MNRAS **000**, 1–13 (2018) Preprint 26 March 2019 Compiled using MNRAS LATEX style file v3.0

The GREATS Hβ+[O III] Luminosity Function and Galaxy Properties at z ∼ 8: Walking the Way of *JWST*

S. De Barros[1]★, P. A. Oesch[1,2], I. Labbé[3], M. Stefanon[4], V. González[5,6], R. Smit[7,8], R. J. Bouwens[4], G. D. Illingworth[9]

[1] *Département d'Astronomie, Université de Genève, 51 Ch. des Maillettes, 1290 Versoix, Switzerland*
[2] *International Associate, Cosmic Dawn Center (DAWN) at the Niels Bohr Institute, University of Copenhagen and DTU-Space, Technical University of Denmark*
[3] *Centre for Astrophysics & Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3112, Australia*
[4] *Leiden Observatory, Leiden University, NL-2300 RA Leiden, Netherlands*
[5] *Departmento de Astronomia, Universidad de Chile, Casilla 36-D, Santiago 7591245, Chile*
[6] *Centro de Astrofisica y Tecnologias Afines (CATA), Camino del Observatorio 1515, Las Condes, Santiago 7591245, Chile*
[7] *Cavendish Laboratory, University of Cambridge, 19 JJ Thomson Avenue, Cambridge CB3 0HE, UK*
[8] *Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[9] *UCO/Lick Observatory, University of California, Santa Cruz, CA 95064, USA*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

The *James Webb Space Telescope* will allow to spectroscopically study an unprecedented number of galaxies deep into the reionization era, notably by detecting [O III]$\lambda\lambda$4959, 5007 and Hβ nebular emission lines. To efficiently prepare such observations, we photometrically select a large sample of galaxies at $z \sim 8$ and study their rest-frame optical emission lines. Combining data from the GOODS Re-ionization Era wide-Area Treasury from *Spitzer* (GREATS) survey and from *HST* we perform spectral energy distribution (SED) fitting, using synthetic SEDs from a large grid of photoionization models. The deep *Spitzer*/IRAC data combined with our models exploring a large parameter space enables to constrain the [O III]+Hβ fluxes and equivalent widths for our sample, as well as the average physical properties of $z \sim 8$ galaxies, such as the ionizing photon production efficiency with $\log(\xi_{\rm ion}/{\rm erg}^{-1}\ {\rm Hz}) \geq 25.77$. We find a relatively tight correlation between the [O III]+Hβ and UV luminosity, which we use to derive for the first time the [O III]$\lambda\lambda$4959, 5007+Hβ luminosity function (LF) at $z \sim 8$. The $z \sim 8$ [O III]+Hβ LF is higher at all luminosities compared to lower redshift, as opposed to the UV LF, due to an increase of the [O III]+Hβ luminosity at a given UV luminosity from $z \sim 3$ to $z \sim 8$. Finally, using the [O III]+Hβ LF, we make predictions for *JWST*/NIRSpec number counts of $z \sim 8$ galaxies. We find that the current wide-area extragalactic legacy fields are too shallow to use *JWST* at maximal efficiency for $z \sim 8$ spectroscopy even at 1hr depth and *JWST* pre-imaging to $\gtrsim 30$ mag will be required.

Key words: galaxies: evolution – galaxies: high-redshift – reionization

1 INTRODUCTION

Great progress has been made over the last two decades in our study of early galaxy mass assembly and the evolution of the cosmic star-formation rate density at $z \geq 4$ (e.g., Madau & Dickinson 2014; Duncan et al. 2014; Salmon et al. 2015). However, so far, these studies mostly rely on the analysis of broad-band photometry only, given that current facilities only provide access to the faint, rest-UV emission lines in a small number of bright galaxies (e.g., Stark et al. 2017). While a wealth of photometric data are now publicly available, photometric studies can suffer from several caveats. The selection of high-redshift galaxies relies on the Lyman break (i.e., dropout selection; Steidel et al. 1996) that can possibly miss a significant fraction of galaxies (e.g., Inami et al. 2017), and the derivation of most of the galaxy physical properties relies on spectral energy distribution (SED) fitting that

★ E-mail: stephane.debarros@unige.ch

is affected by several degeneracies and strongly depends on assumptions (e.g., star formation history, metallicity, dust extinction curve; Finlator et al. 2007; Yabe et al. 2009; De Barros et al. 2014). Furthermore, the photometry can be contaminated by strong nebular emission lines that are ubiquitous at high-redshift (e.g., Chary et al. 2005; Schaerer & De Barros 2010; Shim et al. 2011; Stark et al. 2013; Labbé et al. 2013; Smit et al. 2014; Shivaei et al. 2015; Faisst et al. 2016; Mármol-Queraltó et al. 2016; Rasappu et al. 2016). To account for their impact, either empirical (e.g., Schaerer & De Barros 2009) or dedicated photoionization modeling (e.g., Zackrisson et al. 2001) have been used. However, direct observational access to these emission lines through spectroscopy will have to await the advent of the *James Webb Space Telescope* (*JWST*).

At lower redshift, where most ultraviolet, optical, and near-infrared emission lines can be observed either from the ground or space, lines are efficiently used to determine instantaneous star-formation rates (SFR) and specific star-formation rates (sSFR=SFR/$M_\star$; Kauffmann et al. 2004), to derive gas-phase element abundances (e.g., Tremonti et al. 2004), to accurately derive the dust extinction thanks to the Balmer decrement (e.g., Domínguez et al. 2013; Reddy et al. 2015), and to determine the main source of ionizing photons (star formation or AGN) with the BPT diagram (Baldwin et al. 1981). While direct observations of optical and near-infrared lines are out of reach for $z > 4$ galaxies until the launch of *JWST*, one can take advantage of the impact of nebular emission lines on the photometry to probe these lines indirectly and to uniquely reveal some ISM properties of high-redshift galaxies.

Emission lines can also be useful to derive very accurate photometric redshifts. Smit et al. (2015) exploit extremely blue *Spitzer*/IRAC colors to identify $6.6 \leq z \leq 6.9$ galaxies, for which [O III]$\lambda\lambda$4959, 5007+Hβ lines are expected to fall in the 3.6μm band while the 4.5μm band is free of line contamination. A similar technique applied to $z > 7.1$ galaxies has led to the reliable selection and subsequent spectroscopic confirmation of some of the most distant Lyman-α emitters to date (Roberts-Borsani et al. 2016; Oesch et al. 2015; Zitrin et al. 2015; Stark et al. 2017). SFR and sSFR can be derived from Hα emission for galaxies at $z \sim 4$ where the Hα line is found in the IRAC 3.6μm channel while the IRAC 4.5μm channel is free from line contamination (Shim et al. 2011; Stark et al. 2013; Mármol-Queraltó et al. 2016). Furthermore, the Hα luminosities have also been used to derive the ionizing photon production efficiency ($\xi_{\rm ion}$, defined as the production rate of ionizing photons per unit luminosity in the UV-continuum; Bouwens et al. 2016) at $z \sim 4$. The comparison between uncorrected SFR derived from emission lines and SFR derived from UV+IR shed light on the relative stellar to nebular attenuation (Shivaei et al. 2015; De Barros et al. 2016).

In this paper, we adopt the same approach as these latter studies: we use the impact of nebular emission on the broad-band photometry to indirectly derive emission line fluxes and EWs to gain insight into high-redshift galaxy properties. Specifically, we use a sample of photometrically selected $z \sim 8$ galaxies, for which the [O III]$\lambda\lambda$4959, 5007+Hβ lines fall in the IRAC 4.5μm channel while we do not expect strong lines in the IRAC 3.6μm channel. We take advantage of the *Spitzer* ultra-deep survey covering CANDELS/GOODS South and North fields, the GOODS Reionization Era wide-Area Treasury from Spitzer (GREATS, Labbé et al. 2019, in prep) survey, providing the best constraints on $z \sim 8$ IRAC colors to date. To derive the line fluxes as accurately as possible and account for most of the uncertainties, we use synthetic SEDs produced with dedicated photoionization modeling to fit the observed $z \sim 8$ SEDs. Our aim is to derive the [O III]$\lambda\lambda$4959, 5007+Hβ luminosity function (LF) at $z \sim 8$ to prepare efficient *JWST* observations in the future.

The paper is structured as follows: The photometric data and selection procedure are described in Sec. 2. We provide a description of our photoionization grid in Sec. 3, and Sec. 4 gives the SED fitting method. In Sec. 5, we present the constraints that we obtain on physical properties of $z \sim 8$ galaxies. The resulting [O III]$\lambda\lambda$4959, 5007+Hβ luminosity function is shown and discussed in Sec. 6. We summarize our conclusions in Sec. 7.

We adopt a Λ-CDM cosmological model with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$. All magnitudes are expressed in the AB system (Oke & Gunn 1983).

2 DATA AND SAMPLE

The input sample used for this work is based on the Lyman break galaxy (LBG) catalogs from Bouwens et al. (2015). These are compiled from all the prime extragalactic legacy fields, including the Hubble Ultra Deep Field (HUDF; Ellis et al. 2013; Illingworth et al. 2013) and its parallel fields, as well as all five CANDELS fields (Grogin et al. 2011; Koekemoer et al. 2011).

In addition to deep *HST* near-infrared imaging, all these fields have extensive *Spitzer*/IRAC coverage. We have reduced and combined all the IRAC 3.6 and 4.5 μm data available in each field. In particular, we include the complete data from the GREATS survey (Labbe et al. 2019, in prep). GREATS builds on the vast amount of archival data in the two GOODS fields (Giavalisco et al. 2004) and brings the IRAC 3.6μm and 4.5μm coverage to a near-homogeneous depth of 200-250 hr, corresponding to 5σ sensitivities of 26.8-27.1 mag, over $\sim$ 200 arcmin^2. This is very well matched to the H_{160}-band detection limits, allowing us to detect the rest-frame optical light of nearly all the galaxies identified in the *HST* data. For a complete description of the *HST* and *Spitzer* dataset we refer the reader to Bouwens et al. (2015) and Stefanon et al. (2019, in prep).

Given the much wider point-spread function (PSF) of the IRAC data compared to *HST*, special care is required to derive accurate photometry. We use a custom-made software tool `mophongo`, developed and updated over the last few years (e.g., Labbé et al. 2010, 2015). In short, starting from the *HST* F160W image, `mophongo` uses position-dependent *HST*-to-IRAC PSF kernels to fit and subtract all the neighboring galaxies in a 21″ region around a source of interest, before measuring its flux density in a 2″ aperture.

After discarding objects with IRAC fluxes highly contaminated by neighbours (< 10%), we ran the EAZY code (Brammer et al. 2008) on the full photometric catalog, allowing the photometric redshift to vary from 0 to 10, and deriving the redshift probability distribution for each object in the GREATS data. We took care to

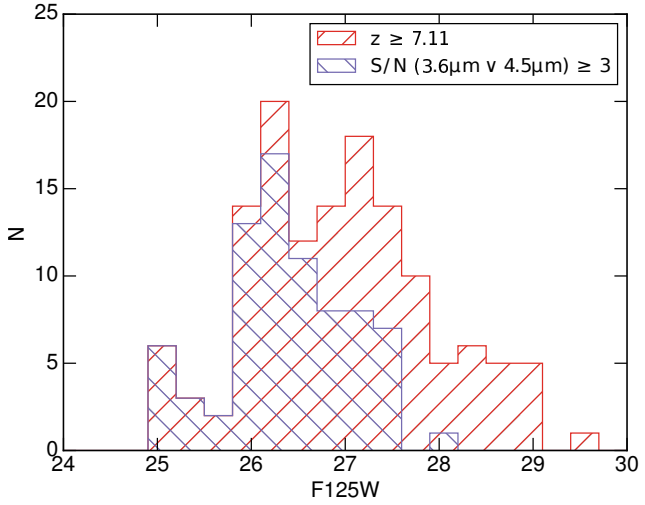

Figure 1. F125W magnitude distribution for the sample before (red) and after applying the 3.6μm and/or 4.5μm detection ($S/N > 3$) requirement (purple). This latter criterion favors UV bright galaxies with F125W< 27.5.

exclude IRAC 3.6μm and IRAC 4.5μm bands from the SED fit, since emission lines at high redshift can affect the photometry in these bands (e.g., Smit et al. 2014) and we want to avoid to be biased toward objects with large EW([O III]$\lambda\lambda$4959,5007+Hβ). We want to focus on $z \sim 8$ galaxies for which the IRAC 3.6μm−IRAC 4.5μm color provides constraints on the [O III]$\lambda\lambda$4959,5007+Hβ flux and these lines have completely entered the *Spitzer*/IRAC 4.5μm channel at $z \geq 7.11$ [1]. Therefore, the selection criterion was defined as $p(z \geq 7.11) \geq 0.68$, with $p(z \geq 7.11)$ the probability for a galaxy to have a redshift $z \geq 7.11$. Additionally to this criterion, we select a subsample of galaxies with at least one detection with $S/N \geq 3$ in either IRAC 3.6μm or IRAC 4.5μm allowing us to derive at least a strong upper or lower limit on the $(3.6-4.5)\mu$m color for each of these galaxies. In this subsample ($N = 76$), 16 galaxies are detected at 4.5μm only and 10 at 3.6μm only, potentially slightly biasing this subsample toward larger EW([O III]$\lambda\lambda$4959,5007+Hβ), with the detection at 4.5μm being consistent with an increase in flux due to [O III]+Hβ lines, while the stellar continuum could remain undetected at 3.6μm. We show in Fig. 1 the F125W magnitude distribution for the galaxy selected based on their photometric redshifts ($N = 135$) and the same distribution after applying the *Spitzer*/IRAC detection criterion. The F125W band probes the UV rest-frame emission of $z \sim 8$ galaxies (~ 1500Å). $\sim 50\%$ of the $z \sim 8$ galaxies are detected with $S/N \geq 3$ in at least one *Spitzer*/IRAC band, mostly the brightest with F125W< 27.5. While we apply a threshold detection in IRAC to select our subsample, we use all available photometry including bands with low S/N (<3) as well as non detections to perform the SED fitting. We use the subsample ($N = 76$) with $S/N(3.6\mu\text{m} \vee 4.5\mu\text{m}) \geq 3$ to constrain the $z \sim 8$ galaxy physical properties (Sec. 5) and the entire photometric sample ($N = 135$) to derive the [O III]+Hβ LF (Sec. 6).

[1] The [O III]+Hβ lines are out of IRAC 4.5μm at $z > 9.05$ and after applying our selection, only one galaxy has a redshift above this limit.

3 PHOTOIONIZATION MODELS

Several works have used photoionization models to study or predict nebular emission line properties of high-redshift galaxies (e.g., Zackrisson et al. 2011; Jaskot & Ravindranath 2016; Steidel et al. 2016; Nakajima et al. 2018; Berg et al. 2018). Since we do not have access with the current facilities to the optical/NIR nebular emission lines for galaxies at $z \sim 8$, the ISM physical conditions at these high redshifts are largely unknown. Therefore we created a grid of photoionization models with a large parameter space to encompass the plausible stellar and ISM physical properties at $z \sim 8$

We used the latest release of the Cloudy photoionization code (C17, Ferland et al. 2017) to build our grid of models. We chose SEDs from the latest BPASSv2.1 models (Eldridge et al. 2017) as input, which account for stellar binaries and stellar rotation effects. Indeed, several recent observations point out that high-redshift galaxies can exhibit UV emission lines requiring hard ionizing photons (e.g., C III]$\lambda\lambda$1907,1909, C IV$\lambda\lambda$1548,1550; Stark et al. 2014; Amorín et al. 2017; Vanzella et al. 2017), and there are mounting evidences that these UV lines are due to star formation since they are spatially associated with star forming regions (Smit et al. 2017). These kind of strong emission lines can be reproduced by including stellar rotation and/or stellar binaries in the stellar models (e.g., Eldridge et al. 2008) or using stellar templates updated with recent UV spectral libraries and stellar evolutionary tracks as in the latest Charlot & Bruzual single star models (e.g., Gutkin et al. 2016). Regarding the ISM properties, we adopt the same interstellar abundances and depletion factors of metals on to dust grains, and dust properties as Gutkin et al. (2016). These authors show that these modeling assumptions span a range that can reproduce most of the observed UV and optical emission lines at low- and high-redshift (Stark et al. 2014, 2015a,b, 2017; Chevallard & Charlot 2016). While Gutkin et al. (2016) use different stellar population synthesis model than used here, namely an updated version of the Bruzual & Charlot (2003) stellar population synthesis model, a comparison of these two SPS models show that they provide similar results in interpreting stellar and nebular emissions of local massive star-clusters (Wofford et al. 2016). For the grid used in this work, we use stellar metallicities from $Z = 0.001$ to $Z = 0.008$ with an initial mass function (IMF) index of −2.35 and an upper mass cutoff of 300M$_\odot$. For each stellar metallicity, for simplicity, we assume the same gas-phase metallicity. We also explore a range of C/O abundance ratio (from $\log \text{C/O} = -1.0$ to -0.4, consistent with the observations of Amorín et al. 2017), three different values of dust-to-metal ratios ($\xi_d = 0.1, 0.3, 0.5$; Gutkin et al. 2016), and a range of hydrogen gas densities (10^2 to 10^3 cm^{-3}). We assume no leakage of ionizing photons. For each set of parameters and each stellar age, we built SEDs, adding to the pure stellar SEDs from BPASS the nebular emission lines and nebular continuum as computed in Cloudy.

4 SED FITTING

We use the spectral energy distributions created with CLOUDY to perform SED fitting of each individual galaxy in our $z \sim 8$ sample. We use a modified version of the SED fitting code *Hyperz* (Bolzonella et al. 2000) allowing us to apply two different attenuation curves to the stellar and nebular components of the SED. We apply a Calzetti attenuation curve (Calzetti et al. 2000) to the stellar component and a Cardelli attenuation curve (Cardelli et al. 1989) to the nebular component, adopting $E(B-V)_{\rm gas} = E(B-V)_{\star}$ for simplicity. Studies of galaxy samples at $z \sim 2$ show that the ratio $E(B-V)_{\rm gas}/E(B-V)_{\star}$ is affected by a large scatter and is increasing with increasing $E(B-V)_{\star}$ and SFR (Reddy et al. 2015; Theios et al. 2019). Since $z \sim 8$ galaxies exhibit blue UV β slopes indicating low dust extinction (e.g., Bouwens et al. 2014), we do not expect dust to have a large impact on our results. Nevertheless we allow $E(B-V)$ to vary from 0.0 to 0.2 in our SED fitting procedure. We assume a constant star formation history. While the choice of the SFH has an impact on the derived physical parameters, this effect is alleviated for young ages (< 100Myr; De Barros et al. 2014). Since the oldest age allowed by the cosmological model adopted in this work is 730 Myr and due to the relatively young best-fit ages found for our sample (Sec. 5), the assumed SFH has a limited impact on our final results.

Minimization of χ^2 over the entire parameter space yields the best-fit SED. For each physical parameter of interest, we derive the median of the marginalized likelihood, and its associated uncertainties. Derived physical parameters include age of the stellar population, stellar mass, SFR, sSFR, as well as observed (i.e., attenuated by dust) [O III]$\lambda\lambda$4959,5007 and Hβ emission line fluxes and EWs, and ISM physical properties (e.g., ionization parameter). All physical properties used in this work such as EW([O III]+Hβ) and L([O III]+Hβ) are SED derived, and correspond to the median of the marginalized likelihood, except stated otherwise.

The median uncertainty regarding the EW([O III]$\lambda\lambda$4959,5007+Hβ) for the entire sample is +0.34/−0.37 dex. We also split the sample in four bins of F125W magnitude to emphasize the reliability of the constraints depending on the UV luminosity. For F125W < 26, 26 < F125W < 26.5, 26.5 < F125W < 27, and F125W > 27, the median uncertainties are +0.27/−0.32 dex, +0.30/−0.32 dex, +0.38/−0.45 dex, and +0.39/−0.52 dex, respectively. We describe how we account for those relatively large uncertainties in the [O III]+Hβ luminosity function (LF) derivation in Sec. 6.1.

5 CONSTRAINTS ON THE ISM AND PHYSICAL PROPERTIES

5.1 Predictions for the (3.6-4.5)μm Color

We show in the top left panel of Fig. 2 the range of (3.6-4.5)μm color which is spanned by our grid of models for $Z = 0.004$. At an age of 1 Myr, (3.6-4.5)μm can vary by 2 magnitudes, nebular emission (mainly [O III]$\lambda\lambda$4959,5007+Hβ lines) boosting the flux in the IRAC 4.5μm channel and producing IRAC 3.6μm−IRAC 4.5μm color redder by $\sim$ 1.5 magnitude in comparison with the color expected from a pure stellar template. However, nebular emission can also have the opposite effect for low ionization parameters, producing a bluer color than expected for pure stellar emission up to $\sim$ 0.5 magnitude. This effect is due to the relation between the ionization parameter and the [O III]/[O II]λ3727 ratio: log([O III]/[O II]λ3727) is increasing with higher ionization parameter (depending on the metallicity, Kewley & Dopita 2002), becoming larger than 1 at $Z = 0.05Z_{\odot}$ for $\log \mathrm{U} \gtrsim -3.0$. However, the number of galaxies for which the IRAC 3.6μm−IRAC 4.5μm color is best fitted with a very low ionization parameter ($\log \mathrm{U} = -4.0$) is small (25%). Furthermore, photometric redshift uncertainties can also account for $z > 7.1$ blue colors with the Balmer jump starting to enter the IRAC 3.6μm channel at $z > 8$.

We show the EW([O III]$\lambda\lambda$4959,5007+Hβ) evolution with age in Fig. 2 (bottom left panel) and examples of SEDs for our final sample for a range of F125W magnitude in Fig. 3.

5.2 Stellar Metallicity and ISM Physical Properties

The models required to reproduce the SEDs, mainly the IRAC colors, give insight in the ISM physical properties at $z \sim 8$. Some parameters, like the C/O ratio or the hydrogen gas density, have little to no impact on the EW([O III]$\lambda\lambda$4959,5007+Hβ) which is the spectral feature with the largest effect on the IRAC colors, and therefore providing the main constraints on the ISM physical conditions. In our grid of models, the (3.6-4.5)μm is mostly defined by the stellar metallicity, the ionization parameter, and the age of the stellar population. For our sample, we find a median stellar metallicity of $Z_{\star} = 0.004^{+0.004}_{-0.002}$, a median ionization parameter $\log \mathrm{U} = -3.0 \pm 1.0$, and a median $\log(\mathrm{age/yr}) = 7.2^{+0.9}_{-0.6}$. The constraints on dust extinction are mostly coming from the fit of the UV β slope and we find a median extinction $A_V = 0.4 \pm 0.2$.

As noted previously, we do not assume any ISM properties of $z \sim 8$ galaxies but compare their photometry with a grid of photoionization models that we consider to encompass the plausible $z \sim 8$ properties. The constraints on the stellar metallicity and the ionization parameter are driven by the number of ionizing photons available to interact with the gas and the gas-phase Oxygen abundance. The ionizing photon flux is increasing with decreasing metallicity (e.g., Stanway et al. 2016) and with increasing ionization parameter, but the gas-phase Oxygen abundance is decreasing with decreasing metallicity, and also decreasing with increasing dust-to-metal ratio (more Oxygen depletes into the ISM dust-phase; Gutkin et al. 2016). Therefore the ISM parameter derivation suffers from several degeneracies: to reproduce (3.6-4.5)μm colors produced by strong [O III]$\lambda\lambda$4959,5007+Hβ emission lines ((3.6 − 4.5)μm > 0, Fig. 2) a large range of metallicities and ionization parameters is allowed, as long as there is the right balance between the number of ionizing photons and the Oxygen abundance. However, the parameter space allowing this balance is smaller for low metallicities due to large ionizing photon production but low Oxygen abundance. The same is true for high metallicities due to lower ionizing photon production and high Oxygen abundance. Then the median metallicity

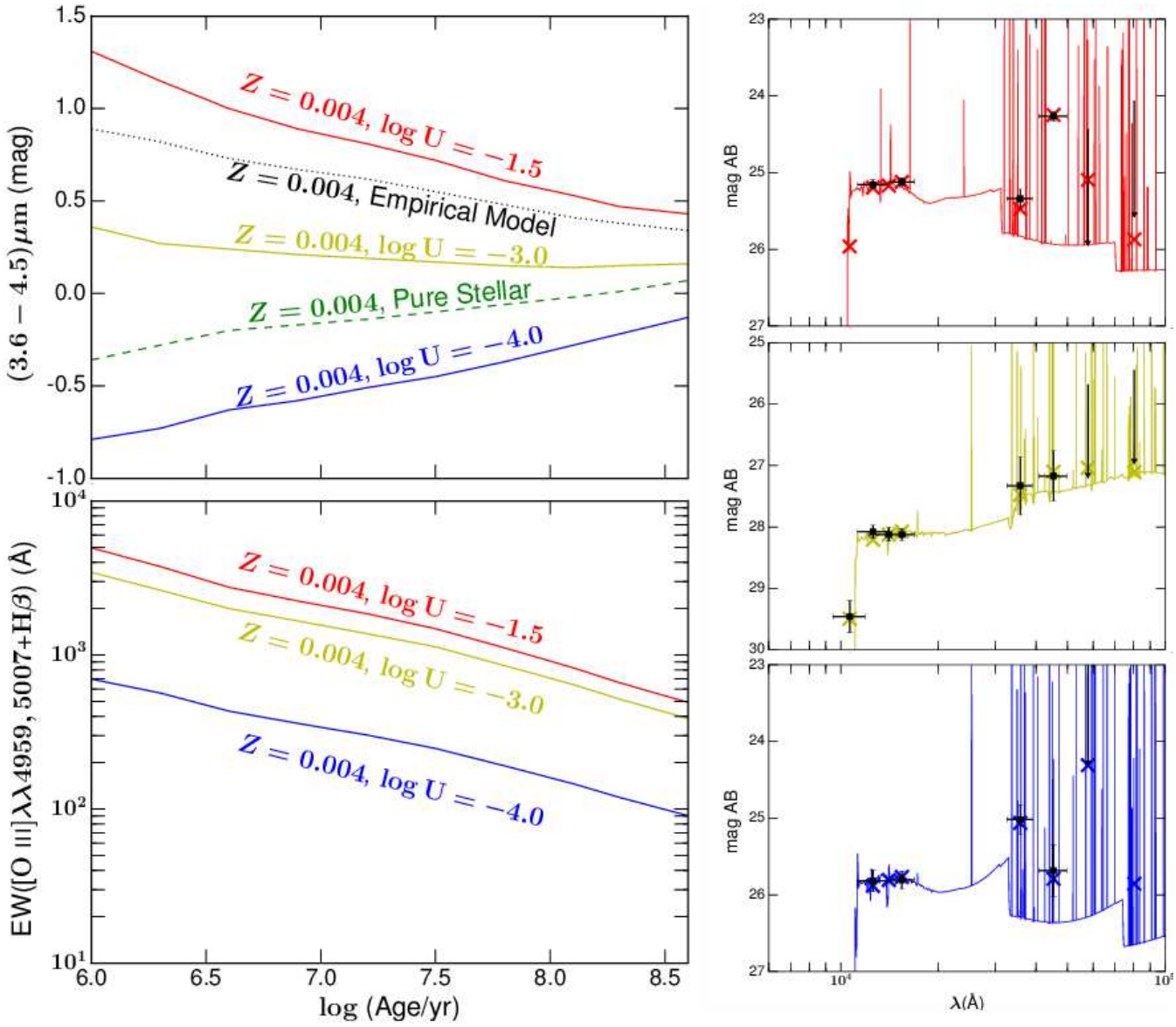

Figure 2. Top and bottom left panels: Range of IRAC 3.6μm–IRAC 4.5μm colors (at $z = 7.5$) and EW([O III]$\lambda\lambda$4959, 5007+Hβ) vs. age probed by our grid of photoionization models for $Z = 0.004$ and $\log U = -1.5$, -3.0, and -4.0 in red, yellow, and blue, respectively. The hydrogen densities n(H) of the models shown lie between 10^2 and 10^3 (the impact of a n(H) variation in this range is small) and we do not specify the carbon to oxygen abundance log(C/O) since this parameter also has no impact on the quantities shown. Also shown are IRAC colors for a pure stellar BPASSv2.1 template (dashed green line) and a template using typical empirical modeling of nebular emission (continuum and lines; Schaerer & De Barros 2009, 2010). **Right panels:** Examples of three best-fit SEDs obtained with the models shown on the left panels (same colors) for $z \sim 7.5-8$ galaxies from our broader LBG sample. Clearly, the IRAC colors are heavily affected by the vast amount of emission lines.

found for our sample only reflects that for $Z = 0.004$ there is a larger parameter space in terms of ionization parameter, age, and dust-to-metal ratio allowing to reproduce the observed (3.6-4.5)μm colors. Indeed, we found that $\sim 40\%$ of our sample has a best-fit SED with $Z \leq 0.002$, while the median of the marginalized likelihood for the metallicity is $Z \geq 0.004$ for the entire sample.

The choice of an IMF upper mass cutoff at 300M$_\odot$ has a negligible impact on our results since a cutoff of 100M$_\odot$ changes the [O III]$\lambda\lambda$4959, 5007 flux by $\sim 10\%$ (1-2% for Hβ and [O II]) which is small compared to the typical uncertainties affecting EWs and line luminosities (Sec. 4). Changing the assumed dust attenuation curve from a Calzetti to an SMC curve (Prevot et al. 1984; Bouchet et al. 1985) has also little to no impact on the overall derived properties, except for dust attenuation.

The $z \sim 8$ subsample with $S/N(3.6\mu\text{m} \vee 4.5\mu\text{m}) \geq 3$ has a median stellar mass $\log(M_\star/M_\odot) = 8.62^{+0.43}_{-0.39}$ and a median SFR $\log(\text{SFR}/M_\odot\text{yr}^{-1}) = 1.26^{+0.42}_{-0.30}$.

5.3 Ionizing Photon Production Efficiency

We are able to derive the ionizing photon production efficiency from SED fitting by computing for each template

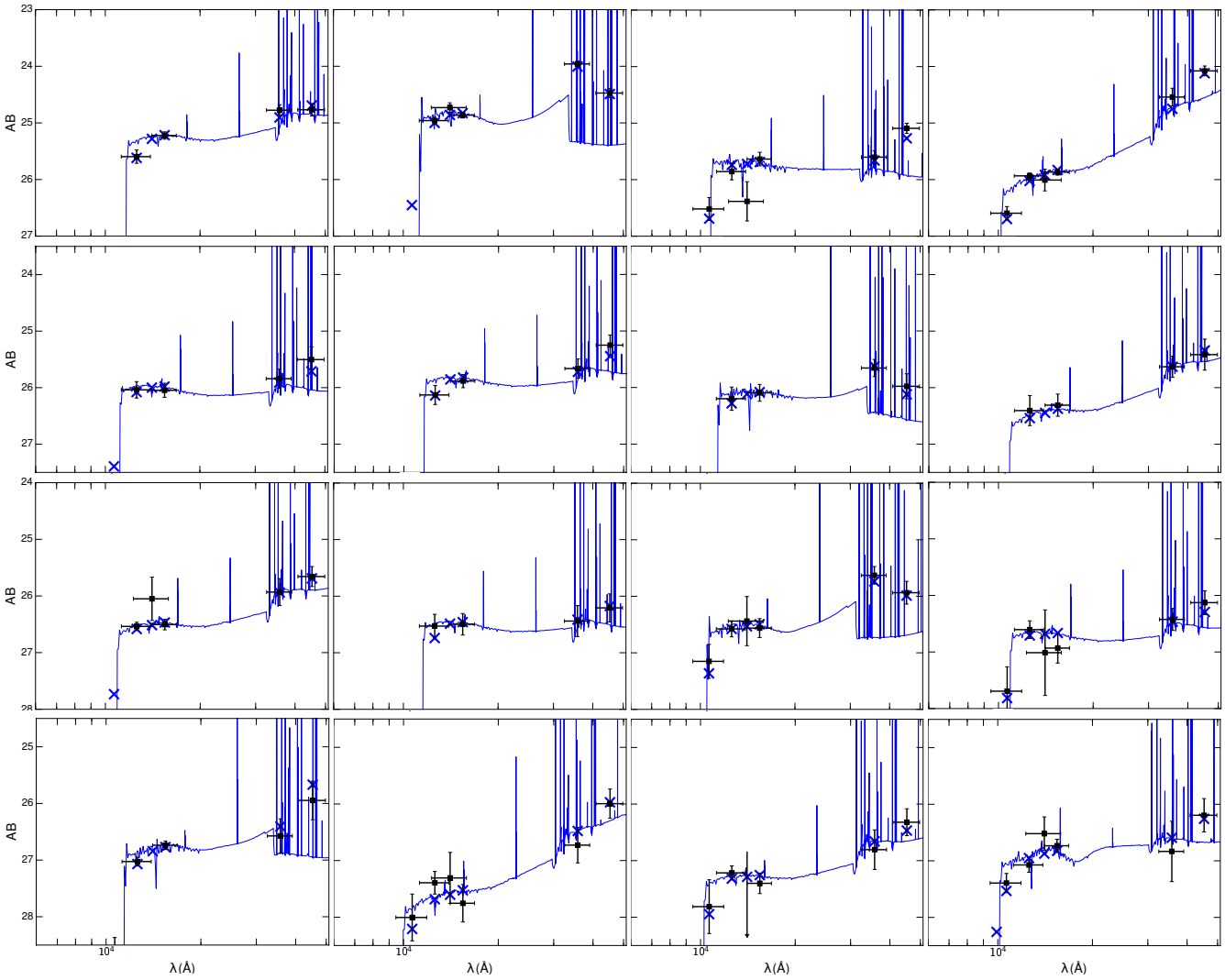

Figure 3. Examples of SEDs for our final sample. Each row shows 4 SEDs randomly selected in 4 bins of F125W magnitude defined from top to bottom as F125W < 26, 26 < F125W < 26.5, 26.5 < F125W < 27, and F125W > 27. The errorbars of the observed wavelength indicate the width of the filter transmission curve. Upper limits in flux indicate 1σ limits. Blue crosses show the synthesised flux in the filters. Our grid of models is able to reproduce the large range of observed (3.6-4.5)μm colors.

used in this work the Lyman continuum photon production rate $N_{\rm LyC}$ and the observed monochromatic UV luminosity L_ν. The ionizing photon production is then $\xi_{\rm ion} = N_{\rm LyC}/L_\nu$. For our final sample, we find $\log(\xi_{\rm ion}/{\rm erg}^{-1}\ {\rm Hz}) = 26.29^{+0.40}_{-0.38}$ assuming a Calzetti dust attenuation curve and $\log(\xi_{\rm ion}/{\rm erg}^{-1}\ {\rm Hz}) = 26.07^{+0.27}_{-0.30}$ assuming an SMC dust attenuation curve. In most lower redshift studies (e.g., Shivaei et al. 2018), $\xi_{\rm ion}$ is inferred from a dust corrected Hydrogen line (e.g., Hα) for which the flux depends mostly on the Lyman continuum photon production rate (e.g., Storey & Hummer 1995). Given the large number of unconstrained parameters going into our analysis (e.g., attenuation curve), we consider that our result set a lower limit to the average $z \sim 8$ ionizing photon production efficiency with $\log(\xi_{\rm ion}/{\rm erg}^{-1}\ {\rm Hz}) \geq 25.77$.

We note that in some studies (e.g., Bouwens et al. 2016), $\xi_{\rm ion}$ is an *intrinsic* quantity since it is derived by using a dust-corrected UV luminosity, while in our work we derive an *observed* $\xi_{\rm ion}$ value since we do not correct the observed UV luminosity for dust. Due to this dust correction, the intrinsic $\xi_{\rm ion}$ sets a lower limit for the observed $\xi_{\rm ion}$. However, thanks to the small dust attenuation that we find for our sample, the difference between intrinsic and observed $\xi_{\rm ion}$ should be small.

The constraints that we obtain on the ionizing photon production efficiency at $z \sim 8$ are consistent with results obtained for the bluest (i.e., least dust attenuated) LBGs and Lyman-α Emitters at $z \sim 2$ (Shivaei et al. 2018; Sobral et al. 2018; Tang et al. 2018) as well as results obtained for low-redshift compact star-forming galaxies (Izotov et al. 2017), and for LBGs at $z \sim 4-5$ (Bouwens et al. 2016; Lam et al. 2019; Ceverino et al. 2019). The observed ionizing photon production efficiency that we find is also consistent with the observed value found for $z \sim 0.3$ Lyman continuum emitters ($\xi_{\rm ion} = 25.6 - 26$, Schaerer et al. 2016).

This is the first time that the ionizing photon production efficiency is estimated for a significant sample of galaxies in the reionization era and this value is higher than the

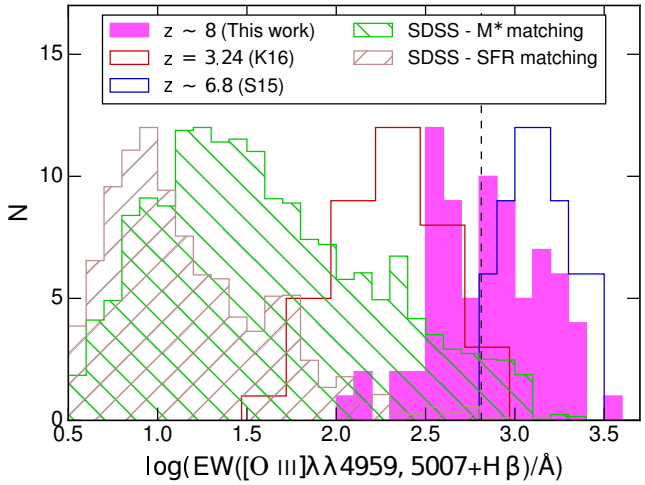

Figure 4. EW([O III]$\lambda\lambda$4959, 5007+Hβ) distribution for our sample. We also show the EW distribution at $z = 3.24$ (Khostovan et al. 2016, K16), $z \sim 6.8$ (Smit et al. 2015, S15), and the distribution drawn from the SDSS sample by matching the stellar mass and SFR distributions of our $z \sim 8$ sample. All distributions have been renormalized to have the same maximum as the $z \sim 8$ distribution. The vertical dashed line shows the median value for our sample. Clearly, the average equivalent width of our $z \sim 8$ sample is significantly higher than the $z \sim 0$ or $z \sim 3$ samples.

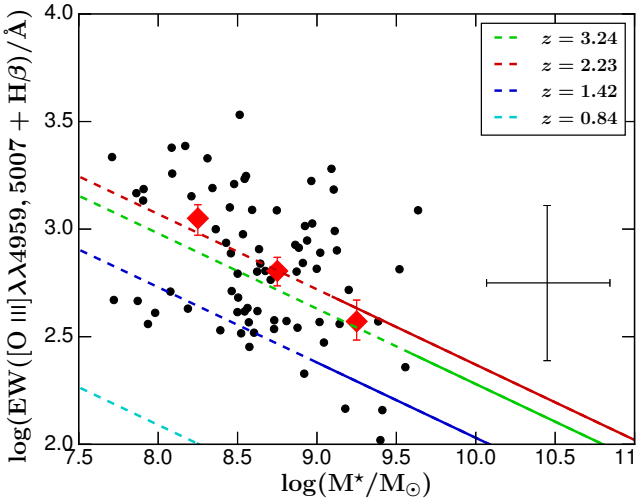

Figure 5. EW vs. $M_\star$ for our $z \sim 8$ sample. The large red diamonds show the median EW in bins of $M_\star$. We also show the power law relations EW $\propto M_\star^\beta$ derived for different redshifts in Khostovan et al. (2016) down to the minimum stellar mass used to derive them (solid lines) and their extrapolations to lower stellar masses (dashed lines). The median error bar for individual objects is shown on the right side of the figure. Our $z \sim 8$ sample is broadly consistent with the $z \sim 2-3$ relation.

canonical value $\log(\xi_{\rm ion}/{\rm erg}^{-1}\ {\rm Hz}) = 25.2 - 25.3$ by a factor ≥ 3. This higher value of $\xi_{\rm ion}$ translates into a lower value of the Lyman continuum escape fraction required in a scenario where star-forming galaxies are driving cosmic reionization (Bouwens et al. 2016; Shivaei et al. 2018; Chevallard et al. 2018b; Matthee et al. 2017b,a; Lam et al. 2019).

5.4 Evolution of the EW(Hβ+[O III]$\lambda\lambda$4959, 5007) with Redshift

We show the EW([O III]$\lambda\lambda$4959, 5007+Hβ) distribution for our $z \sim 8$ sample in Fig. 4 along with the EW distribution at $z = 3.24$ (Khostovan et al. 2016) and the one at $z \sim 6.8$ for the extreme emitter sample of Smit et al. (2015). Our distribution is consistent with the latter, given that the sample of Smit et al. (2015) only included sources with the largest EW. We find a median EW([O III]$\lambda\lambda$4959, 5007 + Hβ) = 649^{+92}_{-49}Å, consistent with the value of 670^{+260}_{-170}Å from Labbé et al. (2013).

Comparing our EW distribution with the Sloan Digital Sky Survey (SDSS, Abolfathi et al. 2018), we find that only 0.23 ± 0.01% galaxies exhibit such strong emission lines (EW([O III]$\lambda\lambda$4959, 5007 + Hβ) $\geq$ 300Å) in the entire SDSS sample. We also compare the $z \sim 8$ distribution with two distributions drawn from SDSS. The SDSS samples were mass- or SFR-matched by randomly picking 50 SDSS galaxies within 0.05 dex and 0.1 dex in terms of stellar mass and SFR, respectively, for each galaxy in our sample. The SDSS sample selected through SFR-matching exhibits only a small overlap with the EWs derived in our work (1.6 ± 0.2%), while the one matched by stellar mass leads to a non-negligible fraction of galaxies with EW as high as found in our $z \sim 8$ sample (9.8 ± 0.4%). Nevertheless, it is clear that the emission lines of $z \sim 8$ galaxies are much more extreme than local galaxies of similar mass or SFR, although specific local and low-z population such as Blue Compact Dwarf galaxies exhibit similar properties (e.g., Izotov et al. 2011; Cardamone et al. 2009; Yang et al. 2017; Rigby et al. 2015; Senchyna et al. 2017).

We have also attempted to match our sample in terms of sSFR but the number of SDSS galaxies to exhibit similarly high sSFR as in our sample (median sSFR = $63^{+188}_{-55}{\rm Gyr}^{-1}$) is extremely small ($< 0.35$%), such that no representative sSFR-matched sample could be constructed. However, the small sample of galaxies with such high sSFR does indeed exhibit EWs as large as derived in our work. This illustrates that finding a significant sample of local galaxies with properties (stellar mass, SFR, sSFR, and EW([O III]$\lambda\lambda$4959, 5007+Hβ) similar to $z \sim 8$ galaxy properties is a difficult task.

From $z = 3.24$ to $z \sim 8$, there is a clear evolution of the median of the EW distribution. However, the EWs have to be compared for a given stellar mass range (e.g., $9.5 < \log(M_\star/M_\odot) < 10.0$, Khostovan et al. 2016). The derived stellar mass for our sample is significantly lower than the $z \sim 3.24$ sample, with only four galaxies (5%) with $\log(M_\star/M_\odot) \geq 9.5$. A comparison of the EW properties as a function of stellar mass is shown in Fig. 5. While individual error bars are relatively large ($\sim \pm 0.4$ dex for both parameters), our sample is consistent with the $z = 3.24$ and 2.23 EW–$M_\star$ relations. Furthermore, using the median stellar mass of our sample with the $z = 2.23$ relation from Khostovan et al. (2016), we predict a median equivalent width EW = 712^{+70}_{-62}Å for our sample, a value consistent with our derivation based on SED fitting.

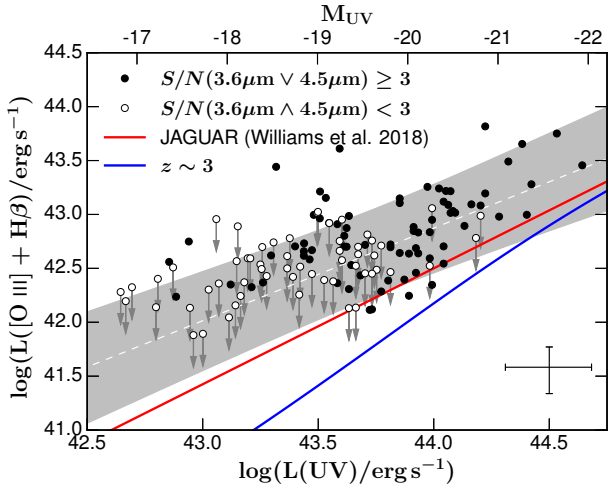

Figure 6. UV luminosity vs. [O III]$\lambda\lambda$4959, 5007+Hβ luminosity for our sample. The dashed line is the derived relation and the grey area shows the 68% confidence interval. The typical errorbar for individual galaxies with $S/N(3.6\mu\text{m} \vee 4.5\mu\text{m}) \geq 3$ is shown on the bottom right corner. For galaxies with $S/N(3.6\mu\text{m} \wedge 4.5\mu\text{m}) < 3$, we show the 90% upper limits. We show in red and blue the same relation derived from the JAGUAR mock catalog (Williams et al. 2018, see Sec. 6.1) and the relation at $z \sim 3$ (as derived through abundance matching; see Sec. 6.2), respectively. The line luminosities increase from $z \sim 3$ to $z \sim 8$ (at a given UV luminosity).

6 THE Hβ+[O III] LUMINOSITY FUNCTION AT $Z \sim 8$

Based on the photometric estimates of the [O III]$\lambda\lambda$4959, 5007+Hβ emission line strengths of all the $z \sim 8$ galaxies in the GREATS sample, we have the opportunity for a first derivation of the [O III]$\lambda\lambda$4959, 5007+Hβ luminosity function at these redshifts, which we describe in the next section.

6.1 Derivation of the Emission Line Luminosity Function

Our approach is based on converting the UV LF to an emission line LF using the relation between the observed UV luminosity, L_{UV}, and the [O III]$\lambda\lambda$4959, 5007+Hβ line luminosity, $L_{OIII+H\beta}$, at $z \sim 8$. This approach is analogous to the one used in the derivation of stellar mass functions at very high redshift (e.g., in González et al. 2010; Song et al. 2016) or the star-formation rate function (Smit et al. 2012, 2016; Mashian et al. 2015).

The relation between the [O III]$\lambda\lambda$4959, 5007+Hβ luminosity and the observed UV luminosity in our sample is calibrated in Fig. 6. To increase the range of UV luminosities probed in our work, we add to our sample with $S/N(3.6\mu\text{m} \vee 4.5\mu\text{m}) \geq 3$ galaxies with lower S/N. We apply to these galaxies the same procedure as the rest of the sample and so we obtain the complete probability distribution function for all the parameters, including the [O III]+Hβ luminosity. While uncertainties remain relatively large for individual galaxies (~ 0.18 dex and $0.19 - 0.24$ dex for the UV and [O III]$\lambda\lambda$4959, 5007+Hβ luminosities, respectively), we find that the UV luminosity and the [O III]+Hβ luminosity are well correlated (Spearman rank correlation coefficient $\rho = 0.56$, standard deviation from null hypothesis $\sigma = 7.1$).

We use a MCMC method to fit the relation with three parameters, a slope and an intercept, plus an intrinsic (Gaussian) dispersion around the relation, σ_{int}. This results in:

$$\log(\text{L}([\text{O III}]\lambda\lambda4959, 5007 + \text{H}\beta)/\text{erg s}^{-1}) = 0.86 \pm 0.12 \times \log(\text{L}_{\text{UV}}/\text{erg s}^{-1}) + 33.92^{+1.23}_{-1.27} \quad (1)$$

Together, with an intrinsic dispersion of $\sigma_{int} = 0.35$ dex around the median fit. The corresponding 68-percentile contours are also shown in Fig. 6.

As a comparison, we also use the publicly available mock catalog JAdes extraGalactic Ultradeep Artificial Realizations (JAGUAR, Williams et al. 2018) to derive the relation between UV and [O III]+Hβ luminosity of simulated $z \sim 8$ galaxies. The JAGUAR mock catalog has been produced by matching luminosity and stellar mass functions as well as the relation between the stellar mass and UV luminosity, mostly at $z \leq 4$. The galaxy properties are then extrapolated up to $z \sim 15$. The JAGUAR catalog provides emission line fluxes and EWs for the main lines based on modeling with the BEAGLE code (Chevallard & Charlot 2016; Chevallard et al. 2018a). We identify all galaxies from the fiducial JAGUAR mock in the redshift range $7.11 < z < 9.05$ and we randomly select 1000 of them to match the absolute UV magnitude distribution of our sample, and then fit the UV-[O III]+Hβ luminosity data. The result is shown in red in Fig. 6. Similarly to our sample, the $z \sim 8$ galaxies from the JAGUAR catalog exhibit a tight relation between UV and [O III]+Hβ luminosity (Spearman rank correlation coefficient $\rho = 0.73$, standard deviation from null hypothesis $\sigma > 40$). However, the mock galaxies exhibit a significantly lower [O III]+Hβ luminosity (~ 0.5 dex) at a given L_{UV} compared to the relation of our galaxies. The detailed reason for this discrepancy relative to the JAGUAR mock is unclear, but one possible reason is differences in the median physical properties. For instance, while the mock galaxies exhibit (3.6-4.5)μm color similar to the ones from our sample at a given UV luminosity, the average F125W-3.6μm color in JAGUAR is smaller by ~ 0.3 magnitude compared to the observed F125W-3.6μm color in our sample. This means that while (3.6-4.5)μm color and EW([O III]+Hβ) are on average similar between JAGUAR and our sample, the absolute [O III]+Hβ line luminosity scales with the 3.6μm flux which is larger in our sample compared to the JAGUAR mock catalog. Furthermore, JAGUAR models a small field comparatively to our data, therefore the overlap in UV luminosity is small.

Using the relation between the UV and [O III]+Hβ luminosity of Eq. 1, we can now derive the [O III]+Hβ LF based on the known $z \sim 8$ UV LF (Bouwens et al. 2015). In order to properly compute errorbars, we adopt a Markov Chain Monte Carlo approach (see Sharma 2017, for a review). In particular, we sample 100'000 points from the UV LF and convert their corresponding L_{UV} values to $L_{OIII+H\beta}$, based on the relation derived above including the appropriate dispersion σ_{int}. Finally, we fit a Schechter function to the resulting $L_{OIII+H\beta}$ values, keeping the three quantities $\Phi^\star$, $\text{L}^\star$, and α as free parameters. We repeat this procedure 10'000 times, and vary the input Schechter function parameters of the UV LF according to their appropriate covariance matrix, which was derived from the contour plots shown in

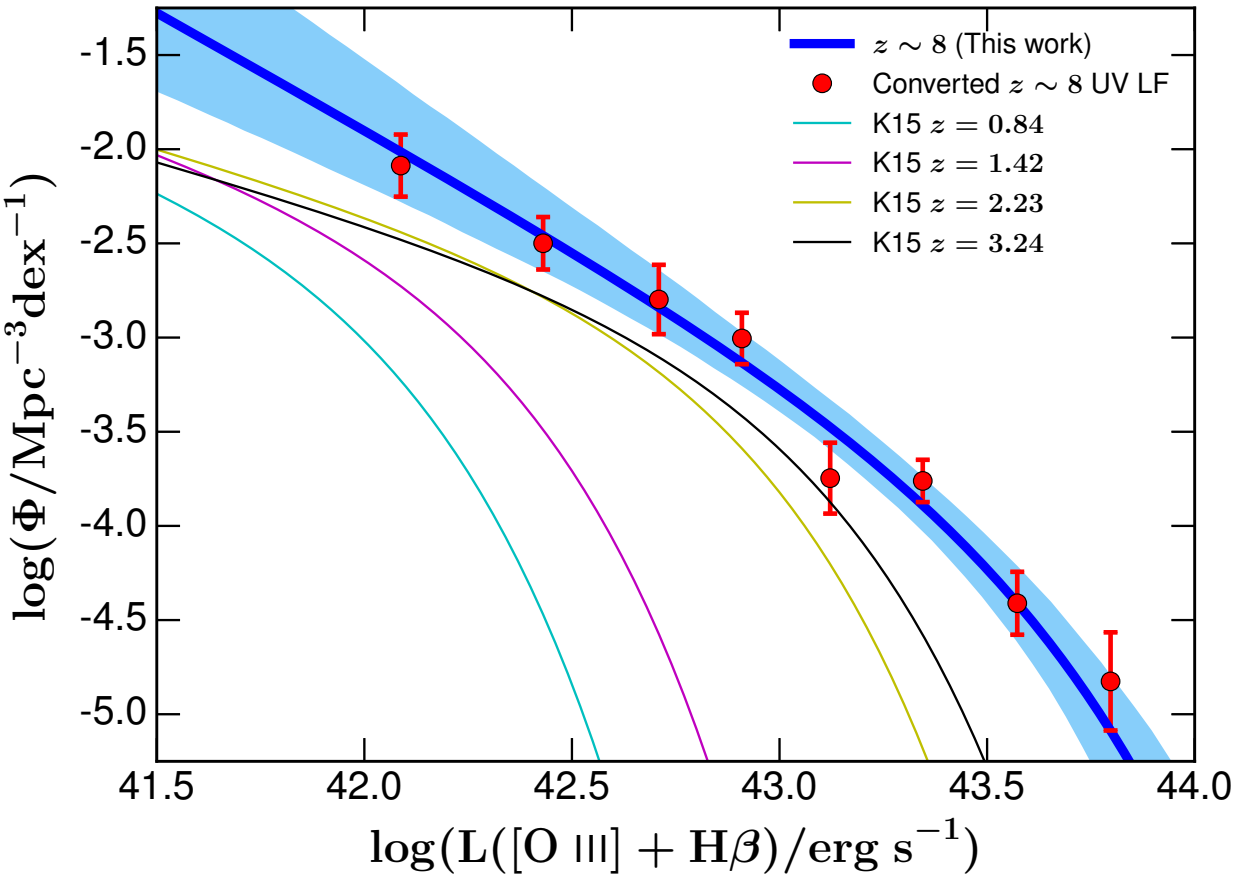

Figure 7. [O III]$\lambda\lambda$4959, 5007+Hβ luminosity function derived in this work for our sample at $z \sim 8$ (blue thick line) with 68% confidence interval (light blue area). We also show the corresponding $z \sim 8$ UV luminosity function converted with the relation from Fig. 6. For comparison, we show the [O III]$\lambda\lambda$4959, 5007+Hβ luminosity functions measured in Khostovan et al. (2015, K15) from $z = 0.84$ to $z = 3.24$. Continuing the trend from lower redshift, the line LF is higher at $z \sim 8$ than at $z \sim 3$ (unlike the UV LF).

Bouwens et al. (2015). This procedure results in 10'000 line LFs, from which we compute the mean and standard deviation for all three Schechter function parameters of the line LF.

The final result is shown in Fig. 7, with Schechter function parameters of the line LF of $\log(\mathrm{L}^\star/\mathrm{erg\ s}^{-1}) = 43.45^{+0.21}_{-0.19}$, $\log(\Phi^\star/\mathrm{Mpc}^{-3}) = -4.05^{+0.40}_{-0.44}$, and $\alpha = -2.22^{+0.28}_{-0.32}$. The points with errorbars in Fig. 7 correspond to the observed UV LF measurements from Bouwens et al. (2015), which were converted to the [O III]$\lambda\lambda$4959, 5007+Hβ LF using the same approach as described above. They clearly agree very well with the mean Schechter function derivation.

6.2 The Evolution of the $[OIII] + H\beta$ Line Luminosity Function to $z \sim 8$

The [O III]$\lambda\lambda$4959, 5007+Hβ LF has previously been derived up to $z \sim 3$ by several authors (Hippelein et al. 2003; Ly et al. 2007; Colbert et al. 2013; Pirzkal et al. 2013; Khostovan et al. 2015). By adding our estimate at $z \sim 8$, we can thus study its evolution across more than 13 Gyr. Fig. 7 shows such a comparison of the $L_{OIII+H\beta}$ LFs derived at different redshift, from $z \sim 0.5$ to $z \sim 8$. Interestingly, the line LF is found to be higher at all luminosities at $z \sim 8$ compared to $z \sim 3$. This is in stark contrast to the evolution of the UV LF, which peaks at $z \sim 2-3$, but then steadily declines to higher (or lower) redshift. We show in Fig. 8 the evolution with redshift of $\mathrm{L}^\star$ and $\Phi^\star$ and by extrapolating the observed trends at lower redshift to $z \sim 8$, especially the Khostovan et al. (2015) results, the $z \sim 8$ values are in remarkable agreement with expectations.

The evolution of the [O III]+Hβ LF from $z \sim 3$ to $z \sim 8$ can be explained by an evolution of the relation between L(UV) and the stellar mass with redshift. It is known that high-redshift galaxies have their stellar mass related to their UV luminosity (Stark et al. 2009; González et al. 2011; Duncan et al. 2014; Grazian et al. 2015; Song et al. 2016; Stefanon et al. 2017) and while at $0 < z \leq 4$ the slope of the $M_{\mathrm{UV}} - \mathrm{M}_\star$ relation is not evolving, the intercept evolves in such way that at a given stellar mass the UV luminosity decreases with increasing redshift (Williams et al. 2018). The evolution of this relation is more uncertain at $z > 4$ because of the uncertainties on the stellar mass estimation due to the nebular emission contamination (e.g., De Barros et al. 2014), but assuming a similar evolution at $z > 4$, combining the evolution of the UV luminosity at a given stellar mass

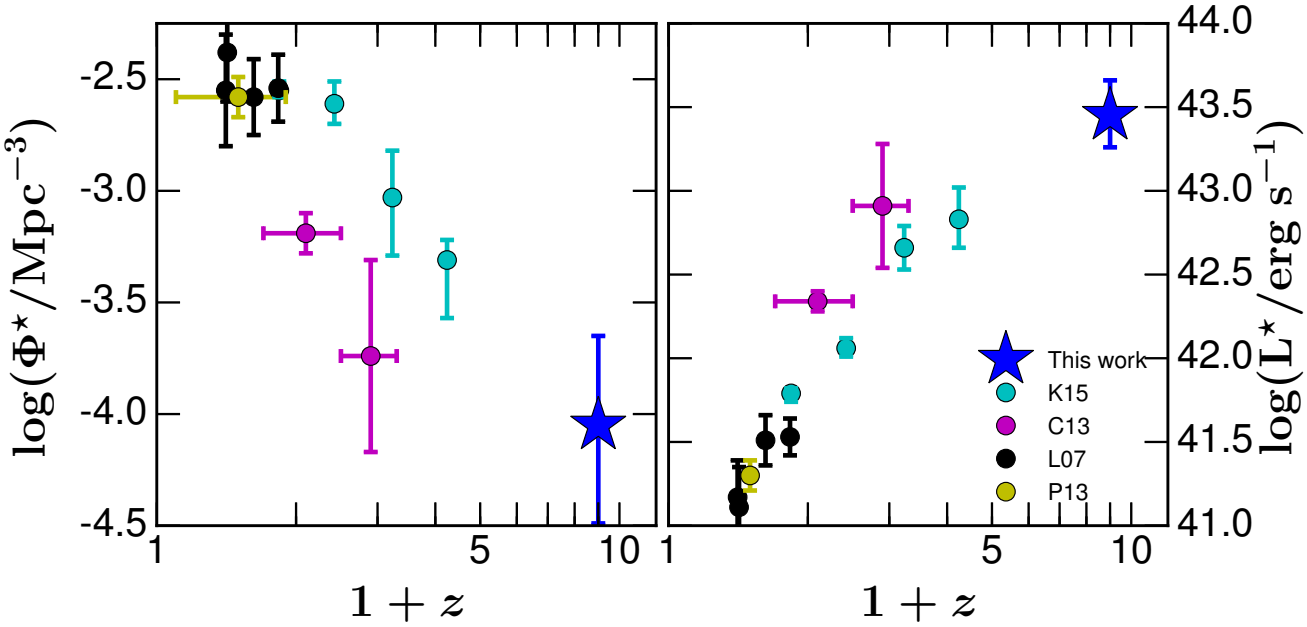

Figure 8. Evolution of $\Phi^\star$ (left panel) and $L^\star$ (right panel) with redshift with data from Khostovan et al. (2015, K15), Colbert et al. (2013, C13), Ly et al. (2007, L07), and Pirzkal et al. (2013, P13). Our values of $\Phi^\star$ and $L^\star$ agree with the extrapolated trends observed at lower redshift, indicating a relatively smooth evolution of the emission line LF.

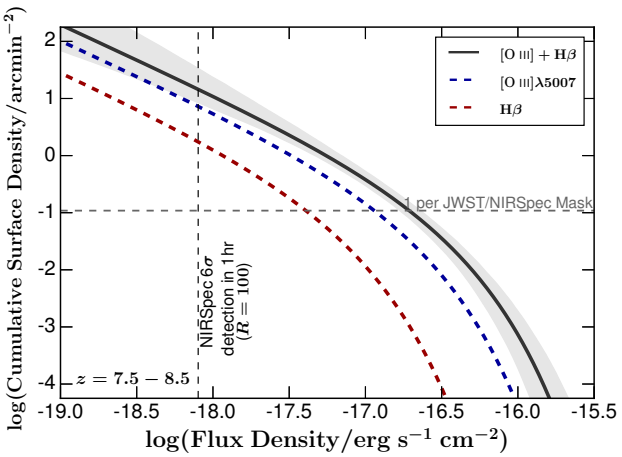

Figure 9. Cumulative surface density as a function of emission line flux densities at $z \sim 8$ derived from the [O III]$\lambda\lambda$4959, 5007+Hβ luminosity function presented in this work. We also show the relations for [O III]λ5007 and Hβ, assuming the median contribution of each of these lines to the total [O III]$\lambda\lambda$4959, 5007+Hβ luminosity obtained from the SED fitting. We use this figure to make *JWST* number counts prediction. Already at the 1hr depth of NIRSpec, we expect to be able to more than fill an entire mask with $z \sim 8$ galaxies, but deeper pre-imaging is required to identify these targets in most fields (see Text).

and the relation observed between nebular emission line EW and stellar mass (Fig. 5; Fumagalli et al. 2012; Sobral et al. 2014; Khostovan et al. 2016), we expect that with increasing redshift, at a given UV luminosity, the stellar mass decreases and the EW([O III]+Hβ) increases accordingly. This means that at a given UV luminosity the [O III]+Hβ luminosity is increasing with increasing redshift. This is the expected trend but uncertainties on the stellar mass estimation (Fig. 5) precludes any further quantification of the UV luminosity vs. stellar mass relation evolution from $z \sim 3$ to $z \sim 8$.

To test this explanation, we derive the relation between L([O III]+Hβ) and L(UV) at $z \sim 3$ through abundance matching. In particular, we use the $z \sim 3$ UV LF from Reddy & Steidel (2009) to derive the cumulative number density of galaxies at a given UV luminosity and match it to the corresponding cumulative number density at a given [O III]+Hβ luminosity based on the $z = 3.2$ [O III]+Hβ LF from Khostovan et al. (2015). We show the resulting relation in Fig. 6 with a blue line. There is a clear evolution of the L(UV) vs. L([O III]+Hβ) relation from $z \sim 3$ to $z \sim 8$, galaxies being brighter in [O III]+Hβ at any given UV luminosity explored in this work and an increasing difference between the $z \sim 3$ and $z \sim 8$ relation with decreasing UV luminosity. This finding supports our explanation for the [O III]+Hβ LF evolution.

6.3 Predictions of JWST Number Counts

One of the most awaited capabilities of the upcoming *JWST* is its unprecedented sensitivity for spectroscopy at $> 2\mu$m. In particular, *JWST* will for the first time provide spectroscopic access to the rest-frame optical emission lines of very high redshift galaxies, including the [O III]$\lambda\lambda$4959, 5007 and Hβ lines at $z \sim 8$ that we constrained through photometry here. In order to decide on the area and depth for the most efficient spectroscopic surveys with *JWST*, an estimate of the [O III]$\lambda\lambda$4959, 5007+Hβ LF as derived above is of critical importance.

Of particular interest for such *JWST* predictions is the

NIRSpec instrument (Bagnasco et al. 2007), which will be the workhorse NIR spectrograph. With four quadrants, each of which covers 2.3 arcmin^2, NIRSpec spans an effective area of ∼ 9.2 arcmin^2. Its sensitivity is exquisite. In only 1hr, NIRSpec will reach 6σ detections for emission lines at ∼4.5μm and fluxes of ∼6×10^{-19} erg s^{-1}cm^{-2} at R = 1000, or ∼8×10^{-19} erg s^{-1}cm^{-2} at R = 100. These numbers were derived with the latest *JWST*/ETC, when integrating over the full extent of the lines (which were assumed to have an intrinsic width of 150 km s^{-1}).

In Fig. 9, we plot the cumulative surface density of galaxies at z = 7.5 − 8.5 as a function of emission line luminosities based on the line LF which we derived in the previous section. In particular, we show the total combined [O III]$\lambda\lambda$4959, 5007+Hβ luminosities (as would be seen, e.g., in R = 100 low-resolution spectroscopy with NIRSpec), and we also split the luminosities into the three different lines, Hβ, [OIII]4959, and [OIII]5007. For the latter step, we employ the median line ratios as found in the SED fitting in Section 4, with Hβ/(Hβ+[O III]$\lambda\lambda$4959, 5007)=0.21, [O III]λ4959/(Hβ+[O III]$\lambda\lambda$4959, 5007)=0.20, and [O III]λ5007/(Hβ+[O III]$\lambda\lambda$4959, 5007)=0.59.

As can be seen, at the 1hr sensitivity limits of NIRSpec with R = 100, we can expect a cumulative surface density of z = 7.5 − 8.5 galaxies of $10^{1.2\pm0.3}$ arcmin^{-2} which have blended [OIII]+Hβ lines that are bright enough to be significantly detected. This means that a single NIRSpec mask (with effective area 9.2 arcmin^2) would, in principle, be able to target on average ∼ 150 galaxies (where the 1σ uncertainties range from 80 to 300 galaxies). The fixed grid of the NIRSpec slitlet masks will reduce this number somewhat. Unfortunately, however, there is an additional limitation of early spectroscopic surveys. The depth reached in terms of [O III]+Hβ luminosity in 1hr for NIRSpec corresponds to an observed UV magnitude of $m_{\rm UV}$ = 29.9 (∼ 29 − 31 when accounting for the uncertainties in the UV-[O III]+Hβ luminosity relation, Eq. 1). This implies that the average surface density of current z ∼ 8 galaxy samples in the prime extragalactic legacy fields such as CANDELS is significantly lower than the number above (Bouwens et al. 2015). Therefore, early *JWST* spectroscopic surveys, which are based on the selection of targets from current *HST* datasets, cannot be maximally efficient for a targeted z ∼ 8 galaxy survey. The best strategy will be to perform deep pre-imaging with *JWST* to identify targets at z ∼ 8.

Nevertheless, our calculation shows that if significantly deep imaging data are available to select targets from, a single NIRSpec mask with R = 100 could be filled with just z ∼ 8 emission line sources for which a 1hr observation can measure a secure redshift. This will result in revolutionary insights of the large scale structure in the heart of the reionization epoch.

Of course, in order to study the physics, more than a simple redshift measurement is required. In particular, the [OIII]+Hβ lines need to be split with observations at R = 1000 or higher. For such surveys, the corresponding number of expected galaxies in 1hr observations are only 24 galaxies for Hβ lines, and 97 galaxies with 6σ [OIII]5007 line detections, per NIRSpec mask.

As a final remark, we compare our predictions with the ones from the JAGUAR mock catalog. Given the lower [O III]+Hβ luminosities at a given L(UV) compared to our observed sample, it is clear that the mocks will significantly underpredict the number of observed sources at a given line luminosity. When computing the line LF from the JAGUAR catalogs, we find a lower normalization by a factor ∼ 8× compared to our observed LF. Hence, the JAGUAR mock catalogs should underpredict the number of rest-frame optical emission lines that can be detected at z ∼ 8 with *JWST*/NIRSpec in the future by the same factor.

7 CONCLUSIONS

We have presented a detailed analysis of a z ∼ 8 galaxy sample with some of the deepest available *Spitzer* observations, from the GREATS survey. The sample has been culled through photometric redshifts to ensure that the selected galaxies are at $z \geq 7.11$, where the IRAC 3.6μm−IRAC 4.5μm colors put strong constraints on the [O III]+Hβ equivalent width. We built and used a photoionization grid with a large parameter space covering a variety of stellar metallicity and ISM conditions, using the BPASS models as stellar emission inputs. We feed the resulting SEDs that include stellar and nebular emission (continuum and lines) to our SED fitting code to derive z ∼ 8 galaxy properties. Accounting for the photometric and model uncertainties, we have specifically derived the [O III]+Hβ luminosities, allowing us to derive the corresponding [O III]+Hβ luminosity function and make predictions for *JWST* observations.

In summary, we find the following.

(i) Our subsample with S/N(3.6μm ∨ 4.5μm) ≥ 3 has the following average properties: $\log(M_\star/M_\odot) = 8.62^{+0.43}_{-0.39}$, $\log(\text{age/yr}) = 7.2^{+0.9}_{-0.6}$, $\log(\text{SFR}/M_\odot\text{yr}^{-1}) = 1.26^{+0.42}_{-0.30}$, A_V = 0.4 ± 0.2, and $\text{sSFR} = 63^{+188}_{-55}\text{Gyr}^{-1}$.

(ii) To reproduce the observed IRAC color of this subsample, which is strongly affected by [O III]$\lambda\lambda$4959, 5007+Hβ emission, the two main parameters driving the EW are the stellar metallicity and the ionization parameter, and they have the following values $Z_\star = 0.004^{+0.004}_{-0.002}$ and $\log U = -3.0 \pm 1.0$.

(iii) We are able to put constraints on the median ionizing photon production efficiency with $\log(\xi_{\rm ion}/\text{erg}^{-1}\ \text{Hz}) \geq 25.77$. This latter value is > 3 times higher than the canonical value, implying that these galaxies have a higher ionizing output than typically assumed and can thus more easily reionize the universe.

(iv) According to our SED fitting which matches the observed IRAC colors, we find a median rest-frame equivalent width $\text{EW}([\text{O III}]\lambda\lambda4959, 5007 + \text{H}\beta) = 649^{+92}_{-49}$Å (Fig. 5).

(v) We find a relatively tight relation between [O III]+Hβ and UV luminosity (Fig. 6), allowing us to derive for the first time the [O III]+Hβ LF at z ∼ 8 based on the z ∼ 8 UV LF. We find that, in contrast with the evolution of the UV LF from z ∼ 3 to z ∼ 8, the z ∼ 8 [O III]+Hβ LF is higher at all luminosities than at z ∼ 3. This is due to the increasing [O III]+Hβ luminosity at a given UV luminosity with increasing redshift.

(vi) Finally, we use the derived [O III]+Hβ LF to predict *JWST* number counts. A single NIRSpec pointing would contain ∼ 150 galaxies at z = 7.5 − 8.5, for which the [O III]+Hβ emission could be detected in only 1hr. However, the current average surface density of z ∼ 8 galaxies in extra-

galactic legacy fields is significantly lower that this number. Therefore, to maximize the efficiency of *JWST* deep pre-imaging to $m_{\rm UV} \sim 30$ mag will be required.

In this work, we have used the deepest *Spitzer* data available on large areas. We have accounted for observational uncertainties on the photometry and we have used a grid of photoionization models with a large parameter space. While we have attempted to minimize the number of assumptions going into our analysis, many modeling uncertainties are still present: the ingredients in the stellar population synthesis models (stellar atmospheres, binaries, rotation), the IMF, the possible presence of multiple stellar populations, the dust attenuation curve, the ratio between nebular and stellar attenuation, interstellar abundances, and depletion factor of metals on to dust grains. Only the unprecedented abilities of *JWST* will allow to alleviate some of these uncertainties.

ACKNOWLEDGEMENTS

We thank the anonymous referee who helped improve this manuscript. The work of SDB has been partially supported by a Flexibility Grant from the Swiss National Science Foundation and by a MERAC Funding and Travel Award from the Swiss Society for Astrophysics and Astronomy. V. G. was supported by CONICYT/FONDECYT initiation grant number 11160832.

This work made use of v2.1 of the Binary Population and Spectra Synthesis (BPASS) models as last described in Eldridge et al. (2017). Calculations were performed with version 17.00 of Cloudy last described by Ferland et al. (2017). This work also made use of Astropy, a community-developed core Python package for Astronomy (The Astropy Collaboration et al. 2018), as well as the pymc3 library (Salvatier et al. 2016).

This paper made use of public catalogs derived from data taken by the Sloan Digital Sky Survey IV. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org.

REFERENCES

Abolfathi B., et al., 2018, ApJS, 235, 42
Amorín R., et al., 2017, Nature Astronomy, 1, 0052
Bagnasco G., et al., 2007, in Cryogenic Optical Systems and Instruments XII. p. 66920M, doi:10.1117/12.735602
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Berg D. A., Erb D. K., Auger M. W., Pettini M., Brammer G. B., 2018, ApJ, 859, 164
Bolzonella M., Miralles J., Pelló R., 2000, A&A, 363, 476
Bouchet P., Lequeux J., Maurice E., Prevot L., Prevot-Burnichon M. L., 1985, A&A, 149, 330
Bouwens R. J., et al., 2014, ApJ, 793, 115
Bouwens R. J., et al., 2015, ApJ, 803, 34
Bouwens R. J., Smit R., Labbé I., Franx M., Caruana J., Oesch P., Stefanon M., Rasappu N., 2016, ApJ, 831, 176
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cardamone C., et al., 2009, MNRAS, 399, 1191
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
Ceverino D., Klessen R. S., Glover S. C. O., 2019, MNRAS, 484, 1366
Chary R.-R., Stern D., Eisenhardt P., 2005, ApJ, 635, L5
Chevallard J., Charlot S., 2016, MNRAS, 462, 1415
Chevallard J., et al., 2018a, MNRAS,
Chevallard J., et al., 2018b, MNRAS, 479, 3264
Colbert J. W., et al., 2013, ApJ, 779, 34
De Barros S., Schaerer D., Stark D. P., 2014, A&A, 563, A81
De Barros S., Reddy N., Shivaei I., 2016, ApJ, 820, 96
Domínguez A., et al., 2013, ApJ, 763, 145
Duncan K., et al., 2014, MNRAS, 444, 2960
Eldridge J. J., Izzard R. G., Tout C. A., 2008, MNRAS, 384, 1109
Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, Publ. Astron. Soc. Australia, 34, e058
Ellis R. S., et al., 2013, ApJ, 763, L7
Faisst A. L., et al., 2016, ApJ, 821, 122
Ferland G. J., et al., 2017, Rev. Mex. Astron. Astrofis., 53, 385
Finlator K., Davé R., Oppenheimer B. D., 2007, MNRAS, 376, 1861
Fumagalli M., et al., 2012, ApJ, 757, L22
Giavalisco M., et al., 2004, ApJ, 600, L103
González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., Brammer G. B., 2010, ApJ, 713, 115
González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., 2011, ApJ, 735, L34+
Grazian A., et al., 2015, A&A, 575, A96
Grogin N. A., et al., 2011, ApJS, 197, 35
Gutkin J., Charlot S., Bruzual G., 2016, MNRAS, 462, 1757
Hippelein H., et al., 2003, A&A, 402, 65
Illingworth G. D., et al., 2013, ApJS, 209, 6
Inami H., et al., 2017, A&A, 608, A2
Izotov Y. I., Guseva N. G., Thuan T. X., 2011, ApJ, 728, 161
Izotov Y. I., Guseva N. G., Fricke K. J., Henkel C., Schaerer D., 2017, MNRAS, 467, 4118
Jaskot A. E., Ravindranath S., 2016, ApJ, 833, 136
Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, MNRAS, 353, 713
Kewley L. J., Dopita M. A., 2002, ApJS, 142, 35
Khostovan A. A., Sobral D., Mobasher B., Best P. N., Smail I., Stott J. P., Hemmati S., Nayyeri H., 2015, MNRAS, 452, 3948
Khostovan A. A., Sobral D., Mobasher B., Smail I., Darvish B., Nayyeri H., Hemmati S., Stott J. P., 2016, MNRAS, 463, 2363
Koekemoer A. M., et al., 2011, ApJS, 197, 36
Labbé I., et al., 2010, ApJ, 716, L103
Labbé I., et al., 2013, ApJ, 777, L19
Labbé I., et al., 2015, ApJS, 221, 23
Lam D., et al., 2019, arXiv e-prints,
Ly C., et al., 2007, ApJ, 657, 738
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Mármol-Queraltó E., McLure R. J., Cullen F., Dunlop J. S., Fontana A., McLeod D. J., 2016, MNRAS, 460, 3587
Mashian N., et al., 2015, ApJ, 802, 81
Matthee J., Sobral D., Best P., Khostovan A. A., Oteo I., Bouwens R., Röttgering H., 2017a, MNRAS, 465, 3637
Matthee J., Sobral D., Darvish B., Santos S., Mobasher B., Paulino-Afonso A., Röttgering H., Alegre L., 2017b, MNRAS, 472, 772
Nakajima K., et al., 2018, A&A, 612, A94
Oesch P. A., et al., 2015, ApJ, 804, L30

Oke J. B., Gunn J. E., 1983, ApJ, 266, 713
Pirzkal N., et al., 2013, ApJ, 772, 48
Prevot M. L., Lequeux J., Prevot L., Maurice E., Rocca-Volmerange B., 1984, A&A, 132, 389
Rasappu N., Smit R., Labbé I., Bouwens R. J., Stark D. P., Ellis R. S., Oesch P. A., 2016, MNRAS, 461, 3886
Reddy N. A., Steidel C. C., 2009, ApJ, 692, 778
Reddy N. A., et al., 2015, ApJ, 806, 259
Rigby J. R., Bayliss M. B., Gladders M. D., Sharon K., Wuyts E., Dahle H., Johnson T., Peña-Guerrero M., 2015, ApJ, 814, L6
Roberts-Borsani G. W., et al., 2016, ApJ, 823, 143
Salmon B., et al., 2015, ApJ, 799, 183
Salvatier J., Wiecki T. V., Fonnesbeck C., 2016, PyMC3: Python probabilistic programming framework, Astrophysics Source Code Library (ascl:1610.016)
Schaerer D., De Barros S., 2009, A&A, 502, 423
Schaerer D., De Barros S., 2010, A&A, 515, A73+
Schaerer D., Izotov Y. I., Verhamme A., Orlitová I., Thuan T. X., Worseck G., Guseva N. G., 2016, A&A, 591, L8
Senchyna P., et al., 2017, MNRAS, 472, 2608
Sharma S., 2017, ARA&A, 55, 213
Shim H., Chary R.-R., Dickinson M., Lin L., Spinrad H., Stern D., Yan C.-H., 2011, ApJ, 738, 69
Shivaei I., Reddy N. A., Steidel C. C., Shapley A. E., 2015, ApJ, 804, 149
Shivaei I., et al., 2018, ApJ, 855, 42
Smit R., Bouwens R. J., Franx M., Illingworth G. D., Labbé I., Oesch P. A., van Dokkum P. G., 2012, ApJ, 756, 14
Smit R., et al., 2014, ApJ, 784, 58
Smit R., et al., 2015, ApJ, 801, 122
Smit R., Bouwens R. J., Labbé I., Franx M., Wilkins S. M., Oesch P. A., 2016, ApJ, 833, 254
Smit R., Swinbank A. M., Massey R., Richard J., Smail I., Kneib J.-P., 2017, MNRAS, 467, 3306
Sobral D., Best P. N., Smail I., Mobasher B., Stott J., Nisbet D., 2014, MNRAS, 437, 3516
Sobral D., et al., 2018, MNRAS, 477, 2817
Song M., Finkelstein S. L., Livermore R. C., Capak P. L., Dickinson M., Fontana A., 2016, ApJ, 826, 113
Stanway E. R., Eldridge J. J., Becker G. D., 2016, MNRAS, 456, 485
Stark D. P., Ellis R. S., Bunker A., Bundy K., Targett T., Benson A., Lacy M., 2009, ApJ, 697, 1493
Stark D. P., Schenker M. A., Ellis R., Robertson B., McLure R., Dunlop J., 2013, ApJ, 763, 129
Stark D. P., et al., 2014, MNRAS, 445, 3200
Stark D. P., et al., 2015a, MNRAS, 450, 1846
Stark D. P., et al., 2015b, MNRAS, 454, 1393
Stark D. P., et al., 2017, MNRAS, 464, 469
Stefanon M., Bouwens R. J., Labbé I., Muzzin A., Marchesini D., Oesch P., Gonzalez V., 2017, ApJ, 843, 36
Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, ApJ, 462, L17
Steidel C. C., Strom A. L., Pettini M., Rudie G. C., Reddy N. A., Trainor R. F., 2016, ApJ, 826, 159
Storey P. J., Hummer D. G., 1995, MNRAS, 272, 41
Tang M., Stark D., Chevallard J., Charlot S., 2018, preprint, (arXiv:1809.09637)
The Astropy Collaboration et al., 2018, AJ, 156, 123
Theios R. L., Steidel C. C., Strom A. L., Rudie G. C., Trainor R. F., Reddy N. A., 2019, ApJ, 871, 128
Tremonti C. A., et al., 2004, ApJ, 613, 898
Vanzella E., et al., 2017, ApJ, 842, 47
Williams C. C., et al., 2018, ApJS, 236, 33
Wofford A., et al., 2016, MNRAS, 457, 4296
Yabe K., Ohta K., Iwata I., Sawicki M., Tamura N., Akiyama M., Aoki K., 2009, ApJ, 693, 507
Yang H., Malhotra S., Rhoads J. E., Wang J., 2017, ApJ, 847, 38
Zackrisson E., Bergvall N., Olofsson K., Siebert A., 2001, A&A, 375, 814
Zackrisson E., Rydberg C.-E., Schaerer D., Östlin G., Tuli M., 2011, ApJ, 740, 13
Zitrin A., et al., 2015, ApJ, 810, L12

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