for the Abell 2744, Abell 1063, Abell 370, and MACS-J0717 HFF fields[10] and ≳1.5× deeper coverage over a ≳4× larger area than the IRAC Dark Field[11] (Krick et al. 2009). The GREATS mosaics therefore constitute the deepest and thus most suitable data set for probing Hα emission in $z \sim 8$ galaxies prior to JWST operations.

In this study, we explore for the first time the intensity of the Hα emission in $z \sim 8$ galaxies by stacking the image stamps in the IRAC bands centered on 102 candidate LBGs at $7.3 < z_{\rm phot} < 8.7$ identified by Bouwens et al. (2015b) in the CANDELS GOODS, Ultra Deep Survey (UDS), and COSMOS fields. This sample was already utilized by Stefanon et al. (2022) to study the rest-frame optical properties of $z \sim 8$ galaxies, and benefits from minimal neighbor contamination (see Section 2 and Stefanon et al. 2022).

The layout of this paper is as follows: in Section 2 we present the data set and the sample adopted in this study; in Section 3 we describe the procedure we followed to estimate the average flux densities; the stacked photometry and the estimate of the rest-frame EW$_0$(Hα) are presented in Section 4; we place our results in the context of the evolution of the EW$_0$(Hα), sSFR, and $\xi_{\rm ion}$ in Section 5. A summary of this study together with our conclusions is presented in Section 6.

Throughout this paper, we adopt $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70$ km s^{-1} Mpc^{-1}, consistent with the most recent estimates from Planck (Planck Collaboration et al. 2020). Magnitudes are given in the AB system (Oke & Gunn 1983), while our $M_\star$ and SFR measurements are expressed in terms of the Salpeter (1955) initial mass function (IMF). Modeling of the Hα emission line is done using only positive equivalent widths. For brevity, we denote the Hubble Space Telescope (HST) F435W, F606W, F775W, F850LP, F105W, F125W, F140W, and F160W as B_{435}, V_{606}, i_{775}, z_{850}, Y_{105}, J_{125}, JH_{140}, and H_{160}.

2. Data and Sample

The sample adopted for this study consists of the 102 candidate $z \sim 8$ Lyman-break galaxies previously discussed in Stefanon et al. (2022). Briefly, this sample is based on the *Y*-dropout LBGs that Bouwens et al. (2015b) identified over the CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) GOODS-N, GOODS-S (Giavalisco et al. 2004), UDS (Lawrence et al. 2007), and COSMOS (Scoville et al. 2007) fields, the Early Release Science (ERS) field (Windhorst et al. 2011), and the Ultra Deep Field (UDF)/Extreme Deep Field (XDF) (Beckwith et al. 2006; Illingworth et al. 2013; Ellis et al. 2013) with the HUDF09-1 and HUDF09-2 parallels (Bouwens et al. 2011).[12] This initial sample included 185 sources. In Table 1 we summarize the main properties of the adopted data sets. The mosaics are characterized by 5σ depths of ≈27.5 mag in the V_{606} and I_{814} bands, ≈26.7–27.5 mag in the Y_{105} (GOODS

[10] https://irsa.ipac.caltech.edu/data/SPITZER/Frontier/overview.html

[11] http://web.ipac.caltech.edu/staff/jason/darkfield/index.html

[12] We excluded CANDELS/Extended Groth Strip because of the lack of deep data in the *Y* band, which makes the selection of candidate $z \sim 8$ LBGs more uncertain.

Table 1
Observational Data Used for Our Measurement of EW(Hα)

Field Name		Area (arcmin^2)	H_{160}[a]	IRAC Data[b]	3.6 μm[c]	4.5 μm[c]	5.8 μm[c]	8.0 μm[c]
5σ (mag)				5σ (mag)		5σ (mag)	5σ (mag)	5σ (mag)
XDF		4.7	29.4	GREATS	∼27.2	∼26.7	∼23.9	23.5–23.8
HUDF09-1		4.7	28.3	GREATS	∼26.3	∼25.8	∼23.7	∼22.2
HUDF09-2		4.7	28.7	GREATS	∼27.0	25.5–26.0	∼22.5	∼22.2
ERS		40.5	27.4	GREATS	26.2–27.0	25.6–26.7	∼23.5	∼23.3
CANDELS	GOODS-N Deep	62.9	27.5	GREATS	27.0–27.3	26.5–26.8	23.5–24.3	23.3–24.0
	GOODS-N Wide	60.9	26.7	GREATS	26.3–27.2	25.8–26.8	23.5–24.1	23.3–23.9
	GOODS-S Deep	64.5	27.5	GREATS	∼27.3	26.6–26.9	23.5–23.8	23.3–23.8
	GOODS-S Wide	34.2	26.8	GREATS	26.5–27.2	26.2–26.7	∼23.5	∼23.3
	COSMOS	151.9	26.8	SEDS+SCANDELS +S-COSMOS	26.4–26.7	26.0–26.3	∼21.2	∼21.0
	UDS	151.2	26.8	SEDS+SCANDELS +SpUDS	25.4–26.3	25.0–25.9	∼21.5	∼21.7
Totals:		580.2						

Note.
[a] 5σ limit from Bouwens et al. (2015b), computed from the median of measured uncertainties of sources.
[b] GREATS: Stefanon et al. (2021b); SEDS: Ashby et al. (2013); SCANDELS: Ashby et al. (2015); S-COSMOS: Sanders et al. (2007); SpUDS: Caputi et al. (2011).
[c] Nominal 5σ limit for point sources from the SENS-PET exposure time calculator, based on the exposure time maps. Due to inhomogeneities in the coverage, a range of values is quoted when the depth varies by more than ∼0.2 mag across the field. Because of the combined effects of broad Spitzer/IRAC PSF and significant exposure times, source blending may reduce the actual depth (see discussion in Labbé et al. 2015 and Stefanon et al. 2021b).

fields) and 26.0 mag in the ground-based Y band (UDS and COSMOS), and ∼26.8–27.8 mag in the J_{125} and H_{160} bands.

A crucial aspect for this study is that these fields have excellent coverage in the Spitzer/IRAC 3.6, 4.5, 5.8, and 8.0 μm bands. In particular, for the GOODS fields we adopted the mosaics and location-dependent point-spread functions (PSFs) from the GREATS program (PI: Labbé, Stefanon et al. 2021b). These mosaics combine all the useful IRAC data acquired across the full scientific life of Spitzer. As a result they are very deep, with 5σ depths of ∼26.0–27.0 mag in the IRAC 3.6 and 4.5 μm bands, and ∼23.0–24.0 mag in the IRAC 5.8 and 8.0 μm bands. While these are deep, we also require accurate PSFs to minimize the contamination from neighbors (see, e.g., Stefanon et al. 2021b). Given the asymmetric nature of the instrumental IRAC PSF, particularly in the 3.6 and 4.5 μm bands, and the variety of programs included in the mosaics, the PSFs can vary significantly across each field. The PSFs for GREATS are reconstructed by combining an empirical template of high signal-to-noise ratio (S/N) rotated according to the position angle and weighted through the coverage depth from each program at the specific location.

Using EAZY (Brammer et al. 2008), we computed photometric redshifts for the LBG candidates from the available HST observations and required the final selection of galaxies to have $7.3 \leqslant z_{\rm phot} \leqslant 8.7$ (where $z_{\rm phot}$ indicates the peak of the redshift likelihood distribution $p(z)$), and $p(z > 6.5) > 0.6$. This selection resulted in 114 sources. At these redshifts, the IRAC colors are sensitive to the intensity of the main rest-frame optical emission lines. A number of studies have shown that the IRAC colors can be successfully used to significantly reduce the photometric redshift uncertainties (e.g., Smit et al. 2014; Roberts-Borsani et al. 2016). However, this could also potentially bias our sample toward sources with strong emission lines. For this reason, we excluded the flux densities in the IRAC bands when running EAZY. Reassuringly, inclusion of IRAC fluxes in estimating the photometric redshifts has no strong impact on the sources we select (103 sources, corresponding to ∼90%, are in common between the two samples).

Because of the broad PSF of IRAC (PSF FWHM ≈ 1.″5–2.″0 from the 3.6 μm to the 8.0 μm band, Stefanon et al. 2021b), the extended light profiles of neighboring objects could systematically affect the measurement of the emission of specific sources. For this reason, in our analysis we subtracted the neighbor emission with MOPHONGO (Labbé et al. 2006, 2010a, 2010b, 2013, 2015), and removed from the sample 12 sources where visual inspection still showed residual contamination (Stefanon et al. 2022).

To allow for a more meaningful comparison of our results with the literature, we also estimated the main stellar population parameters. These were computed by running FAST (Kriek et al. 2009) on the Bruzual & Charlot (2003) SED templates for $Z_{\rm star} = 0.2\ Z_{\odot}$ metallicity with Salpeter (1955) IMF, a constant star formation history (SFH), and a Calzetti et al. (2000) extinction law. The set of templates was postprocessed through CLOUDY version 17.02 (Ferland et al. 2017), assuming a spherical constant-density nebula with $n(H) = 100\ {\rm cm}^{-3}$, $Z_{\rm gas} = 0.2\ Z_{\odot}$ metallicity, an ionization parameter $\log U = -2.5$ (e.g., Stark et al. 2017; De Barros et al. 2019), and a negligible escape fraction. We note here that the present assumption of constant SFH could result in an underestimate of the stellar mass, particularly for young stellar population ages (e.g., Papovich et al. 2001; Leja et al. 2019; Lower et al. 2020; Topping et al. 2022; Whitler et al. 2022). This uncertainty is currently the result of the coarse spectral coverage provided by Spitzer/IRAC broadband imaging in combination with the low sensitivity of the data in the two reddest IRAC bands. Forthcoming JWST observations will dramatically improve our characterization of the SFH by probing the rest-frame optical continuum levels predicted by the different SFH scenarios.

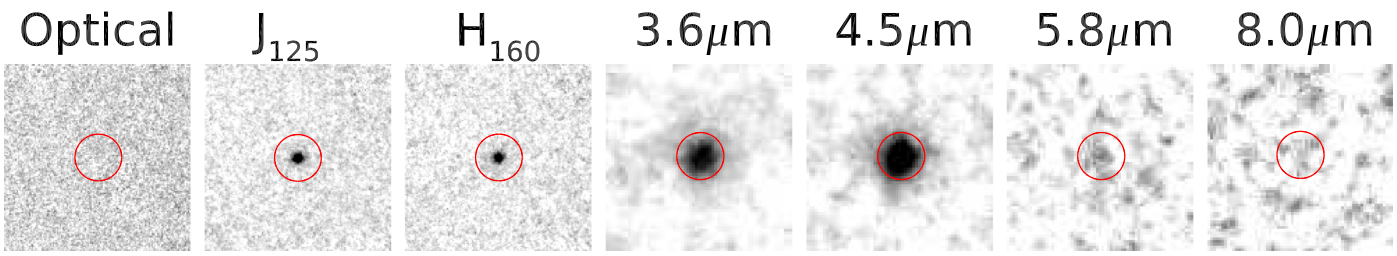

Figure 1. Image stamps (∼8.″0 per side) in the IRAC and HST bands, centered on the median stacks. The red circle marks the 2.″0 diameter aperture adopted for the photometry of the IRAC stacks. The HST stacks are presented to provide a better visual context of the data involved in our study, as the median flux densities in HST bands were estimated from the photometry of individual sources. Each stamp refers to a different band, as labeled at the top; in particular the HST optical stack combines all data available in the B_{435}, V_{606}, i_{775}, and z_{850} bands. Remarkably, we find a 4.3σ detection for the flux density in the 5.8 μm band and a ∼1.8σ measurement in the similarly deep 8.0 μm stack. The striking visual difference between the detection in the 5.8 μm band and those in the 3.6 and 4.5 μm bands is a direct consequence of the ∼40× lower sensitivity available in the 5.8 μm band. Our measurement suggests a significant contribution from Hα emission to the flux density in the 5.8 μm band, as discussed in Section 4.

Table 2
Flux Densities for Our Median-stacked Photometry

	V_{606} (nJy)	i_{775} (nJy)	z_{850} (nJy)	J_{125} (nJy)	H_{160} (nJy)	3.6 μm (nJy)	4.5 μm (nJy)	5.8 μm (nJy)	8.0 μm (nJy)
Stack	-0.4 ± 0.7	-0.5 ± 0.8	-1.3 ± 1.3	46.0 ± 2.4	41.8 ± 1.7	41.5 ± 3.9	68.0 ± 5.4	88.5 ± 20.3	56.3 ± 30.0

Note. We only list the flux densities in those bands available for at least 90% of the sources in our sample.

3. Stacking

Following Stefanon et al. (2022), we adopted distinct stacking procedures for the HST bands and for the IRAC bands, given the different photometric depths. For the HST bands, stacking consisted in evaluating the median of the extracted photometry normalized by the flux density in the H_{160} band of each source, as the generally higher S/N characterizing these data reduces the measurement scatter around the true value. For the IRAC bands, however, the lower S/N than in the HST data could introduce a larger scatter in the final measurement, possibly even systematically affecting it. We therefore constructed image stacks by taking the median of the image stamps centered on each source after they have been cleaned from neighbors using MOPHONGO and normalized by the H_{160} flux density of each source. The stacked IRAC flux densities were measured in 2.″0 diameter apertures. The smaller aperture adopted here compared to what Stefanon et al. (2022) used is a trade-off between optimizing the S/N and minimizing potential flux loss introduced by the challenges in aligning the sources before taking the median and removal of neighbor contamination, particularly in the 5.8 and 8.0 μm data. The aperture photometry was corrected to total using the median of the PSFs reconstructed at the location specific to each source. The applied correction factors are ∼2.2, 2.2, 2.9, and 3.3 for the 3.6, 4.5, 5.8, and 8.0 μm bands, respectively. Uncertainties associated with the flux densities were computed by bootstrapping the sample 1000 times. Finally, all values were rescaled by the median of the flux densities in the H_{160} band. An analysis adopting larger apertures (2.″5 and 3.″6) resulted in measurements consistent at 1σ with those obtained with the smaller aperture, albeit with larger uncertainties.

We further validated our 5.8 μm band measurement through a Monte Carlo simulation, presented in the Appendix. Briefly, we applied the same neighbor removal and stacking procedure we adopted for our main analysis to 102 synthetic sources. We added them to the IRAC 5.8 μm mosaics, after normalizing their flux densities to those expected for the LBGs in our sample, assuming a flat f_ν SED and a rest-frame $\mathrm{EW_0(H\alpha)} = 1900$ Å. This whole process was repeated 100 times. The resulting distribution of flux density measurements shows that, on average, we can recover the input flux density and that the impact of possible contamination by neighbors is negligible, as discussed and shown in more detail in the Appendix and in Figure 6.

The depth of the IRAC 5.8 and 8.0 μm mosaics in the COSMOS and UDS fields is ∼2 mag shallower (corresponding to ∼6× brighter flux limits) than the average depth in the GOODS fields, suggesting we should perhaps exclude or deweight them in our stacks. On the other hand, these two CANDELS-Wide fields do allow us to incorporate 10 of the brightest $z \sim 8$ sources with deep HST imaging from the Bouwens et al. (2015b) catalogs, providing a more comprehensive view of the properties of $z \sim 8$ galaxies. Even so, the median brightness of the sources selected in the COSMOS and UDS fields is only a factor ∼3 brighter than the median for the sources in the GOODS fields. To evaluate the impact of these sources on our stack results, we repeated our stacking analysis excluding the 10 sources in the COSMOS and UDS fields. Reassuringly, the flux densities of the new measurements in the 3.6, 4.5, and 5.8 μm bands differ by ∼5%–10% from those obtained with the full sample, after the change in H_{160} normalization is taken into account. The flux density in the 8.0 μm band is ∼60% fainter but still consistent at the 1σ level with that from the full sample. On balance, we therefore opted for including in our stack the $z \sim 8$ galaxies from the UDS and COSMOS fields.

4. Results

The stacked stamps in the IRAC bands are presented in Figure 1, while the photometry in those bands offering coverage for at least 90% of the sources in our sample is listed in Table 2 and displayed in Figure 2. Our photometric measurements are characterized by $\gtrsim$20σ detections in the HST J_{125} and H_{160} bands, and ∼10σ in the 3.6 and 4.5 μm bands. Remarkably, the stack in the IRAC 5.8 μm band has resulted in a ∼4.3σ detection, while the stack in the 8.0 μm band is

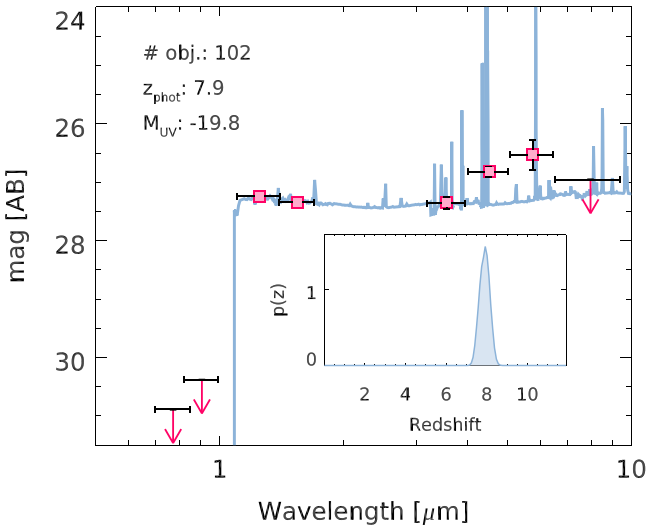

Figure 2. Median-stacked SED resulting from our analysis. The filled red squares with error bars mark the stacked photometry, while the red arrows represent 2σ upper limits. The black horizontal bars indicate the effective width of the bands. The blue curve corresponds to the best-fitting EAZY template. The inset presents the redshift probability distribution computed by EAZY. The labels at the top left corner present the number of objects entering the stack, the median redshift, and the M_{UV} computed by EAZY. Remarkably, the flux density in the 5.8 μm band is comparable to that in the 4.5 μm band, $\sim$2.2$\times$ higher than that in the HST and IRAC 3.6 μm bands, indicative of strong nebular line emission from Hα. Also evident again is the lack of a prominent Balmer break between the H_{160} and 3.6 μm bands, a result previously emphasized in Stefanon et al. (2022), the lead-up study to this analysis.

characterized by a 1.8σ significance. Because the 5.8 μm and 8.0 μm band mosaics adopted in our study have similar depths and image quality, the 8.0 μm band detection, even though of somewhat lower S/N, actually provides valuable added support for our 4.3σ measurement at 5.8 μm as being a genuine detection. Together these detections give added confidence that the detection is real and is not significantly affected by neighbor and/or interloper contamination.

To assist in the interpretation of our measurements and to further validate the consistency of the stacked photometry, we also present the best-fit SED template from EAZY in Figure 2. For this step, we complemented the default set with SEDs of young (age $\sim 10^{6-8}$ yr) star-forming galaxies from BPASS v1.1 (Eldridge et al. 2017), whose nebular emission was computed with CLOUDY (Ferland et al. 2013, 2017). A formal fit with FAST (Kriek et al. 2009) adopting our default configuration (see Section 2) results in a stellar mass of $M_\star = 10^{8.12^{+0.86}_{-0.28}}\,M_\odot$, a dust extinction of $A_V = 0.2^{+0.1}_{-0.2}$ mag, and a stellar population age of $\log(\mathrm{age/yr}) = 7.1^{+1.0}_{-0.5}$. These properties are consistent with the average of the properties Stefanon et al. (2022) found for $z\sim 8$ LBGs from several samples binned by UV luminosity.

The most notable feature evidenced by our stacked photometry is the red $[3.6]-[5.8] = 0.82\pm0.27$ mag color. This is in addition to the lack of a Balmer break between the H_{160} and 3.6 μm band and the robustly red [3.6] – [4.5] colors, indicative of strong [O III]+Hβ emission. Both these results have already been discussed in detail in Stefanon et al. (2022). The $[3.6]-[5.8]$ color was more poorly constrained in that study due to the lower S/N in the 5.8 μm band likely resulting from splitting the sample across four luminosity bins. At $z\sim 6.5$–8.9, a number of optical emission lines fall within the 5.8 μm band, with Hα and [N II] being expected to contribute the most to the flux measurements (e.g., Anders & Alvensleben 2003).

Under the assumption that the flux density in the 5.8 μm band includes a contribution from Hα, [N II], and the continuum (both stellar and nebular), we measured the EW$_0$(Hα+[N II]) directly from the observed $[3.6]-[5.8]$ color, using the following iterative procedure (see also Stefanon et al. 2022). We initially assumed a flat f_ν continuum calibrated to the flux density in the 3.6 μm band, consistent with the observed SED. We then derived a new value for the continuum by removing the contribution of [O II] and other less prominent lines from the 3.6 μm band flux density, estimated from the $[3.6]-[5.8]$ color, using the line ratios of Anders & Alvensleben (2003) for subsolar metallicity ($Z = 0.2\ Z_\odot$) and negligible dust extinction ($[\mathrm{N\,II}]_{\lambda\lambda6548,6583}/\mathrm{H}\beta = 0.234$; $[\mathrm{O\,II}]/\mathrm{H}\beta = 1.791$). The final value for the EW$_0$(Hα+[N II]) was estimated from the ratio between the newly computed continuum and the flux density in the 5.8 μm band. The EW$_0$(Hα) was derived from the measured EW$_0$(Hα+[N II]) by accounting for the line ratios of Anders & Alvensleben (2003). Uncertainties in the EW(Hα) were computed by randomly perturbing the flux densities according to the corresponding 1σ uncertainties 1000 times and then repeating our procedure for deriving the EW(Hα) based on the 3.6 μm and 5.8 μm flux density measurements. The quoted uncertainties correspond to the 16th and 84th percentiles.

Because the red edge of the 5.8 μm band only extends out to $\sim$6.35 μm, we would only observe light from the Hα line in galaxies out to a redshift of $z\lesssim 8.7$. If any sources from our selection lie at redshifts in excess of this, we would not include the Hα emission line from those sources in our stacks, which could dilute our measurement. Fortunately, only three sources in our sample (out of 102) have $z_{\rm phot} > 8.5$ (and nine sources have $z_{\rm phot} > 8.3$), suggesting that any systematics from this issue is likely small ($\lesssim$3%). Additionally, thanks to the very broad 5.8 μm band ($R(\lambda/\delta\lambda)\sim 5$) relative to typical line widths of sources at these redshifts ($\approx$250 km s^{-1}, e.g., Bouwens et al. 2022c; S. Schouws 2022, in preparation; Béthermin et al. 2020; Cassata et al. 2020; Faisst et al. 2020), i.e., $R\sim 1000$, we can largely ignore the finite width of the Hα lines in our analyses.

The measured $[3.6]-[5.8]$ color corresponds to a rest-frame EW$_0$(Hα+[N II]) $= 2328^{+1326}_{-1127}$ Å, while for the EW$_0$(Hα) alone we derive an extreme EW$_0$(Hα) $= 1960^{+1089}_{-927}$ Å. This estimate corresponds to a luminosity $\log(L_{\rm H\alpha}/[\mathrm{erg\ s^{-1}}]) = 42.62^{+0.15}_{-0.23}$ and an SFR $= 36^{+14}_{-15}\,M_\odot$ yr^{-1} (Kennicutt & Evans 2012).

To our knowledge, this measurement constitutes the first detection of Hα from broadband photometry in the population of star-forming galaxies at $z > 6.5$ typically probed by Lyman-break selections (e.g., Bouwens et al. 2015b, 2022a; Finkelstein et al. 2015; Bowler et al. 2017; Harikane et al. 2022). Recent attempts at measuring the intensity of Hα at $z\sim 8$ for individual sources in HFF cluster fields were unsuccessful (e.g., Asada & Ohta 2022), likely due to the much shallower coverage available in the IRAC 5.8 μm band over those fields combined with the relatively low magnification values for the considered sources.

Such an elevated EW$_0$(Hα) could originate from active galactic nuclei (AGNs). Indeed, indication of nuclear activity in $z\sim 8$ galaxies has recently been found (e.g., Laporte et al. 2017; Mainali et al. 2018; Topping et al. 2021). However, the extrapolation to $z\sim 8$ of recent results at $z\lesssim 7$ (e.g., Harikane

et al. 2022) suggests that AGNs would be a marginal population in the $L < L^*$ galaxies that dominate our sample. A more definitive assessment of the fraction of AGNs in sub-L^* galaxies at $z \sim 8$ requires spectroscopic data, still unavailable for statistically significant samples.

The stacked SED also shows a red $[3.6] - [4.5] = 0.54 \pm 0.13$ mag color. The increased flux density in the 4.5 μm band is likely to result from substantial enhancement by [O III]$_{\lambda\lambda 4959,5007}$ and Hβ line emission. Applying the same measurement procedure adopted for the estimate of the EW$_0$(Hα), the measured [3.6] – [4.5] color corresponds to an EW$_0$([O III]+Hβ) $= 1006^{+230}_{-220}$ Å. This value implies an EW$_0$(Hα) $= 697^{+160}_{-153}$ Å, which we obtained assuming Case B recombination with Hα/H$\beta = 2.85$ (Hummer & Storey 1987), [O III]$_{\lambda 4959}$/H$\beta = 1.6$, [O III]$_{\lambda 5007}$/H$\beta = 4.7$, and a flat f_ν continuum at rest-frame optical wavelengths. Our more direct measurement of EW$_0$(Hα) based on the 5.8 μm band excess differs by only $\sim 1.3\sigma$ from this estimate, providing further confirmation that Hα is very strong in these $z \sim 8$ LBGs. Finally, our stacked SED is characterized by a $J_{125} - H_{160} = -0.10 \pm 0.07$ mag color, indicating a blue UV slope ($\beta \sim -2.4$), and a flat $H_{160} - [3.6] = -0.01 \pm 0.11$ mag, suggesting young stellar population ages. Stefanon et al. (2022) already provide an extensive discussion of the interpretation of stack results involving these bands (J_{125}, H_{160}, 3.6 μm, and 4.5 μm).

5. Discussion

5.1. The EW$_0$(Hα) $\sim$ 2000 Å in the Context of Star-forming Galaxies at z $\sim$ 8

It is interesting to consider where galaxies from our sample fit into the population of star-forming galaxies at $z \sim 8$. We do so by using the parameterization of the main sequence of star-forming galaxies (e.g., Brinchmann et al. 2004; Noeske et al. 2007; Elbaz et al. 2007; Daddi et al. 2007; Speagle et al. 2014; Salmon et al. 2015; Tomczak et al. 2016) of Speagle et al. (2014), extrapolated to $z \sim 8$ by Topping et al. (2022).

The position of the star-forming galaxies in our sample depends on how the SFR is inferred for galaxies in our sample. If the Hα luminosity of our stack is used, our stacked SED lie $\sim$1 dex above the assumed main sequence, suggesting that very active star formation could have been ubiquitous in early galaxies. If on the other hand the UV luminosity of our stack is used, we derive SFR$_{\rm UV} \sim 4.2\ M_\odot\ {\rm yr}^{-1}$, which is in much better agreement with the considered main sequence. However, because the parameters of the star-forming main sequence are still quite uncertain at $z > 7$–8 (see, e.g., Bhatawdekar et al. 2019; Kikuchihara et al. 2020; Stefanon et al. 2021a), it is difficult to come to definitive conclusions.

Interestingly, a number of studies have reported star formation rates estimated from the stellar continuum to systematically differ from those estimated from nebular emission in the early universe (e.g., Shim et al. 2011; Smit et al. 2015; Shivaei et al. 2015; Stefanon et al. 2022). The lack of spectroscopic rest-optical observations precludes our gaining greater insight into the physical mechanisms related to this apparent discrepancy. Possible explanations include, but are not limited to, evolving line ratios, differing dust extinction affecting stellar and nebular light, and rapidly increasing star formation histories. We refer the reader to Shim et al. (2011), Smit et al. (2015), Shivaei et al. (2015), Harikane et al. (2018b), and Stefanon et al. (2022) for more detailed discussions.

5.2. Evolution of the EW$_0$(Hα)

The large EW$_0$(Hα) we infer requires very young stellar populations, $\lesssim$few $\times 10^7$ yr (e.g., Inoue 2011; Wilkins et al. 2020), consistent with our age estimate based on multiband photometry and with recent measurements at similar epochs (e.g., Stefanon et al. 2022; Endsley et al. 2021b; Strait et al. 2020, but see, e.g., Roberts-Borsani et al. 2020, 2022; Tacchella et al. 2022).

In Figure 3 we compare the EW$_0$(Hα) estimate from this study with measurements at $z < 7$ from the literature. Specifically, we considered the estimates from Fumagalli et al. (2012), Sobral et al. (2014), Smit et al. (2016), Bouwens et al. (2016c), Rasappu et al. (2016), Mármol-Queraltó et al. (2016), Faisst et al. (2016, 2019a), Reddy et al. (2018), Harikane et al. (2018a), Lam et al. (2019), and Nanayakkara et al. (2020) (see also Shim et al. 2011; Stark et al. 2013—not plotted). We included only measurements with $M_{\rm UV}$ within $\sim \pm 0.75$ mag of $M_{\rm UV} = -19.9$ mag, or whose stellar masses $M_\star$ lie within ± 1 dex of the stellar mass we estimated from our stacked photometry, when a typical $M_{\rm UV}$ was not quoted with a result. Our estimate constitutes one of the highest EW$_0$(Hα) measurements across the $0 < z < 8$ redshift range. Nonetheless, the present large uncertainty of our measurement makes it consistent at $\sim 1.5\sigma$ with the average EW$_0$(Hα) existing at $z \sim 4$–6.

Given that the samples in Figure 3 were comparably selected over the full redshift range, we can also address the question of the sSFR over this wide time baseline. To a first approximation the EW(Hα) is proportional to the sSFR, and so we can also use the trends seen in Figure 3 to characterize the evolution with redshift of the sSFR. This is shown in the form sSFR $\propto (1+z)^{5/2}$, as derived by Dekel et al. (2013), applying a constant conversion factor to the analytical expression for the evolution of the specific accretion rate of the dark matter halo under the hypothesis of a non-evolving ratio between the stellar mass and the mass of the host dark matter halo. We applied an overall normalization by fitting the curve to the observations, obtaining sSFR $= 5.41 \times (1+z)^{5/2}$. This simple relation can reproduce well the observations at $z \gtrsim 4$, but underestimates the expected EW(Hα) at lower redshifts, with larger gaps for lower redshift values.

While the uncertainty on the present EW$_0$(Hα) measurement is large, one possible explanation for the differential evolution observed between the sSFR and the EW$_0$(Hα) is an $M_\star/L_{\rm optical}$ ratio (where $L_{\rm optical}$ refers to the luminosity of the continuum at wavelengths close to that of Hα) evolving with cosmic time. This is indeed expected considering the increasingly larger fractions of evolved stellar populations at later cosmic times. The black solid curve in Figure 3 presents the result of applying a redshift-dependent $M_\star/L_{\rm optical}$ ratio to the analytical expression for the sSFR(z) evolution. This factor was estimated from our default template set (Section 2), assuming galaxies started forming stars at $z \sim 20$ (e.g., Mawatari et al. 2020; Harikane et al. 2022).[13] We applied a global normalization factor from fitting the curve to the available observations ($\chi_{\rm red} = 4.7$). Here

[13] The exact burst redshift does not significantly influence our conclusions because the time difference between $z \sim 20$ and, e.g., $z \sim 15$ is just less than 100 Myr.

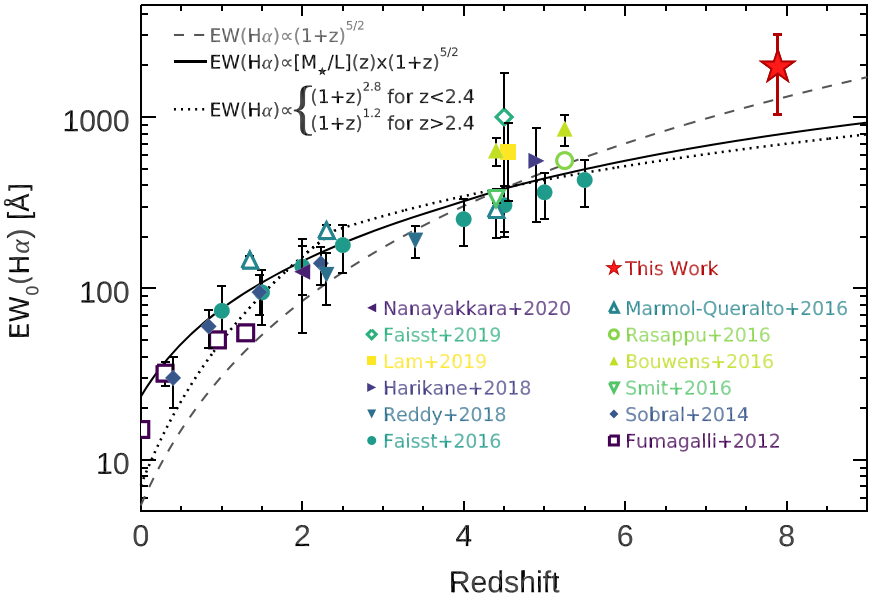

Figure 3. Evolution of the rest-frame $EW_0(H\alpha)$ since $z \sim 8$. Our new $z \sim 8$ measurement is shown relative to other measurements from the literature for $z < 6$, as indicated by the legend. Specifically, we included estimates from Fumagalli et al. (2012), Sobral et al. (2014), Smit et al. (2016), Bouwens et al. (2016c), Rasappu et al. (2016), Mármol-Queraltó et al. (2016), Faisst et al. (2016, 2019a), Reddy et al. (2018), Harikane et al. (2018a), Lam et al. (2019), and Nanayakkara et al. (2020). The dashed gray curve marks the $EW_0(H\alpha)$ expected when its evolution follows that of the sSFR under a scenario of non-evolving star formation efficiency (e.g., Dekel et al. 2013) and constant $M_\star/L_{\rm optical}$ ratio, while the solid black curve is for an $M_\star/L_{\rm optical}$ ratio that increases with decreasing redshift according to a constant star formation history. Finally, the dotted curve corresponds to the double-power-law evolution of the sSFR from Tasca et al. (2015).

we do not consider the effects of dust attenuation given the growing indication that at the stellar masses considered here they are not a significant factor at $z > 2$ (e.g., Bouwens et al. 2016b, 2020; Dunlop et al. 2017; McLure et al. 2018) and marginal at $z < 2$ (e.g., Garn & Best 2010). The curve matches the observations reasonably well, in particular for $z \lesssim 6$, and it can be represented across $0 \lesssim z \lesssim 9$ by the following power law:

$$\mathrm{sSFR} = 25.04 \times (1+z)^{1.59} \quad (1)$$

Our measurement at $z \sim 8$ is consistent at $\gtrsim 1\sigma$ with the values expected from the new relation, although this is due, at least in part, to the large uncertainties. Nonetheless, the overall agreement in the recovered $EW_0(H\alpha)$ with the sSFR evolution to $z \sim 8$ supports a scenario of a marginally evolving star formation efficiency, as suggested by recent observational studies (e.g., Stefanon et al. 2017, 2021a, 2022; Oesch et al. 2018; Harikane et al. 2018a; Bouwens et al. 2021b).

Finally, we note that the compilation of measurements at $z \lesssim 6$ can be modeled similarly well ($\chi_{\rm red} = 4.6$) using the double-power-law evolution of the sSFR presented by Tasca et al. (2015). This sSFR parameterization is characterized by a steep ($\mathrm{sSFR} \propto (1+z)^{2.8}$) growth at $z \leqslant 2.4$, but a significantly shallower evolution ($\mathrm{sSFR} \propto (1+z)^{1.2}$) for $z \geqslant 2.4$. We show this in Figure 3 with a dotted curve, after we applied a normalization factor estimated by fitting the double power law to the measurements (we obtain $EW_0(H\alpha) = 7.06 \times (1+z)^{2.8}$ and $EW_0(H\alpha) = 50.03 \times (1+z)^{1.2}$ for $z \leqslant 2.4$ and $z \geqslant 2.4$, respectively). In particular, the shallower slope at high redshifts, combined with the large uncertainties in the $EW_0(H\alpha)$ measurements, mitigate the requirement for the additional redshift-dependent factor based on the $M_\star/L_{\rm optical}$ ratio we discussed previously. This indicates that on average, the stellar population ages could have evolved only marginally since $z \sim 6$, or, equivalently, it suggests the existence of a significant fraction of evolved stellar populations in $z \gtrsim 5$–6 galaxies. However, we note that for their stellar population analysis, Tasca et al. (2015) assumed an exponentially declining SFH, which might lead to their overestimating the ages of the stellar populations (and consequently underestimating the sSFR), particularly for the highest-redshift galaxies in their sample ($z \sim 5$–6). Forthcoming JWST/NIRCam and NIRSpec observations will enable more accurate determinations of stellar population age by probing the continuum in between lines and therefore providing more robust answers to this tension.

5.3. Implications of the High $EW_0(H\alpha)$ for Reionization

5.3.1. Constraints on ξ_{ion}

We can now use our new determination of the $EW_0(H\alpha)$ to derive the efficiency of production of H-ionizing photons ($\xi_{\rm ion}$). This enables us to quantify a key parameter, the total ionizing power of galaxies, in the heart of the reionization epoch and close to the time of instantaneous reionization ($z = 8.8$, Planck Collaboration et al. 2016). This is particularly valuable since only a few, less direct, measurements exist at $z > 7$ (.e.g., Stark et al. 2015, 2017; De Barros et al. 2019; Endsley et al. 2021a), inferred either from spectral analysis of rest-UV emission lines or from SED fitting to broadband photometry.

Following Bouwens et al. (2016a), we compute $\xi_{\rm ion}$ from the production rate of Lyman-continuum photons, $\dot{N}(\mathrm{H}^0)$. This can be inferred from the Hα luminosity $L(H\alpha)$, using the relation of Leitherer & Heckman (1995):

$$L(H\alpha) = 1.36 \times 10^{-12} \dot{N}(\mathrm{H}^0) \quad (2)$$

where $L(H\alpha)$ has units of erg s^{-1} and $\dot{N}(\mathrm{H}^0)$ units of s^{-1}. This relation has a small ($\lesssim$15% or 0.06 dex) dependence on the metallicity and electron temperature (e.g., Charlot & Longhetti 2001), which we assume as systematic uncertainty. This uncertainty is significantly smaller than the stochastic

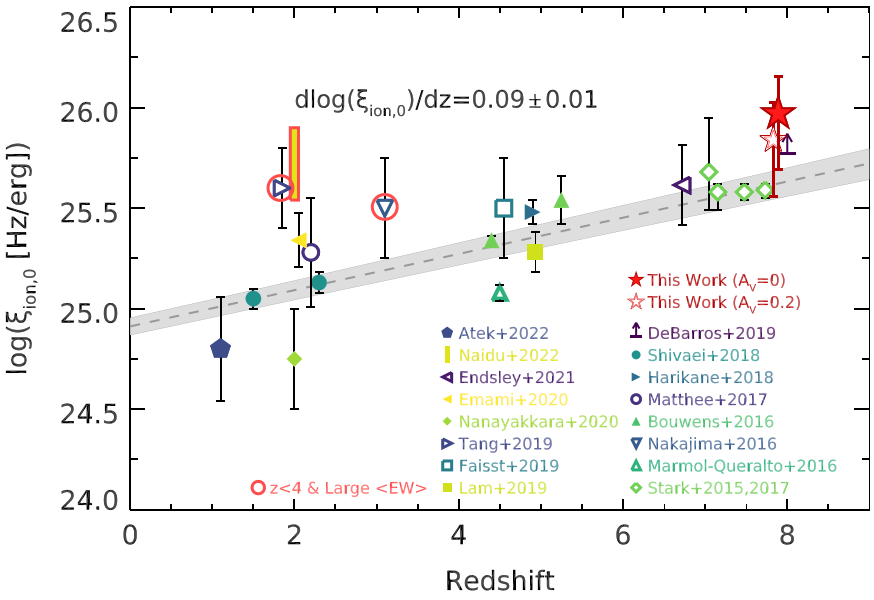

Figure 4. Evolution of $\xi_{\rm ion}$ for $1.5 \lesssim z \lesssim 8$. The compilation of measurements, listed in the legend, includes the results of Stark et al. (2015, 2017), Mármol-Queraltó et al. (2016), Nakajima et al. (2016), Bouwens et al. (2016c), Matthee et al. (2017), Harikane et al. (2018a), Shivaei et al. (2018), De Barros et al. (2019), Lam et al. (2019), Faisst et al. (2019a), Tang et al. (2019), Nanayakkara et al. (2020), Emami et al. (2020), Endsley et al. (2021a), Naidu et al. (2022), and Atek et al. (2022). We only considered measurements corresponding to $M_{\rm UV} \sim -19.9$ mag or $\log(M_\star/M_\odot) \sim 8.1 \pm 1.0$ when the $M_{\rm UV}$ information was missing. The red open circles flag those results at $z < 4$ whose sample was explicitly selected to have rest-optical lines with $\rm EW_0 \gtrsim 1000$ Å (Nakajima et al. 2016; Tang et al. 2019) or with likely a hard ionizing spectrum (Naidu et al. 2022). We arbitrarily shifted by $\Delta z = -0.06$ our measurement for $A_V = 0.2$ mag (open star) to improve the readability. The dashed line marks the result of a linear fit, with the 68% confidence interval encompassed by the gray shaded area. This composite set of measurements suggests a steady increase of $\xi_{\rm ion}$ with increasing redshift.

uncertainties associated with the $\rm EW_0(H\alpha)$ measurement. The LyC photon production efficiency $\xi_{\rm ion,0}$ (where the subscript 0 indicates the assumption of an escape fraction $f_{\rm esc} = 0$, i.e., this is the actual production rate, in the galaxy, excluding any losses) can then be computed as

$$\xi_{\rm ion,0} = \frac{\dot{N}({\rm H}^0)}{L_{\rm UV}} \quad (3)$$

where $L_{\rm UV}$ is the UV-continuum luminosity computed from the stacked SED. The application of the above relations to our measurements yields $\xi_{\rm ion,0} = 10^{25.97^{+0.18}_{-0.28}}$ Hz erg^{-1}, assuming negligible dust attenuation, as expected for $L < L^*$ LBGs at $z > 4$ (e.g., Dunlop et al. 2017; Bouwens et al. 2021b; Casey et al. 2021) and from the extrapolation of the results for Lyα emitters at lower redshifts (e.g., Naidu et al. 2022). If instead we consider a case with a small amount of dust attenuation, we obtain $\xi_{\rm ion,0} = 10^{25.84^{+0.18}_{-0.28}}$ Hz erg^{-1}. In deriving this dust-impacted value we assumed, for simplicity, the Calzetti et al. (2000) attenuation law and the same $A_V = 0.2$ mag value for both the stellar continuum and nebular emission, given the relative contribution of the two components is still quite uncertain (e.g., Buat et al. 2018; Shivaei et al. 2020; Reddy et al. 2020; Li et al. 2021 and references therein). These high values of $\xi_{\rm ion}$ require young stellar populations (ages $\lesssim 10^7$ yr, e.g., Robertson 2021), consistent with the values we find from our SED fitting (see also Stefanon et al. 2022).

In Figure 4 we compare the value of $\xi_{\rm ion,0}$ from this study to previous estimates at similar redshifts and down to $z \sim 2$ (Stark et al. 2015, 2017; Mármol-Queraltó et al. 2016; Nakajima et al. 2016; Bouwens et al. 2016c; Matthee et al. 2017; Harikane et al. 2018a; Shivaei et al. 2018; De Barros et al. 2019; Lam et al. 2019; Faisst et al. 2019; Tang et al. 2019; Nanayakkara et al. 2020; Emami et al. 2020; Endsley et al. 2021a; Naidu et al. 2022; Atek et al. 2022). We only considered $\xi_{\rm ion}$ estimates that refer to samples with either $M_{\rm UV}$ within ± 1 mag of the UV luminosity for our stack or $\log M_\star$ within ± 1 dex of the stellar mass we estimated with FAST.

Our measurement is consistent with the estimates existing at $z \sim 6.5$–8 (Stark et al. 2015, 2017; De Barros et al. 2019; Endsley et al. 2021a). Their estimates derived, respectively, from modeling the intense C IV λ1548 Å line identified in the spectrum of a $z = 7.045$ galaxy (Stark et al. 2015), from the [C III] and Lyα lines in three $z \sim 7$ galaxies with evidence for significant [O III] emission as suggested by their IRAC colors (Stark et al. 2017), and from SED fitting to multiwavelength photometry (De Barros et al. 2019; Endsley et al. 2021a). Our measurement is also broadly consistent with the estimates at $z \sim 5$ (Bouwens et al. 2016c; Harikane et al. 2018a; Lam et al. 2019; Faisst et al. 2019); see also Vanzella et al. (2018) for a similarly high $\xi_{\rm ion}$ value in a LyC leaker at $z \sim 4$ and Maseda et al. (2020) for exceptionally high $\xi_{\rm ion} \approx 10^{26.3}$ Hz erg^{-1} in lower-mass galaxies at $z \sim 4$–5). Overall, these results suggest that $\xi_{\rm ion} \approx 10^{25.6\text{–}25.8}$ Hz erg^{-1} could be typical at these epochs, and that $\xi_{\rm ion} \approx 10^{25.7}$ might represent a reasonable estimate.

The values for $\xi_{\rm ion}$ at $z \sim 2$–3 are characterized by a large dispersion of $\approx 10^{24.7}$–$10^{26.0}$ Hz erg^{-1}. This distribution could be explained at least in part by selection effects. Remarkably, the values of $\xi_{\rm ion}$ from samples characterized by high EW line emission (low-z analogues: Nakajima et al. 2016; Tang et al. 2019, see also Chevallard et al. 2018 for similar values at $z \sim 0$), and from Lyα emitters (Naidu et al. 2022) are consistent with those found at $z \sim 7$–8. Instead, the $\xi_{\rm ion}$ estimated from more inclusive samples are generally lower (Matthee et al. 2017; Shivaei et al. 2018; Nanayakkara et al. 2020; Emami et al. 2020; Atek et al. 2022). A formal fit to the evolution of $\xi_{\rm ion}$, after excluding values from high EW samples, results in

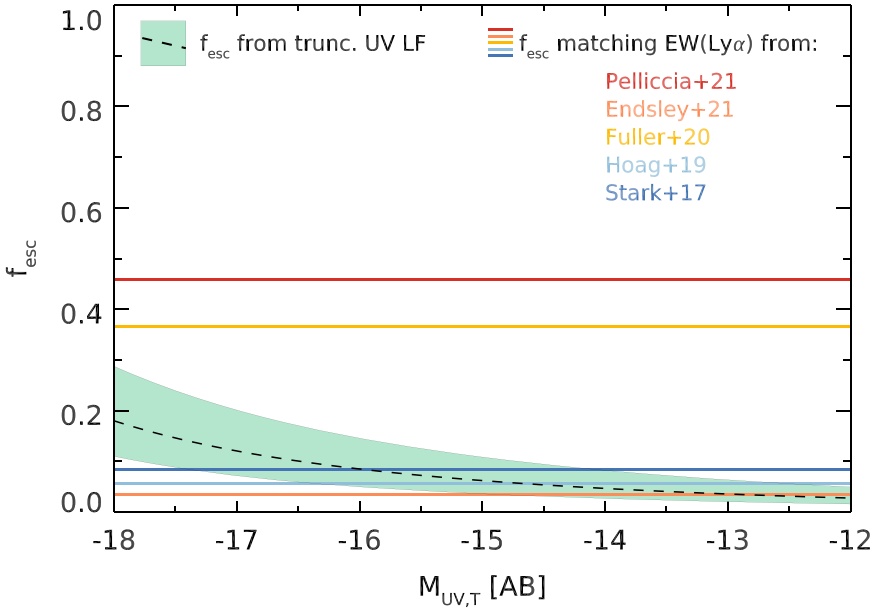

Figure 5. The black dashed curve and green shaded area mark the escape fraction and 68% confidence interval estimated as a function of the UV LF turnover magnitude M_T, as set for the required number density of ionizing photons per second $n_{\rm ion}$, and based on the result for $\xi_{\rm ion,0}$ from this study. The horizontal lines correspond to the $f_{\rm esc}$ estimated by matching the rest-frame EW$_0$(Lyα) inferred from our Hα measurement to a compilation of values from the literature. Specifically, we considered the sample averages from Stark et al. (2017), Hoag et al. (2019), and Endsley et al. (2021a), the single $z > 7$ galaxy in the sample of Fuller et al. (2020) (C14215A1), and RELICS-DP7 from Pelliccia et al. (2021).

$\log(\xi_{\rm ion}/[\rm Hz\ erg^{-1}]) = (0.09 \pm 0.01)z + (24.82 \pm 0.08)$, whose slope is consistent with predictions from recent models (e.g., Finkelstein et al. 2019; Yung et al. 2020, but see, e.g., Matthee et al. 2022 for a non-evolving $\xi_{\rm ion}$ model). The high value for $\xi_{\rm ion}$ resulting from large EW samples at all redshifts $z > 2$ does indicate that we may be settling on a value broadly appropriate for early times for sub-L^* LBGs.

5.3.2. *Implications for* f_{esc}

The number density of ionizing photons per second that can reionize the intergalactic hydrogen is generally expressed as $n_{\rm ion} = \rho_{\rm UV}\xi_{\rm ion}f_{\rm esc}$ (e.g., Robertson et al. 2013), where $\rho_{\rm UV}$ is the luminosity density at rest-frame UV (≈1500–1600 Å) and $f_{\rm esc}$ is the fraction of ionizing photons escaping into the intergalactic medium. Because it is still uncertain whether the faint end of the $z \sim 7$–8 UV luminosity function (LF) presents a turnover (e.g., Atek et al. 2015; Livermore et al. 2017; Bouwens et al. 2017, 2021b; Yue et al. 2018; Bhatawdekar et al. 2019), here we explore the impact that different values of turnover have on $f_{\rm esc}$. Our estimates are based on the requirement that all the necessary H-ionizing radiation for reionization is generated by stars. For this, we assumed $n_{\rm ion} = 10^{50.75}$ s^{-1} Mpc^{-3} (e.g., Bouwens et al. 2015a; Finkelstein et al. 2019; Mason et al. 2019; Naidu et al. 2020) and $\xi_{\rm ion,0}$ from this study. In the computation of $\rho_{\rm UV}$, we approximated the turnover by truncating the $z \sim 8$ UV LF of Bouwens et al. (2021b) at values spanning $-18 < M_T < -12$ mag.

The result of this procedure is presented in Figure 5. The larger value we find for $\xi_{\rm ion}$ translates into $f_{\rm esc} \lesssim 30\%$ for $M_{\rm UV} > -18$ mag. Values $f_{\rm esc} \lesssim 20\%$ have been inferred at $z > 6$ by recent studies (e.g., Castellano et al. 2017). In particular, absence of a turnover in the faint-end slope down to $M_T \sim -13$ mag would only require $f_{\rm esc} \sim 5\%$–10% to fully ionize the neutral H at $z \sim 8$. These values are consistent with the $f_{\rm esc} \sim 5\%$–10% inferred for sub-L^* LBGs at $z \lesssim 4$ by an increasing number of studies (e.g., Marchi et al. 2017; Naidu et al. 2018; Pahl et al. 2021; Siana et al. 2010; Grazian et al. 2016, 2017; Rutkowski et al. 2016; Steidel et al. 2018). Together they suggest only a marginal evolution of $f_{\rm esc}$ with cosmic time for the average galaxy population. Furthermore, and qualitatively, such small values of escape fraction can more easily be reconciled with the strong emission lines inferred at rest-frame optical for $z \gtrsim 7$ galaxies (EW$_0$([O III]+Hβ) > 1000 Å, e.g., Smit et al. 2014; Castellano et al. 2017; De Barros et al. 2019; Stefanon et al. 2019, 2022; Bowler et al. 2020; Strait et al. 2020, 2021; Endsley et al. 2021b). The overall consistency is reinforced by considering that photoionization modeling suggests that the production of such strong emission lines already requires a very young ($\lesssim 10^7$ yr) stellar population (e.g., Inoue 2011; Wilkins et al. 2020).

An increasing number of studies are identifying Lyα emission in $z \gtrsim 7$ galaxies, with rest-frame EW ranging from ≈5–20 Å to >100–200 Å (e.g., Pentericci et al. 2014; Stark et al. 2017; Hoag et al. 2019; Fuller et al. 2020; Endsley et al. 2021b; Pelliccia et al. 2021; Larson et al. 2022). A direct conversion of our EW$_0$(Hα) using Case B recombination coefficients ($L_{\rm Ly\alpha}/L_{\rm H\alpha} = 8.7$) suggests an *intrinsic* EW$_0$(Lyα)$_{\rm intrinsic}$ $= 517^{+287}_{-244}$ Å. Under the assumption that the fraction of escaping LyC photons is approximately similar to that of Lyα emission (e.g., Steidel et al. 2018; Izotov et al. 2020), the implied $f_{\rm esc}$ ranges between $\lesssim$10% and ≈40%–50% (Figure 5). These estimates are likely upper limits, given the still significant fractions of nondetection, particularly for sub-L^* galaxies (e.g., Pentericci et al. 2014; Jung et al. 2022). Thus the preliminary indications from Lyα studies are consistent with the more direct estimates, and reinforce the likely ready availability of adequate reionizing photons from star-forming galaxies for reionization. To give a sense of sensitivity to any dust, a thin $A_V = 0.2$ mag Calzetti et al. (2000) dust screen would lower the Lyα flux by ∼1.7× (resulting in EW$_0$(Lyα) =

298^{+165}_{-141} Å), and increase the requirement on f_{esc} by the same factor. Since very low dust absorption is likely, this suggests that any likely levels of dust would not change our conclusions significantly.

6. Conclusions

Our analysis of the deepest Spitzer/IRAC data available over extragalactic fields for a large sample of $z \sim 8$ LBGs has allowed us to detect and measure for the first time the flux in the Hα line in the early universe, and to explore the resulting implications for reionization. Specifically, we obtained this measurement through a median stacking of Hubble and IRAC data for 102 LBGs initially identified by Bouwens et al. (2015b) from Hubble imaging over the CANDELS GOODS-N/S, ERS, XDF, CANDELS/UDS, and CANDELS/COSMOS fields. Stefanon et al. (2022) had previously used a similar median stacking procedure to study the main properties of this sample of $z \sim 8$ star-forming galaxies as a function of UV luminosity. These fields have deep coverage in the HST/ACS V_{606} and I_{814} and HST/WFC3 Y_{105}, J_{125}, JH_{140}, and H_{160} bands. Key for our current study are that these fields also have deep Spitzer/IRAC mosaics from the GOODS Reionization Era Wide Area Treasury from Spitzer (GREATS, PI: Labbé; Stefanon et al. 2021b). These mosaics combine all the relevant observations acquired with IRAC in the 3.6, 4.5, 5.8, and 8.0 μm bands over the GOODS-N/S fields across the full scientific lifetime of Spitzer. In particular, the GREATS 5.8 μm imaging is the deepest data available at $\approx$6 μm before JWST, and represents a unique opportunity to probe Hα at $6.8 \lesssim z \lesssim 8.7$.

We extracted median flux densities in the IRAC bands after combining the image stamps cleaned from neighbor contamination through MOPHONGO (Labbé et al. 2006, 2010a, 2010b, 2013, 2015), and recovered total flux densities using the location-specific PSFs from GREATS. Our main results are the following.

1. Our stack results for 102 galaxies at $z \sim 8$ show a 4.3σ detection in the 5.8 μm band, and a red $[3.6]-[5.8] = 0.82 \pm 0.27$ mag color.
2. Interpreting the excess in the 5.8 μm band as due to emission from Hα, we infer a rest-frame $EW_0(H\alpha) = 1960^{+1089}_{-927}$ Å, corresponding to a luminosity of $\log(L_{H\alpha}/[\text{erg s}^{-1}]) = 42.62^{+0.15}_{-0.23}$. Our result represents the first direct determination of the Hα intensity at $z > 6.5$.

These results allow us to draw the following conclusions.

1. Comparison of our new $EW_0(H\alpha)$ measurement with previous determinations at lower redshifts from the literature suggests that the trend of increasing $EW_0(H\alpha)$ with redshift (e.g., Faisst et al. 2016) can be extended up to $z \sim 8$.
2. After accounting for an $M_\star/L_{optical}$ ratio that depends on cosmic time, the observed evolution with redshift of $EW_0(H\alpha)$ is consistent with the evolution of the specific accretion rate of the dark matter halos, providing further evidence that the star formation efficiency is at most marginally evolving with cosmic time in the early universe.
3. Following the formalism of Bouwens et al. (2016a), our new measurement of $L_{H\alpha}$ implies an efficiency of production of LyC photon $\xi_{ion,0} = 10^{25.97^{+0.18}_{-0.28}}$ Hz erg^{-1}. This constitutes one of the largest ξ_{ion} estimates at $0 < z < 8$ for sub-L^* galaxies ($M_{UV} \sim -19.8$ mag, $M_\star \approx 10^8$ $M_\odot$). While the uncertainties are large, our new measurement is very consistent with previous estimates at similar redshifts, at $z \sim 5$, and with those values at lower redshift inferred from samples with significant nebular line emission. This consistency is not only reassuring but also points to a surprising uniformity across billions of years for star-forming galaxies.
4. The large value of ξ_{ion} we find suggests that escape fractions $f_{esc} \lesssim 10\%$ are sufficient for star-forming galaxies to fully ionize the neutral H at $z \sim 8$ through escaping LyC radiation. The small value of f_{esc} is consistent with what is seen at lower redshifts $z \sim 2$–6 in star-forming galaxies, reinforcing the likelihood that galaxies alone are responsible for reionization.

It is remarkable to step back and realize that this study was enabled by observations in the 5.8 μm band, acquired during the first few years of Spitzer scientific operations, a decade and a half ago. The present results highlight once again how powerful and pivotal a small telescope like Spitzer has been, especially when able to leverage robust selections made possible with HST. Fortunately, JWST/NIRSpec, NIRCam, and MIRI combine and enhance the capabilities of HST and Spitzer, providing the potential for absolutely game-changing science in the coming years.

The authors would like to thank the referee for their careful reading and for the positive and constructive comments that helped to improve the quality of the paper. M.S. acknowledges support from the CIDEGENT grant CIDEGENT/2021/059, and from project PID2019-109592GB-I00/AEI/10.13039/501100011033 from the Spanish Ministerio de Ciencia e Innovación – Agencia Estatal de Investigación. M.S. and R.J.B. acknowledge support from TOP grant TOP1.16.057. P.A.O. acknowledges support from the Swiss National Science Foundation through the SNSF Professorship grant 190079 "Galaxy Build-up at Cosmic Dawn." The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant No. 140. We also acknowledge the support of NASA grants HSTAR-13252, HST-GO-13872, HST-GO-13792, and NWO grant 600.065.140.11N211 (vrij competitie). G.D.I. acknowledges support for GREATS under RSA No. 1525754. V.G. gratefully acknowledges support by the ANID BASAL projects ACE210002 and FB210003. This paper utilizes observations obtained with the NASA/ESA Hubble Space Telescope, retrieved from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute (STScI). STScI is operated by the Association of Universities for Research in Astronomy, Inc. under NASA contract NAS 5-26555. This work is based [in part] on observations made with the Spitzer Space Telescope, which was operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. Support for this work was provided by NASA through an award issued by JPL/Caltech.

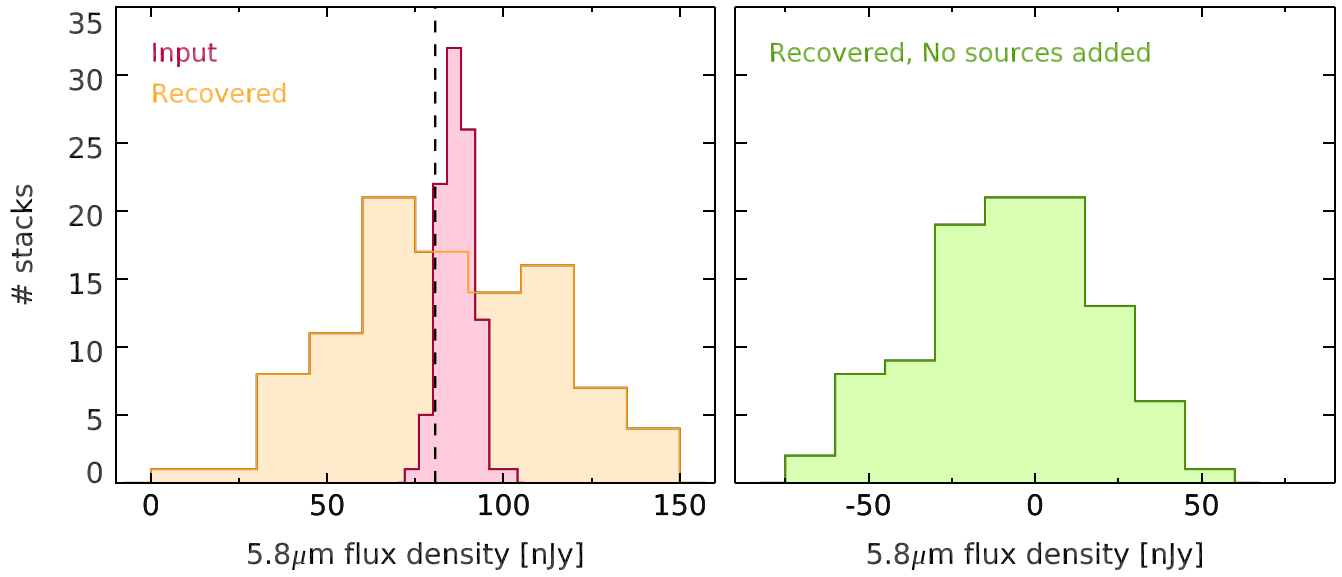

Figure 6. Left panel: comparison between the distribution of flux densities adopted as input for our Monte Carlo simulations (red histogram) and that after recovering the synthetic sources following the same procedure we adopted for our main analysis (yellow histogram—see also Section 3). The vertical black dashed line indicates the median of our measurements. Right panel: distribution of flux densities measured by stacking neighbor-cleaned stamps centered at locations free from existing sources. These two results clearly indicate that our 5.8 μm flux density measurement is genuine and any systematics resulting from the imperfect subtraction of neighboring sources are negligible.

Appendix
Validation of the Flux Density Measured in the 5.8 μm Band

We performed a Monte Carlo simulation to ascertain whether the signal detected in the 5.8 μm band is genuine emission from the sample of LBGs at $z \sim 8$ and not the result of residual contamination from neighboring sources. We generated a new set of 102 flux densities in the H_{160} band by randomly scattering the H_{160} measurements of our sample according to their associated uncertainties. We then computed the flux densities in the 5.8 μm band by assuming a constant ratio $f_{5.8}/f_{160} = 2.2$ between the flux density in the 5.8 μm band ($f_{5.8}$) and that in the H_{160} band (f_{160}), consistent with what we measure in our stack. This assumption is equivalent to a $z \sim 8$ flat f_ν SED, with a rest-frame EW$_0$(Hα+[N II]) = 2300 Å line emission contributing to the 5.8 μm flux density. Point sources having the previously computed 5.8 μm flux densities were then added at random locations across the four mosaics, adopting the location-specific PSFs from GREATS. In doing this, we preserved the relative fraction and luminosity distribution of sources in each field present in our original sample. Our adoption of point sources is supported by the smaller sizes ($R_e \lesssim 1$ kpc, corresponding to $\approx 0\farcs2$ at $z \sim 8$) of sub-L^* galaxies at these redshifts (e.g., Shibuya et al. 2015; Bouwens et al. 2021a, 2022b), compared to the 5.8 μm PSF FWHM ($\approx 2\farcs0$). Following the same procedure we implemented for our main analysis (Section 3), we then constructed neighbor-cleaned stamps using MOPHONGO, and extracted the photometry of the median stack adopting a $2\farcs0$ diameter aperture, correcting to total using the reconstructed PSF. All these steps were repeated 100 times. The results of this simulation are presented in the left panel of Figure 6. The simulated measurements have a dispersion of $\sigma = 26$ nJy, after the dispersion of the input measurements ($\sigma \sim 6$ nJy) has been subtracted in quadrature, and are only modestly larger than the 20 nJy we estimated from our main stacking analysis. This clearly shows that our analysis is able to recover the median of the input flux densities. To test the amount of systematics that non-optimal removal of neighboring sources could introduce into our measurements, we also extracted the photometry from neighbor-cleaned stacks centered on locations free from sources, as inferred from the combination of the J_{125}, JH_{140}, and H_{160} band mosaics (i.e., this is equivalent to adopting $f_{5.8}/f_{160} = 0$). The outcome of this second experiment is shown in the right panel of Figure 6. As we might expect if there is negligible contamination from the neighboring sources, the measurements are normally distributed around 0 nJy. The present Monte Carlo simulation results significantly increase our confidence in the overall robustness of our 5.8 μm band flux measurements.

ORCID iDs

Mauro Stefanon https://orcid.org/0000-0001-7768-5309
Rychard J. Bouwens https://orcid.org/0000-0002-4989-2471
Garth D. Illingworth https://orcid.org/0000-0002-8096-2837
Ivo Labbé https://orcid.org/0000-0002-2057-5376
Pascal A. Oesch https://orcid.org/0000-0001-5851-6649
Valentino Gonzalez https://orcid.org/0000-0002-3120-0510

References

Anders, P., Fritze, v., & Alvensleben, U. 2003, A&A, 401, 1063
Asada, Y., & Ohta, K. 2022, ApJ, 924, 71
Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2013, ApJ, 769, 80
Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2015, ApJS, 218, 33
Atek, H., Furtak, L. J., Oesch, P., et al. 2022, MNRAS, 511, 4464
Atek, H., Richard, J., Kneib, J.-P., et al. 2015, ApJ, 800, 18
Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., et al. 2006, AJ, 132, 1729
Béthermin, M., Fudamoto, Y., Ginolfi, M., et al. 2020, A&A, 643, A2
Bhatawdekar, R., Conselice, C. J., Margalef-Bentabol, B., & Duncan, K. 2019, MNRAS, 486, 3805
Bouwens, R., González-López, J., Aravena, M., et al. 2020, ApJ, 902, 112
Bouwens, R. J., Aravena, M., Decarli, R., et al. 2016b, ApJ, 833, 72

Bouwens, R. J., Illingworth, G. D., Ellis, R. S., Oesch, P. A., & Stefanon, M. 2022a, arXiv:2205.11526
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011, ApJ, 737, 90
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015a, ApJ, 811, 140
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015b, ApJ, 803, 34
Bouwens, R. J., Illingworth, G. D., van Dokkum, P. G., et al. 2021a, AJ, 162, 255
Bouwens, R. J., Illingworth, G. D., van Dokkum, P. G., et al. 2022b, ApJ, 927, 81
Bouwens, R. J., Oesch, P. A., Illingworth, G. D., Ellis, R. S., & Stefanon, M. 2017, ApJ, 843, 129
Bouwens, R. J., Oesch, P. A., Labbé, I., et al. 2016c, ApJ, 830, 67
Bouwens, R. J., Oesch, P. A., Stefanon, M., et al. 2021b, AJ, 162, 47
Bouwens, R. J., Smit, R., Labbé, I., et al. 2016a, ApJ, 831, 176
Bouwens, R. J., Smit, R., Schouws, S., et al. 2022c, ApJ, 931, 160
Bowler, R. A. A., Dunlop, J. S., McLure, R. J., & McLeod, D. J. 2017, MNRAS, 466, 3612
Bowler, R. A. A., Jarvis, M. J., Dunlop, J. S., et al. 2020, MNRAS, 493, 2059
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151
Brinchmann, J., Pettini, M., & Charlot, S. 2008, MNRAS, 385, 769
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Buat, V., Boquien, M., Małek, K., et al. 2018, A&A, 619, A135
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Caputi, K. I., Cirasuolo, M., Dunlop, J. S., et al. 2011, MNRAS, 413, 162
Caputi, K. I., Deshmukh, S., Ashby, M. L. N., et al. 2017, ApJ, 849, 45
Casey, C. M., Zavala, J. A., Manning, S. M., et al. 2021, ApJ, 923, 215
Cassata, P., Morselli, L., Faisst, A., et al. 2020, A&A, 643, A6
Castellano, M., Pentericci, L., Fontana, A., et al. 2017, ApJ, 839, 73
Charlot, S., & Longhetti, M. 2001, MNRAS, 323, 887
Chevallard, J., Charlot, S., Senchyna, P., et al. 2018, MNRAS, 479, 3264
Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ, 670, 156
De Barros, S., Oesch, P. A., Labbé, I., et al. 2019, MNRAS, 489, 2355
Dekel, A., Zolotov, A., Tweed, D., et al. 2013, MNRAS, 435, 999
Dunlop, J. S., McLure, R. J., Biggs, A. D., et al. 2017, MNRAS, 466, 861
Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, A&A, 468, 33
Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017, PASA, 34, e058
Ellis, R. S., McLure, R. J., Dunlop, J. S., et al. 2013, ApJL, 763, L7
Emami, N., Siana, B., Alavi, A., et al. 2020, ApJ, 895, 116
Endsley, R., Stark, D. P., Charlot, S., et al. 2021a, MNRAS, 502, 6044
Endsley, R., Stark, D. P., Chevallard, J., & Charlot, S. 2021b, MNRAS, 500, 5229
Faisst, A. L., Capak, P. L., Emami, N., Tacchella, S., & Larson, K. L. 2019, ApJ, 884, 133
Faisst, A. L., Capak, P., Hsieh, B. C., et al. 2016, ApJ, 821, 122
Faisst, A. L., Schaerer, D., Lemaux, B. C., et al. 2020, ApJS, 247, 61
Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, RMxAA, 53, 385
Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, RMxAA, 49, 137
Finkelstein, S. L., D'Aloisio, A., Paardekooper, J.-P., et al. 2019, ApJ, 879, 36
Finkelstein, S. L., Ryan, R. E., Jr., Papovich, C., et al. 2015, ApJ, 810, 71
Fuller, S., Lemaux, B. C., Bradač, M., et al. 2020, ApJ, 896, 156
Fumagalli, M., Patel, S. G., Franx, M., et al. 2012, ApJL, 757, L22
Garn, T., & Best, P. N. 2010, MNRAS, 409, 421
Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJL, 600, L93
Grazian, A., Giallongo, E., Gerbasi, R., et al. 2016, A&A, 585, A48
Grazian, A., Giallongo, E., Paris, D., et al. 2017, A&A, 602, A18
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Harikane, Y., Ono, Y., Ouchi, M., et al. 2022, ApJS, 259, 20
Harikane, Y., Ouchi, M., Ono, Y., et al. 2018a, PASJ, 70, S11
Harikane, Y., Ouchi, M., Shibuya, T., et al. 2018b, ApJ, 859, 84
Hoag, A., Bradač, M., Huang, K., et al. 2019, ApJ, 878, 12
Hummer, D. G., & Storey, P. J. 1987, MNRAS, 224, 801
Illingworth, G. D., Magee, D., Oesch, P. A., et al. 2013, ApJS, 209, 6
Inoue, A. K. 2011, MNRAS, 415, 2920
Izotov, Y. I., Schaerer, D., Worseck, G., et al. 2020, MNRAS, 491, 468
Jung, I., Papovich, C., Finkelstein, S. L., et al. 2022, ApJ, 933, 87
Kennicutt, R. C., & Evans, N. J. 2012, ARA&A, 50, 531
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Kewley, L. J., Zahid, H. J., Geller, M. J., et al. 2015, ApJL, 812, L20
Kikuchihara, S., Ouchi, M., Ono, Y., et al. 2020, ApJ, 893, 60
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Krick, J. E., Surace, J. A., Thompson, D., et al. 2009, ApJS, 185, 85
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Labbé, I., Bouwens, R., Illingworth, G. D., & Franx, M. 2006, ApJL, 649, L67
Labbé, I., González, V., Bouwens, R. J., et al. 2010a, ApJL, 716, L103
Labbé, I., González, V., Bouwens, R. J., et al. 2010b, ApJL, 708, L26
Labbé, I., Oesch, P. A., Bouwens, R. J., et al. 2013, ApJL, 777, L19
Labbé, I., Oesch, P. A., Illingworth, G. D., et al. 2015, ApJS, 221, 23
Lam, D., Bouwens, R. J., Labbé, I., et al. 2019, A&A, 627, A164
Laporte, N., Nakajima, K., Ellis, R. S., et al. 2017, ApJ, 851, 40
Larson, R. L., Finkelstein, S. L., Hutchison, T. A., et al. 2022, ApJ, 930, 104
Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599
Leitherer, C., & Heckman, T. M. 1995, ApJS, 96, 9
Leja, J., Johnson, B. D., Conroy, C., et al. 2019, ApJ, 877, 140
Li, N., Li, C., Mo, H., et al. 2021, ApJ, 917, 72
Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, ApJ, 835, 113
Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, ApJ, 837, 97
Lower, S., Narayanan, D., Leja, J., et al. 2020, ApJ, 904, 33
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Mainali, R., Zitrin, A., Stark, D. P., et al. 2018, MNRAS, 479, 1180
Marchi, F., Pentericci, L., Guaita, L., et al. 2017, A&A, 601, A73
Mármol-Queraltó, E., McLure, R. J., Cullen, F., et al. 2016, MNRAS, 460, 3587
Maseda, M. V., Bacon, R., Lam, D., et al. 2020, MNRAS, 493, 5120
Mason, C. A., Naidu, R. P., Tacchella, S., & Leja, J. 2019, MNRAS, 489, 2669
Matthee, J., Naidu, R. P., Pezzulli, G., et al. 2022, MNRAS, 512, 5960
Matthee, J., Sobral, D., Best, P., et al. 2017, MNRAS, 465, 3637
Mawatari, K., Inoue, A. K., Hashimoto, T., et al. 2020, ApJ, 889, 137
McLure, R. J., Dunlop, J. S., Cullen, F., et al. 2018, MNRAS, 476, 3991
Moustakas, J., Kennicutt, R. C. J., & Tremonti, C. A. 2006, ApJ, 642, 775
Naidu, R. P., Forrest, B., Oesch, P. A., Tran, K.-V. H., & Holden, B. P. 2018, MNRAS, 478, 791
Naidu, R. P., Matthee, J., Oesch, P. A., et al. 2022, MNRAS, 510, 4582
Naidu, R. P., Tacchella, S., Mason, C. A., et al. 2020, ApJ, 892, 109
Nakajima, K., Ellis, R. S., Iwata, I., et al. 2016, ApJL, 831, L9
Nanayakkara, T., Brinchmann, J., Glazebrook, K., et al. 2020, ApJ, 889, 180
Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJL, 660, L43
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Labbé, I., & Stefanon, M. 2018, ApJ, 855, 105
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Pahl, A. J., Shapley, A., Steidel, C. C., Chen, Y., & Reddy, N. A. 2021, MNRAS, 505, 2447
Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, ApJ, 559, 620
Pelliccia, D., Strait, V., Lemaux, B. C., et al. 2021, ApJL, 908, L30
Pentericci, L., Vanzella, E., Fontana, A., et al. 2014, ApJ, 793, 113
Planck Collaboration, Adam, R., Aghanim, N., et al. 2016, A&A, 596, A108
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
Rasappu, N., Smit, R., Labbé, I., et al. 2016, MNRAS, 461, 3886
Reddy, N. A., Shapley, A. E., Kriek, M., et al. 2020, ApJ, 902, 123
Reddy, N. A., Shapley, A. E., Sanders, R. L., et al. 2018, ApJ, 869, 92
Roberts-Borsani, G., Morishita, T., Treu, T., Leethochawalit, N., & Trenti, M. 2022, ApJ, 927, 236
Roberts-Borsani, G. W., Bouwens, R. J., Oesch, P. A., et al. 2016, ApJ, 823, 143
Roberts-Borsani, G. W., Ellis, R. S., & Laporte, N. 2020, MNRAS, 497, 3440
Robertson, B. E. 2021, arXiv:2110.13160
Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, ApJ, 768, 71
Rutkowski, M. J., Scarlata, C., Haardt, F., et al. 2016, ApJ, 819, 81
Salmon, B., Papovich, C., Finkelstein, S. L., et al. 2015, ApJ, 799, 183
Salpeter, E. E. 1955, ApJ, 121, 161
Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJS, 172, 86
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Shibuya, T., Ouchi, M., & Harikane, Y. 2015, ApJS, 219, 15
Shim, H., Chary, R.-R., Dickinson, M., et al. 2011, ApJ, 738, 69
Shivaei, I., Reddy, N., Rieke, G., et al. 2020, ApJ, 899, 117
Shivaei, I., Reddy, N. A., Siana, B., et al. 2018, ApJ, 855, 42
Shivaei, I., Reddy, N. A., Steidel, C. C., & Shapley, A. E. 2015, ApJ, 804, 149
Siana, B., Teplitz, H. I., Ferguson, H. C., et al. 2010, ApJ, 723, 241
Smit, R., Bouwens, R. J., Franx, M., et al. 2015, ApJ, 801, 122
Smit, R., Bouwens, R. J., Labbé, I., et al. 2014, ApJ, 784, 58
Smit, R., Bouwens, R. J., Labbé, I., et al. 2016, ApJ, 833, 254
Sobral, D., Best, P. N., Smail, I., et al. 2014, MNRAS, 437, 3516
Sobral, D., Kohn, S. A., Best, P. N., et al. 2016, MNRAS, 457, 1739
Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15
Stark, D. P., Ellis, R. S., Charlot, S., et al. 2017, MNRAS, 464, 469
Stark, D. P., Schenker, M. A., Ellis, R., et al. 2013, ApJ, 763, 129
Stark, D. P., Walth, G., Charlot, S., et al. 2015, MNRAS, 454, 1393
Stefanon, M., Bouwens, R. J., Labbé, I., et al. 2017, ApJ, 843, 36
Stefanon, M., Bouwens, R. J., Labbé, I., et al. 2021a, ApJ, 922, 29
Stefanon, M., Bouwens, R. J., Labbé, I., et al. 2022, ApJ, 927, 48

Stefanon, M., Labbé, I., Bouwens, R. J., et al. 2019, ApJ, 883, 99
Stefanon, M., Labbé, I., Oesch, P. A., et al. 2021b, ApJS, 257, 68
Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2018, ApJ, 869, 123
Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, ApJ, 795, 165
Strait, V., Bradač, M., Coe, D., et al. 2020, ApJ, 888, 124
Strait, V., Bradač, M., Coe, D., et al. 2021, ApJ, 910, 135
Tacchella, S., Finkelstein, S. L., Bagley, M., et al. 2022, ApJ, 927, 170
Tang, M., Stark, D. P., Chevallard, J., & Charlot, S. 2019, MNRAS, 489, 2572
Tasca, L. A. M., Le Fèvre, O., Hathi, N. P., et al. 2015, A&A, 581, A54
Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2016, ApJ, 817, 118
Topping, M. W., Shapley, A. E., Stark, D. P., et al. 2021, ApJL, 917, L36
Topping, M. W., Stark, D. P., Endsley, R., et al. 2022, arXiv:2203.07392
Vanzella, E., Nonino, M., Cupani, G., et al. 2018, MNRAS, 476, L15
Whitler, L., Stark, D. P., Endsley, R., et al. 2022, arXiv:2206.05315
Wilkins, S. M., Lovell, C. C., Fairhurst, C., et al. 2020, MNRAS, 493, 6079
Windhorst, R. A., Cohen, S. H., Hathi, N. P., et al. 2011, ApJS, 193, 27
Yue, B., Castellano, M., Ferrara, A., et al. 2018, ApJ, 868, 115
Yung, L. Y. A., Somerville, R. S., Popping, G., & Finkelstein, S. L. 2020, MNRAS, 494, 1002