CAMBRIDGE
UNIVERSITY PRESS

Research Article

MOSEL survey: Unwrapping the Epoch of Reionisation through mimic galaxies at Cosmic Noon

Ravi Jaiswar[1,2] , Anshu Gupta[1,2], Elisabete da Cunha[2,3], Cathryn M. Trott[1,2] , Anishya Harshan[2,4], Andrew Battisti[2,5] and Ben Forrest[6]

[1]International Centre for Radio Astronomy Research, Curtin University, Bentley, Australia, [2]ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions, Bentley, Australia, [3]International Centre for Radio Astronomy Research, University of Western Australia, Crawley, Australia, [4]Department of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia, [5]Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT, Australia and [6]Department of Physics and Astronomy, University of California Davis, Davis, CA, USA

Abstract

The nature of the first galaxies that reionised the universe during the Epoch of Reionisation (EoR) remains unclear. Attempts to directly determine spectral properties of these early galaxies are affected by both limited photometric constraints across the spectrum and by the opacity of the intergalactic medium to the Lyman Continuum (LyC) at high redshift. We approach this by analysing properties of analogous extreme emission line galaxies (EELGs, [OIII]+Hbeta EW > 400) at $2.5 < z < 4$ from the ZFOURGE survey using the Multi-wavelength Analysis of Galaxy Physical Properties (MAGPHYS) SED fitting code. We compare these to galaxies at $z > 5.5$ observed with the James Webb Space Telesope with self-consistent spectral energy distribution fitting methodology. This work focuses on the comparison of their UV slopes (β_P), ionising photon production efficiencies ξ_{ion}, star formation rates and dust properties to determine the effectiveness of this analogue selection technique. We report the median ionising photon production efficiencies as $\log_{10}(\xi_{ion}/(Hz\ erg^{-1})) = 25.14^{+0.06}_{-0.04}$, $25.16^{+0.06}_{-0.05}$, $25.16^{+0.04}_{-0.05}$, $25.18^{+0.06}_{-0.07}$ for our ZFOURGE control, ZFOURGE EELG, JADES, and CEERS samples, respectively. ZFOURGE EELGs are 0.57 dex lower in stellar mass and have half the dust extinction, compared to their ZFOURGE control counterparts. They also have a similar specific star formation rates and β_P to the $z > 5.5$ samples. We find that EELGs at low redshift ($2.5 < z < 4$) are analogous to EoR galaxies in their dust attenuation and specific star formation rates. Their extensive photometric coverage and the accessibility of their LyC region opens pathways to infer stellar population properties in the EoR.

Keywords: Galaxies: intergalactic medium; general; high-redshift; photometry; starburst; evolution

1. Introduction

The model of the Universe's evolution through the Epoch of Reionisation (EoR) is directly affected by our limited understanding of the first galaxies. Debate of their ionising capabilities persists due to their relatively unconstrained ionising photon production and escape fractions (Robertson et al. 2010; Madau & Haardt 2015; Naidu et al. 2019; Finkelstein et al. 2019). While AGN-quasars are unlikely ionisation source candidates due to their infrequency beyond $z > 3$ (Kulkarni et al. 2018), bright, highly star-forming galaxies could have produced copious ionising photons for reionisation (Naidu et al. 2019). Alternatively, numerous faint galaxies could provide sufficient ionising flux over a longer history to reionise the intergalactic medium (Bruton et al. 2023).

Operation of the James Webb Space Telescope (JWST) has, in a short time, revolutionised our understanding of galaxies in the EoR and this landscape is constantly changing with new revelations (Bunker et al. 2023; Bagley et al. 2023; Paris et al. 2023). Early empirical studies had noticed potentially large [OIII]5007 + H_β

emissions at higher redshift (Schaerer & de Barros 2009; Raiter et al. 2010). Now with recent redshift evolution models (Zhai et al. 2019) and JWST/NIRCam number density studies $z > 5.3$ (Matthee et al. 2023) this is evidently ubiquitous in the early universe. Probing galaxies within the EoR reveals almost no dust attenuation as expected, particularly approaching the highest redshift limits (Robertson et al. 2023; Hsiao et al. 2023; Tacchella et al. 2023; Haro et al. 2023b). However, galaxies with unexpectedly high stellar masses have been found (Boylan-Kolchin 2022; Labbe et al. 2022) bordering the possible limits set by ΛCDM and suggesting extremely efficient star formation. Exceptional findings such as a highly quenched, dusty galaxy at $z \sim 5$ (Donnan et al. 2022; Harikane et al. 2022; Naidu et al. 2022; Haro et al. 2023a; Yung et al. 2023) and spectroscopic confirmation of a $z \sim 13.2$ (Robertson et al. 2022) galaxy have both challenged and affirmed ideas of the timescale over which the Universe evolves.

The LyC is the component of stellar emission with sufficient energy to ionise neutral hydrogen atoms and is a key parameter in theoretical models of EoR. However, deriving physical properties relating to the production of LyC radiation and their escape from the interstellar medium of galaxies comes with both technical and physical challenges. Direct detection of LyC photons is impeded by the opacity of neutral hydrogen in EoR galaxies, instead requiring models based on lower redshift analogues (Izotov et al. 2018).

LyC leakers at low redshift exhibit high [OIII]5007 EWs and high O_{32} ([OIII]5007/[OII]3727) ratios albeit with large scatter (Cardamone et al. 2009; Izotov et al. 2018). For galaxies $z < 0.5$, a Lyman-α peak separation of less than 200 km/s is the strongest predictor of significant LyC escape at $z \sim 0$ ($f_{esc} > 0.1$ up to 0.8) (Izotov et al. 2022). This is the basis of EoR-Lyman-α correlated f_{esc} predictions, with > 10 Gyr between the two epochs.

It is possible that these low redshift f_{esc} probes may have a different relationship with LyC escape at high redshifts due to alternate escape methods. Ji et al. (2020) find no Lyman-α emission (only absorption) despite strong LyC detections in a $z \sim 3.8$ galaxy; favouring a model of f_{esc} through LyC-transparent holes rather than through an optically thin interstellar medium (ISM). The assumption that $z < 0.5$ and $z > 5$ galaxies are comparable is in contention with our understanding of these epochs (Madau & Dickinson 2014). Studies using one epoch to analyse another must contend with their fundamental differences and what their comparisons really reveal. For example, the Lyman-α probe or UV slope may have a different relative dependence on dust extinction and the stellar population due to morphological differences (Hayes et al. 2011; Meng & Gnedin 2020).

With the advent of the JWST, recent observations targeting the H_α and H_β Balmer lines have been used to estimate the production efficiency of hydrogen ionising photons (ξ_{ion}) for many $z > 6$ sources (Tang et al. 2023; Simmonds et al. 2023). However, with the existing uncertainty in dust content of high redshift sources, emission line corrections may not accurately reproduce the original luminosity as intended. Furthermore, limited spectroscopic availability means statistically significant studies come to rely on photoionisation models (Ferland et al. 1998; Levesque et al. 2010) to estimate ξ_{ion} (Bouwens et al. 2015) and compare its correlation to other ISM probes such as [OIII]5007+H_β EW (Kewley et al. 2015; Tang et al. 2019, 2023)

The lack of photometric measurements/coverage between the UV and NIR diminishes constraints on the modelled physical properties at $z > 6$. Quantities such as metal abundances (Vincenzo & Kobayashi 2018; Torrey et al. 2018; Berg et al. 2019), the escape fraction f_{esc} or the number of ionising photons $\dot{n}$ (Anderson et al. 2017; Iyer & Gawiser 2017) rely heavily on star formation history (SFH) models to infer the absorbed light, which depends strongly on properly sampling the spectrum.

Having a well sampled SED for galaxies with well constrained redshifts allows for the determination of a photometric UV slope (β_p). This probes the empirical emission 'blueness' which is used in dust corrections (Calzetti et al. 1994; Reddy et al. 2018) and as a proxy for the stellar ages and metal/dust content. Studies of local universe starbursts find that galaxies with a bluer, more negative model β (determined from f_λ) have lower dust obscuration, similar to $z \sim 2$ studies (Takeuchi et al. 2012; Sklias et al. 2013). The relationship of β and dust attenuation is degenerate with metallicity and star formation history (Bouwens et al. 2016) which complicates the relative contributions of each component. However, correlations to dust content are robust at fixed UV luminosity. Reddy et al. (2018) and Nanayakkara et al. (2022) find that galaxies at $4 < z < 7$ are consistently blue and have very low dust attenuation.

The approach this paper takes to these issues in studying the potential influence of the brightest emitters on the EoR is to use an established analogue sample with the stellar and [OIII]5007 emission properties of higher redshift galaxies at the slightly lower redshift range of $2.5 < z < 4$ (11.1–12.2 Gyr ago).

The selected redshift range is temporally closer to the EoR (12.7 Gyr ago) than many low redshift comparisons, while still having an accessible LyC region for follow up observational f_{esc} constraints. The sample is taken from the FourStar Galaxy Evolution Survey (ZFOURGE) (Straatman et al. 2016), which combines deep imaging with medium and narrow-band near-IR filters with multiwavelength observations from several public surveys to accurately determine the redshifts of $\sim 70\,000$ objects to $\sigma_z \sim 2\%$ accuracy (Nanayakkara et al. 2016; Tran et al. 2020).

We will be exploring a subsample of these galaxies presented by Forrest et al. (2018) as having extreme H_β+[OIII]5007 EW (>400) and therefore considered analogues of EoR galaxies (Tang et al. 2023). Follow up spectroscopy with the KMOS /VLT as part of the Mutli-Object Spectroscopic Emission Line (MOSEL) survey confirms the photometric selection for a subset of these as EELGs (Tran et al. 2020; Gupta et al. 2022). Of 19 galaxies targeted, 16 had bright emission lines, where 14 of these had [OIII]5007 as the brightest line and two had H_α. This sample (40 filters) will be compared to direct EoR galaxy measurements made by the CEERS (14 filters) (Yang et al. 2020; Bagley et al. 2023; Yang et al. 2023) and JADES (23 filters) (Bunker et al. 2023; Eisenstein et al. 2023; Hainline et al. 2023; Rieke et al. 2023) surveys using legacy HST and the recent JWST observations to determine the overlap in their observed and the SED model-derived properties. For consistency, we use the Multi-wavelength Analysis of Galaxy Physical Properties (MAGPHYS) SED fitting code (da Cunha et al. 2008, 2015) across all samples.

This work is formatted so that Section 2 describes the data and selection philosophies used to identify Extreme Emission galaxy analogues while Section 3 delves into our methods of analysis and the quantities we use to compare the samples. Results of how representative the analogues are and what their high energy emission properties look like are presented in Section 4. We summarize the work in Section 5.

Throughout this paper we assume a flat universe with $\Omega_m = 0.3$, $\Omega_\lambda = 0.7$ and $H_0 = 70$ km/s/Mpc for our models.

2. Data

2.1 The analogues

In this section we break down the selection criteria and processing of the $2.5 < z < 4$ ZFOURGE control sample Section 2.1.1 as well as the EELG analogues Section 2.1.2. See Fig. 1 for the sample processing.

2.1.1 ZFOURGE control

ZFOURGE is a survey combining legacy photometric UV to NIR data from 3 well studied regions (CDFS, COSMOS and UDS) with the FOURSTAR instrument upon the Magellan telescope (Straatman et al. 2016); its J, H and K medium band filters spanning the 1–1.8 μm range. It creates a photometrically well sampled survey and with the inclusion of two narrowband filters to optimally constrain the 4 000 Å break, enables robust 2% accuracy redshifts for over 70 000 galaxies (Nanayakkara et al. 2016). In this work the ZFOURGE survey data was reduced to only retain CDFS sources as this field had the best narrowband filter depth (NB118 and NB209) which is important in isolating the [OIII]5007 line flux for $z \sim 3$ sources. Galaxies were selected so that their K-band SNR>20 and were between the 2.5–4 redshift range. We used the use = 1 flag which eliminated catastrophic FAST (Kriek et al. 2009)
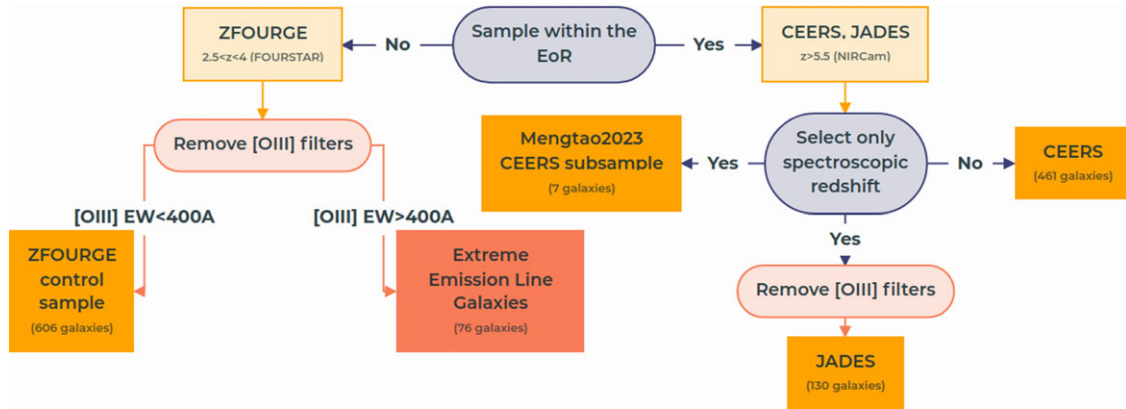
Figure 1. Summary of samples and their constraining quantities as well as some processing information.

and EAZY (Brammer et al. 2008) SED fits, interloper stars, AGN and sources too close to bright stars. The errors for IRAC bands were set to a floor of 25% (if not already above this) as the large PSF requires that the errors be large at high redshift. (Straatman et al. 2016). These were otherwise unphysically small and were found to significantly reduce the number of galaxies with well constrained physical stellar parameters. This created a total sample of 682 galaxies which includes both the control sample (606 galaxies) and the Extreme Emission Line Galaxy sample (76 galaxies) (see Section 2.1.2).

2.1.2 Filtering ZFOURGE EELGs at $z \sim 3$

ZFOURGE EELGs were identified in Forrest et al. (2018) using a stacked superposition of similar galaxy SEDs. The EELGs were determined as having an [OIII]5007 EW of at least 400 Å and were selected between $2.5 < z < 4$. This creates a subsample of 76 EELGs among the 682 total galaxies within the CDFS region, described in the Section 2.1.1. We also use 18 of the 19 reliable spectroscopic redshifts from the MOSEL survey subsample which replace the corresponding ZFOURGE photometric redshifts.

2.2 Epoch of reionisation sample

In this section, we look at the selection criteria and processing of the $z > 5.5$ control samples. This is broken down into the photometric redshift sample taken from the CEERS survey (Section 2.2.1) as well as the subsample of this which have spectroscopic redshifts. This is followed by the JADES survey sample which is only has spectroscopic redshifts (Section 2.2.3).

2.2.1 CEERS

The Cosmic Evolution Early Release Science Survey (Yang et al. 2020; Bagley et al. 2023; Yang et al. 2023) combines JWST NIRCAM photometry with legacy Extended Groth Strip (EGS) field HST data with the science goal of scouting the emergence of galaxies at cosmic dawn. We use the September 2022 data release processed with Grizli[a] where photometric redshifts were derived with EAZY-py. We select galaxies within $5.5 < z < 14$ and eliminate sources with a 95% confidence interval of the χ^2 value above 0.2 Gyr ($t_{z[97.5]} - t_{z[2.5]} < 0.2$ Gyr), which similarly constrains the

---

[a]https://s3.amazonaws.com/grizli-v2/JwstMosaics/v4/index.html.

low and high ends of the redshift bounds. We further require that the determination of the UV slope be based on at least 3 filters in the UV window (Calzetti et al. 1994) and remove sources which do not fit this criteria. This selects only the galaxies with a well constrained photometric redshifts and UV regions, giving us a sample of 461 galaxies within the EoR to compare to the ZFOURGE galaxies, particularly the EELG subsample.

2.2.2 Spectroscopic redshifts

For a small subsample of the CEERS catalogue, we have collected published spectroscopic redshifts from Tang et al. (2023) to minimise fitting errors within MAGPHYS and compare with the results of the publication. These will be referred to as the Tang23 sample throughout the paper.

2.2.3 JADES

The JWST Advanced Deep Extragalactic Survey (JADES, Bunker et al. 2023; Eisenstein et al. 2023; Hainline et al. 2023; Rieke et al. 2023) combines JWST NIRCAM photometry with legacy HST Ultra Deep Field (UDF) data. We only select spectroscopically confirmed galaxies at $z > 5.5$, deriving a sample of 130 galaxies between $5.5 < z < 13.2$. The JADES datasets depth and additional filters help to better constrain the physical parameters, and thus we compare mostly to this sample.

3. Methods

In this section we discuss the derivation of parameters used to describe the 'blueness' (UV slope, β) of a galaxy's emission profile and its likelihood to contribute LyC to the IGM from both photometric data. We describe the SED models used, their limitations and any modifications we have made.

3.1 MAGPHYS

Multi-wavelength Analysis of Galaxy Physical Properties (MAGPHYS) is an SED fitting package which derives the physical properties of a galaxy from the supplied photometry (da Cunha et al. 2008, 2015). It does this by assembling a library of dust and stellar models at a predetermined redshift (if not allowed to vary) and then approaches the closest model using a marginalised likelihood distribution of each physical parameter (for more information see the documentation of MAGPHYS).
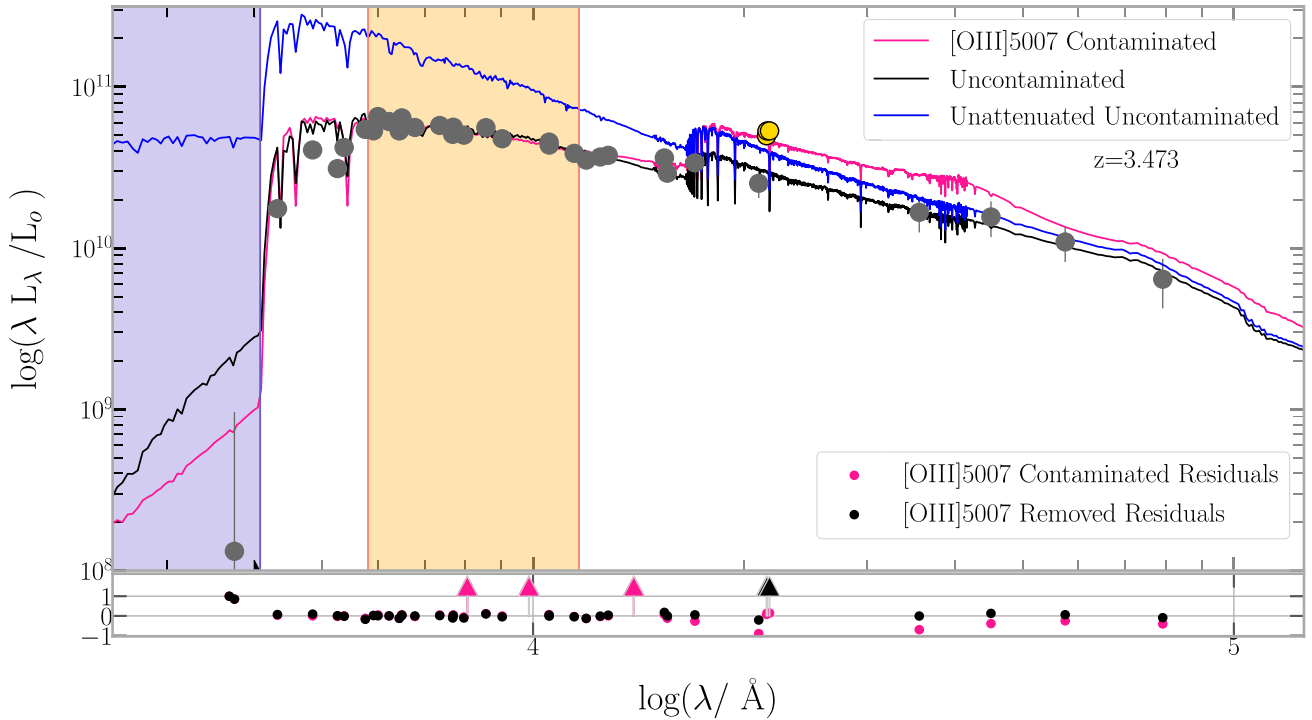
Figure 2. Observed frame MAGPHYS SED of $z = 3.473$ ZFOURGE EELG indicating the models fit to the photometric filter fluxes (grey circles) and their residuals (black circles, arrows if residual $\sigma_{res} > 1$). Blue region represents the LyC and is integrated under the unattenuated uncontaminated line (blue) to derive $\log_{10}(\xi_{ion}/(Hz \, erg^{-1}))$ photometrically. Orange region is the UV region from which the UV slope is determined from either the attenuated SED β_a or from the photometry β_p. Gold filters represent the [OIII]5007 contaminated filters and the difference these make to the attenuated SED is shown in the continuum luminosity difference between the 'uncontaminated' (black) and 'contaminated' (pink) lines. Residuals of the contaminated bestfit are represented by pink circles where $\sigma_{res} > 1$ are represented by pink arrows.

We use the BC03 stellar models (Bruzual & Charlot 2003), (Charlot & Fall 2000) dust models, (Hildebrand 1983) grey body dust emission, (Chabrier 2003) IMF, (Madau 1995) IGM attenuation model and an exponentially declining star formation history model. The SFR timescale is over the past 100 Myr.

In our analysis, we use a modified version of the high_z version of the code with a lower dust prior, an increased range of available redshifts (beyond $z > 10$) and an approximately 10-fold increased model sampling at higher redshifts ($z \sim 8$). Dust attenuation is expected to be low at high redshifts due to the early galactic stars being mostly composed of hydrogen and not having yet seeded a significant metallicity content into the ISM, of which dust is made (Shapley et al. 2023; Cameron et al. 2023). While some dusty galaxies have been discovered in early epochs (Donnan et al. 2022), the original prior is based on local universe observations which are more consistently dusty and so required lowering for improved fits to the higher redshift sample. The introduction of additional models at high redshift allows us to more finely sample the parameter space and derive a better fit to the given photometry. This was found to be particularly necessary to prevent the highest redshift surveys with low sampling from sharing similar ill-fitting models across many different galaxies.

In order to determine the intrinsic production efficiency of ionizing photons we also remove the IGM absorption component of the SED <u>after</u> the other parameters are fit such that their determination is unaffected. The integration of the $\lambda_{rest} < 912$ Å region is then used to derive the flux of the Lyman Continuum region. The physical parameters are still modelled using the Madau (1995)

prescription and determined by their likelihood distributions. While MAGPHYS simultaneously determines dust and stellar components, we do not report the dust mass or luminosity as our photometric sampling does not constrain the mid-far IR range.

[OIII]5007 'contaminated' filters also limit the accuracy of the SED fit, due to these bright lines distorting the flux output in their containing filters. For a discussion of the [OIII]5007 contamination and the method by which we account for this see Section 3.4. Fig. 2 breaks down a MAGPHYS SED fit to a ZFOURGE EELG photometry which has been sectioned into wavelength regions of interest and identifies the 'contaminated' filters.

3.2 Ionizing photon production efficiency $\log_{10}(\xi_{ion})$

The contribution of a source to reionisation is described by the equation

$$\dot{n}_{ion} = f_{esc} \times \xi_{ion} \times \rho_{UV} \qquad (1)$$

where $\dot{n}_{ion}$ is the production rate of ionizing radiation, f_{esc} is the escape fraction of LyC light, $\log_{10}(\xi_{ion}/(Hz \, erg^{-1}))$ is the production efficiency of ionizing radiation in Hz erg^{-1} and ρ_{UV} is the comoving UV luminosity density in erg s^{-1} Hz^{-1} Mpc^{-3}.

The production efficiency of ionizing radiation $\log_{10}(\xi_{ion}/(Hz \, erg^{-1}))$ is a key determinant of a galaxy's potential contribution to reionisation; describing how much of the integrated spectrum is in the LyC region relative to the non ionizing UV component which represents the young stellar population. Spectroscopically it is determined by the Balmer

line luminosities (H_α, H_β or H_γ) line luminosities in ratio with the luminosity of a non-ionizing UV wavelength such as 1 500 Å $\log_{10}(\xi_{ion}/(Hz\ erg^{-1})) = \frac{N_{H^0}}{L_{UV} \times c_{rec}}$ where the N_{H^0} uses the Leitherer & Heckman (1995) conversion $N(H^0)[s^{-1}] = \frac{1}{1.36} \times 10^{12}L(H_\alpha)[erg\ s^{-1}]$ from luminosity to a production rate of ionizing photons, assuming that the recombination and ionisation of the nebula are balanced. $c_{rec} = 2.89$ refers to the case B recombination constant (Osterbrock & Ferland 2006) for hydrogen that allows use of either the H_α or H_β lines depending on what is available observationally. It should be noted that a dust correction is to be applied to the L_{UV} as this is the intrinsic value. Photometrically however, we instead use equation 2 from Wilkins et al. (2016):

$$(\xi_{ion}/(Hz\ erg^{-1})) = \int_\infty^{c/912} L_\nu(h\nu)^{-1}/L_\nu(1\ 500)d\nu \qquad (2)$$

where c is the speed of light in Angstroms/s and h is the Planck constant in erg Hz^{-1}. This essentially divides the Lyman Continuum region luminosity density by a non ionizing luminosity constant, indicating the relative production of ionizing radiation. This is a value strongly dependent on the stellar population synthesis model used requiring a consistent method for meaningful comparison, so we explore only the MAGPHYS SED code for all samples. The unattenuated (dust corrected) luminosity is used for this calculation as it should reflect the intrinsic production efficiency.

3.3 UV slope β

The UV slope (β where $f_\lambda \propto \lambda^\beta$) is defined as the slope of the SED between rest frame UV wavelengths selected to exclude the LyC /Lyman-α features and to set an upper limit still considered to be in the UV region. It is commonly used as an indicator of dust attenuation and the age of the stellar population (Wilkins et al. 2013; Williams et al. 2018; Reddy et al. 2018; Nanayakkara et al. 2022) which are particularly important for the analogue analysis in this work.

We further define two versions of the UV slope β parameter that reflect each the blueness of the direct (not dust corrected) photometry β_P and that of the model derived attenuated SED β_A between rest wavelengths 1 300–2 600 (the orange highlighted region of Fig. 2) (see Calzetti et al. (1994) for a discussion of the chosen wavelength range). These are determined using the scipy.optimize.curve_fit package and under the constraint that at least 3 filters between this wavelength range have a non-zero flux.

The direct photometry is less commonly considered in these analyses (Rogers et al. 2013) as the model dependence of the redshift determination creates a degeneracy with the star formation history model parameters that determine the attenuated UV fit. However, our data includes both spectroscopic redshifts and highly accurate photometric redshifts ($<2\%$ errors), therefore we can estimate the attenuated UV slope while considering the errors in the observed photometry without relying on the underlying models. A discussion on the value of this parameter in contrast to the model derived attenuated slope can be found in Dunlop et al. (2013). Fig. 3 highlights the key differences between these.

The depth of the CEERS data as well as the limited photometric sampling could significantly distort the value derived from the direct photometric method, however, the model slope is not a solution to this. SED models are typically optimised for typical star forming galaxies at $z < 6$ and hence the modelled UV
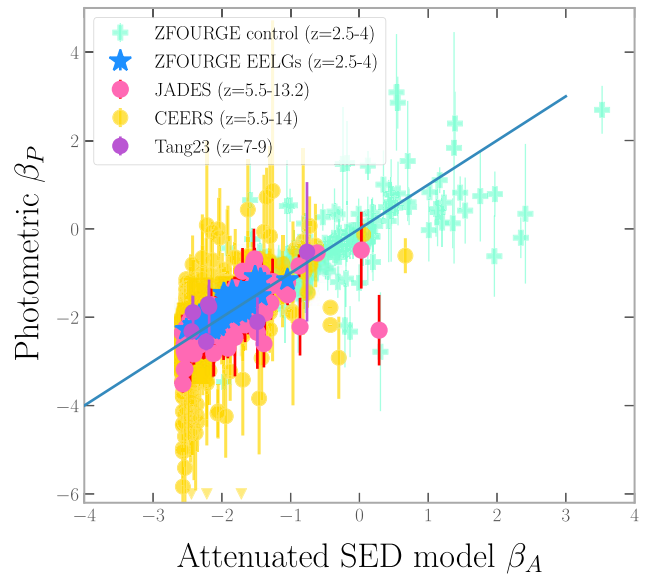

Figure 3. Attenuated model (β_a vs direct photometry β_p UV slopes for each subsample, with the ZFOURGE EELGs as blue stars, ZFOURGE control in cyan pluses, and the EoR samples in pink (JADES), gold (CEERS) and purple (Tang23) circles. The 1 to 1 line is shown in blue. General agreement is found with the exception of the CEERS sample around the $\beta_p < -2.6$ region, where the model attempts to constrain the relative number of blue stars that could have formed in the galaxy's lifetime, limiting the final 'blueness' of the slope.

slope may not be representative of the observed photometry for extreme cases. The model-independent UV slope is robust as it is estimated directly from the photometry and reflects the associated errors. Thus, it provides a consistent, comparative tool between the analogues and EoR galaxies.

3.4 Mass correction in EELGs [OIII]5007 contamination

The [OIII]5007 emission line flux poses as a significant contaminant to the photometric filters in which it resides if not taken into consideration (Forrest et al. 2018). The filter containing the emission line adds this to the total measured continuum flux, which is then read by the SED fitting code to determine physical parameters. Significantly bright emission lines raise the measured continuum flux above the real continuum, resulting in a flux excess and causing the physical parameters associated with that portion of the SED to be distorted if not correctly accounted for, as is the case with the current version of MAGPHYS.

Our sample selection is based on the [OIII]5007 EW, though we entirely remove the [OIII]5007 contaminated filters while fitting the SED of galaxies with MAGPHYS. Fig. 2 shows the effects of this on an example SED while Fig. 4 shows the effect of contamination on the stellar mass of the ZFOURGE galaxies and the EELG subsample. We note that EELG galaxies around $z \sim 3.5$ are most significantly affected by this. The mass determination is dependent largely on the optical region of the SED and the heightened [OIII]5007 flux in this subsample is expected to contribute to the four filters in the K-band ($\sim$2 micron) more-so than in most galaxies. The inclusion of emission lines in the SED fitting models themselves is another way of accounting for this effect which is employed in other SED fitting codes (e.g. Chevallard & Charlot 2016). The same removal method was applied to the JADES sample but not the CEERS sample, as its further limited
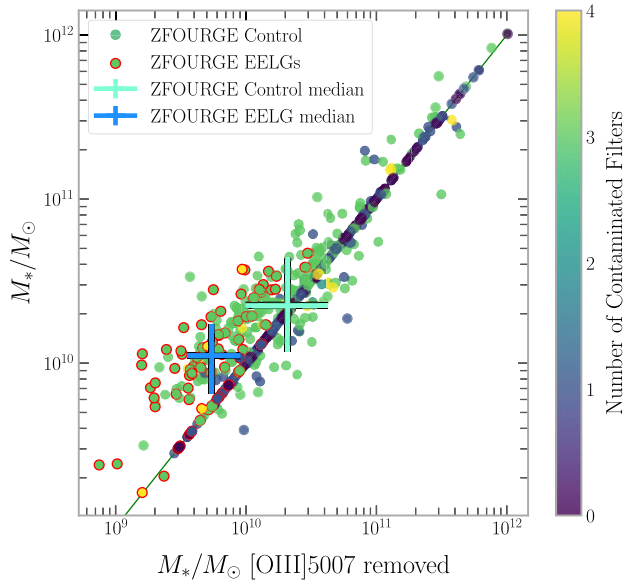
Figure 4. Stellar Mass with and without inclusion of the [OIII]5007 contaminated filters for the ZFOURGE control and ZFOURGE EELG samples. We also note a trend that $2.9 < z < 3.5$ galaxies tend to be more affected by the contamination and our EELG subsample is on the higher end of the discrepancy. This is due to the tendency for the [OIII]5007 line to fall into multiple filters at this redshift, and this is represented by the colour bar.

Figure 5. Specific star formation rate vs Stellar Mass plot for each of our samples. Colour-scheme is the same as Fig. 3 The median values and interquartile ranges are indicated by the coloured errorbars with orange:CEERS, pink:JADES, dark blue :ZFOURGE EELGs, teal:ZFOURGE control. The upper boundary is caused by the chosen SFR timescale. Popesso et al. (2022) z-sSFR-M_* relation selected at $z \sim 3.4$ and $M_*/M_\odot = 10^{10}$ shown as a red cross, and Davidzon et al. (2018) at $z = 2.22$ and $z = 6$ shown as a blue and green diamonds for comparison.

photometric coverage would significantly reduce the model reliability. While this may result in a moderately overestimated mass in the CEERS sample due to the expected high average [OIII]5007 EW, this alone would not account for the extreme values of some of these galaxies which require better photometric constraints (see Section 4.1 for further discussion).

4. Results and discussion

The physical parameters relating to the stellar population and dust attenuation are analysed in the ZFOURGE EELG sample and compared to the control samples within the EoR (JADES, CEERS, Tang23) and at similar redshift (ZFOURGE control). The comparisons selected explore the consistency with which the $2.5 < z < 4$ [OIII]5007 analogues match the physical parameters and internal processes of EoR galaxies.

4.1 Main sequence

We find that the ZFOURGE EELG subsample is consistently on the lower mass, higher specific star formation rate end of this parameter space when compared to the ZFOURGE control sample (Fig. 5). While ZFOURGE EELG stellar masses are generally an order of magnitude above the EoR counterparts, the sSFR median and 25–75 quartile region is comparable in the ZFOURGE EELG subsample and the JWST survey samples between redshifts 5.5–14 (Table 1). Our ZFOURGE control sample at $z \sim 3.4$ lies close to the stellar mass-redshift-sSFR relation developed by Popesso et al. (2022) using 27 other studies between $0 < z < 6$ and $8.5 < \log_{10}(M_*) < 11.5$. The sSFR of the ZFOURGE EELG sample, however, is 0.2 dex above this relation (Popesso et al. 2019; Leslie et al. 2020) and more closely related to the semi-empirical samples at $z \sim 6$ (Grazian et al. 2015; Davidzon et al. 2018). This reflects the strength of the [OIII]5007 EW selection technique at low redshift in collating the most highly star forming galaxies.
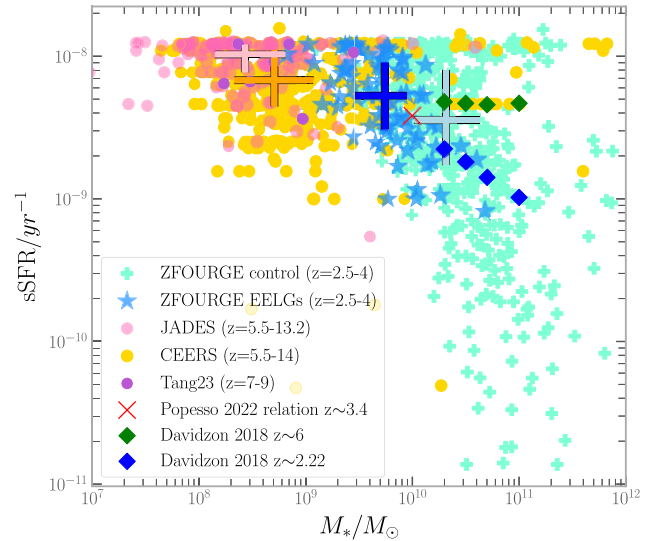
We note the appearance of striations visible in the CEERS sample. This is due to the relatively lower SNR and fewer photometric filters across the spectrum; preventing the code from fitting models to the more intricate variations in the continuum and thus recycling similar models. This effect continues through the other physical parameters to which the CEERS sample is fit and should be considered accordingly.

The stellar mass and star formation rates of both the ZFOURGE control and EELG samples are uniformly higher than that of the EoR samples, with the EELG sample being slightly less massive and less star forming. While the ZFOURGE EELG mass is higher than the EoR samples, it is still ~25% less massive than the average population at $z \sim 3$. The emission of LyC radiation is also not directly dependent on this parameter, so the functionality as an analogue is maintained. The heightened SFR is related to the emission, however, when normalizing by mass (sSFR) we will see that the discrepancy between EELG analogues and EoR galaxies is minimised.

4.2 Dust extinction star formation relation

Dust and star formation are inextricably linked galactic processes. Dust cools and catalyses molecular gas collapse while stellar death synthesizes more dust and transports it through the galaxy (Popping et al. 2017). The correlation of these properties across different epochs could be indicative of their evolving relationship as galaxies age and their dust content increases. Studies such as Zahid et al. (2012) link the correlation between A_V and SFR to the processes that quench star formation. A positive Spearman rank correlation between the dust extinction and star formation rate is observed by Li et al. (2019) in nearby galaxies, while Zahid et al. (2012) found a mass dependence on the correlation. They found more massive galaxies to be positively correlated while finding an anticorrelation below 10^{10} M$_\odot$ when the quiescent high

Table 1. Parameter medians with 25–75 percentiles for the ZFOURGE control, ZFOURGE EELGs, JADES and CEERS samples.

	ZFOURGE	EELGs	JADES	CEERS
$\log_{10}(M_*/M_\odot)$	$10.32^{0.32}_{-0.29}$	$9.75^{0.20}_{-0.28}$	$8.43^{0.37}_{-0.28}$	$8.71^{+0.37}_{-0.36}$
$\log_{10}(\mathrm{sSFR}/yr^{-1})$	$-8.48^{+0.35}_{-0.37}$	$-8.28^{+0.24}_{-0.23}$	$-7.99^{+0.12}_{-0.07}$	$-8.17^{+0.21}_{-0.18}$
A_V	$0.57^{0.30}_{-0.20}$	$0.25^{0.12}_{-0.14}$	$0.22^{0.42}_{-0.16}$	$0.08^{0.09}_{-0.04}$
β_P	$-1.27^{0.31}_{-0.28}$	$-1.83^{0.21}_{-0.15}$	$-2.02^{0.38}_{-0.31}$	$-2.20^{0.50}_{-0.49}$
$\log_{10}(\xi_{ion}/(\mathrm{Hzerg}^{-1}))$	$25.14^{0.06}_{-0.04}$	$25.16^{0.06}_{-0.05}$	$25.13^{0.06}_{-0.05}$	$25.18^{0.06}_{-0.07}$

Table 2. Spearman rank correlations for each subsample (ZFOURGE control, ZFOURGE EELGs, JADES and CEERS) and the combined dataset (Total). Blank spaces reflect correlations above a significance factor of 0.05.

	ZFOURGE	EELGs	JADES	CEERS	Total
Correlation with β_P					
$\log_{10}(M_*/M_\odot)$	0.48	. . .	0.45	0.45	0.67
$\log_{10}(\mathrm{sSFR}/yr^{-1})$	−0.08	. . .	0.32	−0.12	−0.24
A_V	0.52	0.63	0.44	0.21	0.58
Correlation with $\log_{10}(\xi_{ion}/(Hz\,\mathrm{erg}^{-1}))$					
$\log_{10}(M_*/M_\odot)$	−0.48	−0.36	−0.24	−0.36	−0.32
$\log_{10}(\mathrm{sSFR}/yr^{-1})$	0.73	0.57	. . .	0.46	0.55
A_V	0.32	0.53	. . .	. . .	. . .
β_P	−0.18	. . .	−0.42	−0.44	−0.33
Correlation with A_V					
$\log_{10}(\mathrm{SFR}/M_\odot/\mathrm{yr})$	0.70	0.48	0.56	0.45	0.68
$\log_{10}(\mathrm{sSFR}/yr^{-1})$	0.39	0.59	0.51	0.29	0.16

Figure 6. Star formation rate vs Dust attenuation plot for each of our samples. Marker sizes correlate to the Stellar Mass of the galaxy, with larger markers representing more massive galaxies consistent across all samples. Colour scheme represents the same samples as Fig. 3. median values and interquartile ranges are indicated by the coloured errorbars following the same colour-scheme as Fig. 5.

mass sample was removed, indicative of differing internal processes governing this relation between the low and high mass samples. Interestingly, the Balmer decrement (A_V tracer) stellar mass relation shows minimal evolution with redshift for $z < 2$ sources (Battisti et al. 2022).

For galaxies to be analogues of EoR galaxies, they should exhibit similar correlations across various physical parameters, owing to similar internal physical environments. Here we compare our findings across different epochs with our self-consistent, single SED fitting method to determine if the [OIII]5007 selected ZFOURGE EELG sample has a similar relationship between these properties to EoR galaxies (see Fig. 6). We find that ZFOURGE EELGs have a consistently lower dust extinction value when compared to the ZFOURGE control sample Table 1. The median value is half that of the ZFOURGE control and is well within the 25–75th percentiles of the EoR JADES sample.

We find a positive correlation between SFR and dust attenuation across all our samples. The correlation appears strongest in the ZFOURGE control sample, which tend to have higher A_V and SFR than the others (See Table 2). This finding is similar to Sakurai et al. (2013) who also find a stronger correlation between A_V and SFR for galaxies with SFR $> 20\,\mathrm{M}_\odot/\mathrm{yr}$. However, the exact values of our correlations may be biased by the intrinsic scatter and parameter ranges. For example, the weaker correlations observed in ZFOURGE EELGs and EoR (JADES, CEERS) samples could be due to the lack of dusty, low star-forming galaxies in the samples.
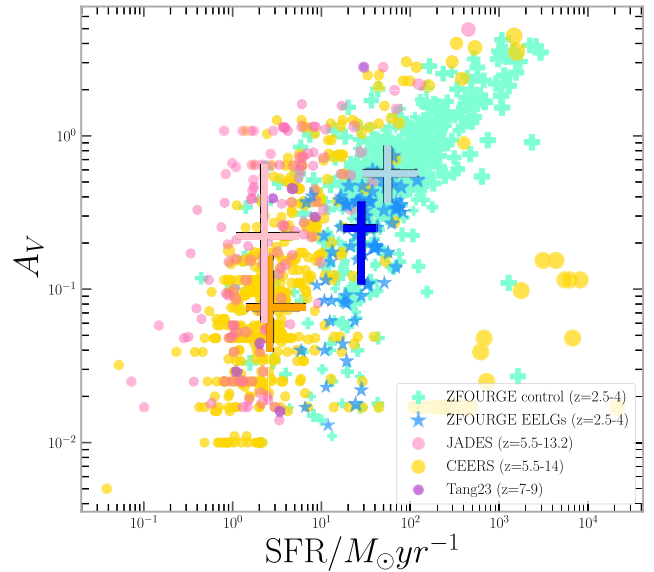
Even after accounting for stellar mass by comparing sSFR and A_V, we still find significant positive correlations across the four samples. The weaker correlations in the ZFOURGE and CEERS samples could be due to their relatively large parameter ranges compared to the ZFOURGE EELG and JADES samples. The similar correlation in the EELG and JADES samples in contrast to the ZFOURGE control suggests that the physical characteristics of ZFOURGE EELGs are more comparable to the $z > 5.5$ JADES sample than to the control $z \sim 3.4$ sample.

4.3 UV slope (β_p) relations

We compare the different samples parameter space relations in Fig. 7 and identify the correlations between each subsample in Table 2. The UV slope interquartile range of the ZFOURGE EELGs matches with the upper quartile of the JADES sample and is well below the ZFOURGE control (Fig. 7 top panel). This identification of the ZFOURGE EELG sample as being similarly blue and dust free as the direct EoR samples while having a higher specific star formation rate than the ZFOURGE control sample (see Table 1) indicates a strong potential for these to be good EoR analogues.

The β_P vs M_* relation (Fig. 8) appears to be discontinuous between the low and high redshift samples. The observed correlation between stellar mass and slope agrees with the conclusions of Pannella et al. (2009) and with the consensus that increasing stellar mass correlates with increasing dust attenuation and therefore a redder (observed) slope (Bouwens et al. 2016). This is supported by the lower dust extinction (Table 1) of the high redshift samples, as well as explains the tight parameter space occupancy of the ZFOURGE EELG sample in this figure. The correlation between UV slope and stellar mass is not present in the ZFOURGE EELG sample (Table 2) as it occupies a very constrained portion of the parameter space compared to the other samples (1.8 dex in stellar mass for ZFOURGE EELGs vs 3.6 dex for JADES for example). The UV slope appears to be sensitive to the mass with a strong
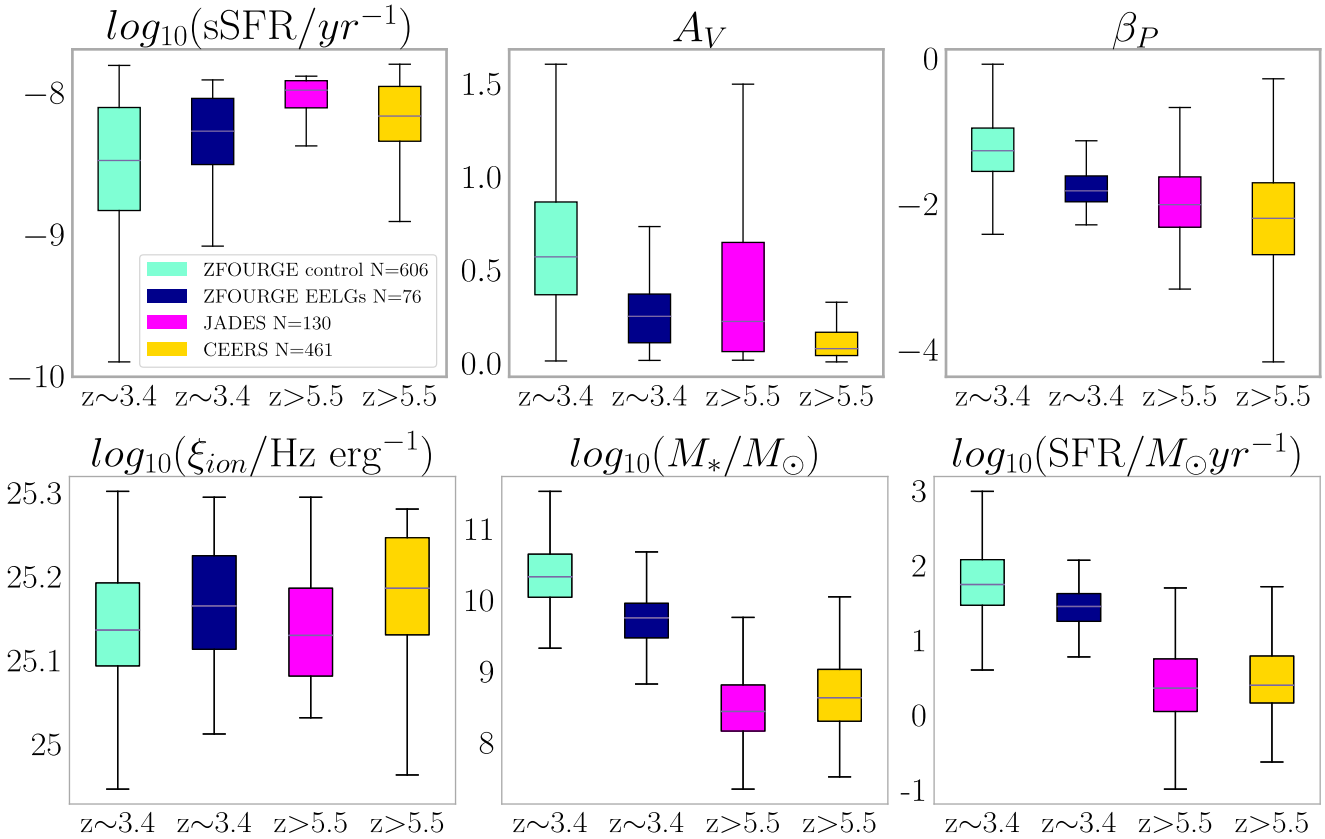
Figure 7. MAGPHYS derived physical parameters for each of the samples following the same colour scheme as previous figures. Median values and correlations can be found in Tables 1, 2.

overall correlation of 0.67 and strong correlations with the EoR and control samples. This suggests that the ZFOURGE EELG sample is consistently blue despite the mass being an order of magnitude greater than the EoR JADES median.

The correlation between the UV slope and dust extinction is readily identifiable, being strongest in the ZFOURGE control and EELG samples, which is similar to the conclusion drawn in Wilkins et al. (2013). The parameter space occupied by the ZFOURGE EELGs is the same as that of the JADES and Tang23 samples which could indicate the 'blueness' of a galaxy being attributed to the absence of attenuating dust in the model, though better dust constraints are required.

We also note the negative correlation of the β_P with ξ_{ion} plot (Fig. 8). The ξ_{ion} does not evolve significantly between the different redshift samples, however it does appear to negatively correlate with UV slope in both the EoR samples. This suggests that at least for the EoR samples, bluer galaxies tend to have a higher ionizing to non ionizing photon ratio though further investigation of the model dependence of the ξ_{ion} parameter are necessary.

4.4 Ionizing photon production

This section explores the correlations between the ionizing photon production efficiency ξ_{ion} and the physical parameters determined

by our model (Fig. 8). We refer to Table 2 for the correlation coefficients and their significance.

The ξ_{ion} anticorrelates with stellar mass in the low redshift samples (−0.48 ZFOURGE control, −0.36 ZFOURGE EELGs), suggesting that a lower mass corresponds to a higher production efficiency, though this is not clear from its visual appearance (Fig. 8). This weak dependence on stellar mass agrees with the findings of Emami et al. (2020) and Shivaei et al. (2018) for this mass range (10^7–10^{11}M$_\odot$) at a similar redshift range to our study, as well as with Lam et al. (2019) for the EoR samples. This weak anticorrelation with stellar mass potentially implicates smaller systems at early epochs as the powerhouses of ionizing photon production.

The production efficiency correlates most strongly with the sSFR in the ZFOURGE control and its EELG subsample. The assertion that a high specific star formation rate would be reflected in higher LyC emission is known (Castellano et al. 2023) between $2 < z < 5$. While a significant correlation could not be determined for the JADES sample due to the small dynamical range of sSFR values, the CEERS sample does reveal a similarly strong relationship in this parameter space.

A notable correlation between ξ_{ion} and dust extinction is found only in the ZFOURGE EELG sample. The result suggests a relationship between the dust within [OIII]5007 selected galaxies and their efficiency of LyC photon production. This is likely due to the
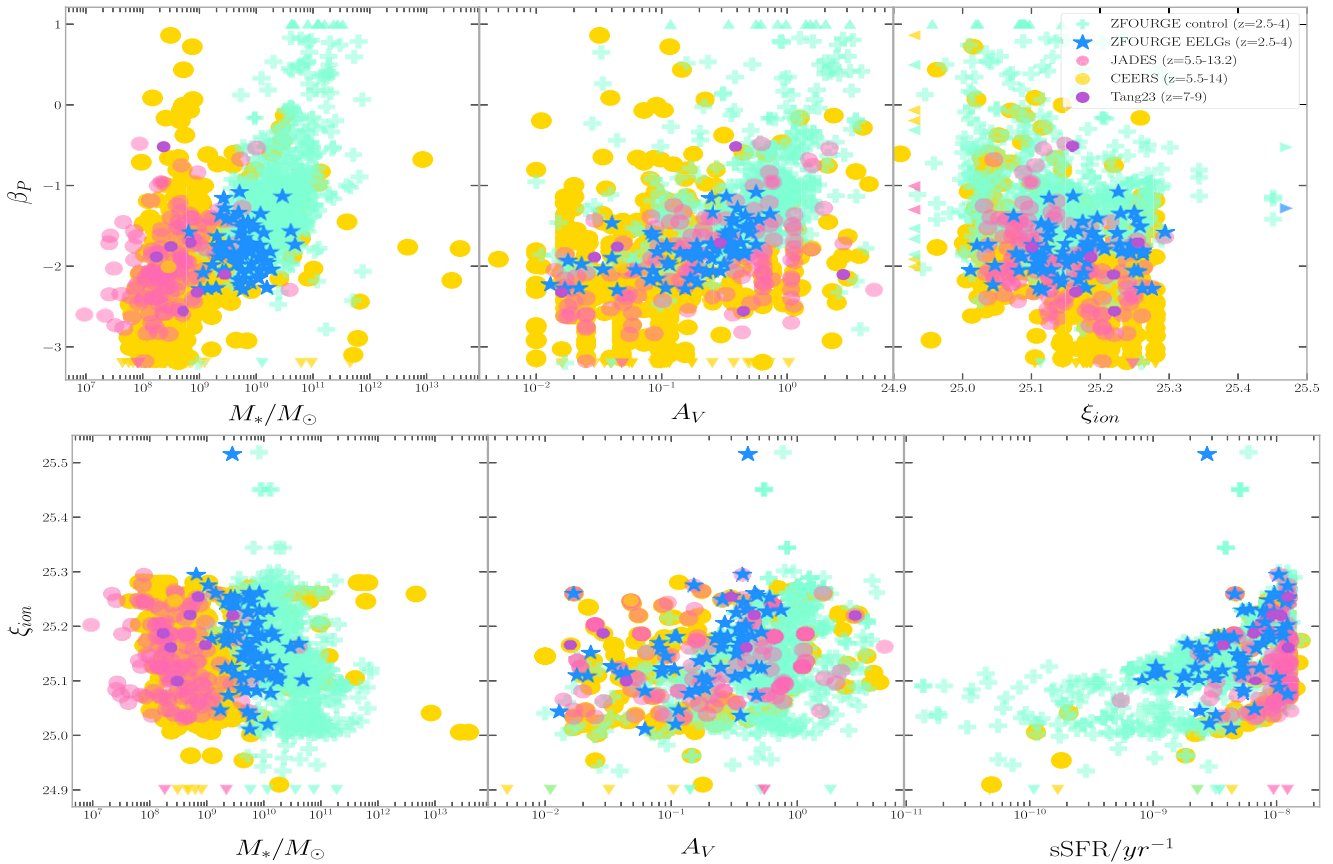
Figure 8. MAGPHYS derived physical parameters vs β_P (uncorrected for dust, top panel) vs photometric $\log_{10}(\xi_{ion}/(Hz\ erg^{-1}))$ (dust corrected Charlot & Fall 2000, bottom panel) for each of the samples following the same colour scheme as previous figures. Median values and correlations can be found in Tables 1, 2. Model limitations manifest as 'walls' in ξ_{ion} and sSFR.

underlying degeneracy correlating A_V, age, and SFR. Young galaxies with emerging O star populations haven't seeded as much dust when compared to their older counterparts, and it is these massive stars that are the driving force for the strong [OIII]5007 emission lines in the ZFOURGE EELG sample.

5. Summary and conclusion

This paper uses SED modelling to derive the physical parameters of EELGs with analogous [OIII]5007 emissions to EoR galaxies at $z \sim 3$ and compares them to both a control sample as well as the EoR galaxies observed at $z > 5.5$ with JWST. In this section we summarize the findings of our analysis and the potential avenues further research can take.

1. The combined high sSFR (-8.28 yr^{-1}), comparatively low stellar mass (9.75 M$_\odot$), low dust extinction ($A_V = 0.25$) and blue UV slope ($\beta_P = -1.84$) suggest that the [OIII]5007 selection technique is a good analogue selector. The high SFR and low dust are critical for their analogue status as they must produce copious ionizing radiation that is not attenuated by dust. We need to further investigate the escape of LyC light by studying the f_{esc} to complete the picture.

2. We confirm the strength of the [OIII]5007 selection technique for finding EoR analogues with lower stellar mass, high relative star formation rates for our sample between $2.5 < z < 4$. These galaxies are an order of magnitude higher in mass than their EoR counterparts but maintain similar sSFR despite this.

3. We find that the [OIII]5007 selected EELGs have a more similar correlation between their dust and star formation parameters to EoR galaxies than to the control sample between $2.5 < z < 4$, suggestive of similar internal processes relating to the dust distribution and star formation between the analogues and EoR galaxies.

4. We find a correlation between the UV slope and dust extinction/attenuation which is strongest in our low redshift samples. We also see a similar correlation with the stellar mass, particularly in our control sample and the CEERS data. Experiencing similar attenuation between these samples indicates low dust obscuration however this is only supported by UV photometry. Further studies using deep ALMA observations are required to determine if the steep UV slopes are due to the stellar population or an absence of dust.

5. We find that the median β_P of our ZFOURGE EELGs matches more closely with the higher redshift counterparts

than with the $2.5 < z < 4$ control sample, suggesting that our selection technique has recovered galaxies with a similar starbursty nature as the EoR galaxies.

6. We find the most significant correlation of $\log_{10}(\xi_{ion}/(Hz\ erg^{-1}))$ to be with its sSFR which coincides with the findings of Castellano et al. (2023) for bright $M_U V \leq 20$, $2 < z < 5$ galaxies. This correlation appears the strongest with our low redshift samples, likely due to the wider parameter space.

Data Availability. All data and software mentioned is publicly available.

References

Anderson, L., Governato, F., Karcher, M., Quinn, T., & Wadsley, J. 2017, MNRAS, 468, 4077

Bagley, M. B., et al. 2023, ApJL, 946, L12

Battisti, A. J., et al. 2022, MNRAS, 513, 4431

Berg, D. A., Erb, D. K., Henry, R. B. C., Skillman, E. D., & McQuinn, K. B. W. 2019, arXiv:1901.08160

Bouwens, R., et al. 2016, ApJ, 833, 72

Bouwens, R. J., et al. 2015, arXiv:1511.08504

Boylan-Kolchin, M. 2022, arXiv:2208.01611

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503

Bruton, S., Lin, Y.-H., Scarlata, C., & Hayes, M. J. 2023 arXiv:2303.03419

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000

Bunker, A. J., et al. 2023, arXiv:2302.07256

Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582

Cameron, A. J., et al. 2023, A&A, 677, A115

Cardamone, C., et al. 2009, MNRAS, 399, 1191

Castellano, M., et al. 2023, A&A, 675, A121

Chabrier, G. 2003, PASP, 115, 763

Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718

Chevallard, J., & Charlot, S. 2016, MNRAS, 462, 1415

da Cunha, E., Charlot, S., & Elbaz, D. 2008, arXiv:0806.1020

da Cunha, E., et al. 2015, arXiv:1504.04376

Davidzon, I., Ilbert, O., Faisst, A. L., Sparre, M., & Capak, P. L. 2018, ApJ, 852, 107

Donnan, C. T., et al. 2022, arXiv:2207.12356

Dunlop, J. S., et al. 2013, MNRAS, 432, 3520

Eisenstein, D. J., et al. 2023, arXiv:2306.02465

Emami, N., et al. 2020, ApJ, 895, 116

Ferland, G. J., et al. 1998, PASP, 110, 761

Finkelstein, S. L., et al. 2019, ApJ, 879, 36

Forrest, B., et al. 2018, arXiv:1807.03785

Grazian, A., et al. 2015, A&A, 575, A96

Gupta, A., et al. 2022, MNRAS, 519, 980

Hainline, K. N., et al. 2023, arXiv:2306.02468

Harikane, Y., et al. 2022, arXiv:2208.01612

Haro, P. A., et al. 2023a, arXiv:2303.15431

Haro, P. A., et al. 2023b, arXiv:2304.05378

Hayes, M., et al. 2011, ApJ, 730, 8

Hildebrand, R. H. 1983, qjras, 24, 267

Hsiao, T. Y.-Y., et al. 2023, arXiv:2305.03042

Iyer, K., & Gawiser, E. 2017, ApJ, 838, 127

Izotov, Y. I., et al. 2022, arXiv:2207.04483

Izotov, Y. I., et al. 2018, MNRAS, 478, 4851

Ji, Z., et al. 2020, ApJ, 888, 109

Kewley, L. J., et al. 2015, arXiv:1507.07932

Kriek, M., et al. 2009, ApJ, 700, 221

Kulkarni, G., Worseck, G., & Hennawi, J. F. 2018, arXiv:1807.09774

Labbe, I., et al. 2022, arXiv:2207.12446

Lam, D., et al. 2019, arXiv:1902.02786

Leitherer, C., & Heckman, T. M. 1995, ApJS, 96, 9

Leslie, S. K., et al. 2020, ApJ, 899, 58

Levesque, E. M., Kewley, L. J., & Larson, K. L. 2010, AJ, 139, 712

Li, H., et al. 2019, ApJ, 872, 63

Madau, P. 1995, ApJ, 441, 18

Madau, P., & Dickinson, M. 2014, Cosmic star-formation history, doi: 10.1146/annurev-astro-081811-125615

Madau, P., & Haardt, F. 2015, arXiv:1507.07678

Matthee, J., et al. 2023, ApJ, 950, 67

Meng, X., & Gnedin, O. Y. 2020, MNRAS, 494, 1263

Naidu, R. P., et al. 2019, arXiv:1907.13130

Naidu, R. P., et al. 2022, arXiv:2208.02794

Nanayakkara, T., et al. 2016, ApJ, 828, 21

Nanayakkara, T., et al. 2022, arXiv:2207.13860

Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei

Pannella, M., et al. 2009, ApJ, 698, L116

Paris, D., et al. 2023, arXiv:2301.02179

Popesso, P., et al. 2019, MNRAS, 490, 5285

Popesso, P., et al. 2022, arXiv:2203.10487

Popping, G., Somerville, R. S., & Galametz, M. 2017, MNRAS, 471, 3152

Raiter, A., Fosbury, R. A. E., & Teimoorinia, H. 2010, A&A, 510, A109

Reddy, N. A., et al. 2018, ApJ, 853, 56

Rieke, M. J., et al. 2023, arXiv:2306.02466

Robertson, B. E., Ellis, R. S., Dunlop, J. S., McLure, R. J., & Stark, D. P. 2010, Early star-forming galaxies and the reionization of the Universe, doi: 10.1038/nature09527

Robertson, B. E., et al. 2022, arXiv:2212.04480

Robertson, B. E., et al. 2023, NatAs, 7, 611

Rogers, A. B., McLure, R. J., & Dunlop, J. S. 2013, MNRAS, 429

Sakurai, A., Takeuchi, T. T., Yuan, F.-T., Buat, V., & Burgarella, D. 2013, Earth, Planets and Space, 65, 203

Schaerer, D., & de Barros, S. 2009, A&A, 502, 423

Shapley, A. E., Sanders, R. L., Reddy, N. A., Topping, M. W., & Brammer, G. B. 2023, ApJ, 954, 157

Shivaei, I., et al. 2018, ApJ, 855, 42

Simmonds, C., et al. 2023, arXiv:2303.07931

Sklias, P., et al. 2013, arXiv:1310.2655

Straatman, C. M. S., et al. 2016, ApJ, 830, 51

Tacchella, S., et al. 2023, arXiv:2302.07234

Takeuchi, T. T., Yuan, F.-T., Ikeyama, A., Murata, K. L., & Inoue, A. K. 2012, ApJ, 755, 144

Tang, M., Stark, D. P., Chevallard, J., & Charlot, S. 2019, MNRAS, 489, 2572

Tang, M., et al. 2023, arXiv:2301.07072

Torrey, P., et al. 2018, MNRAS, 477, L16

Tran, K.-V. H., et al. 2020, arXiv:2004.13079

Vincenzo, F., & Kobayashi, C. 2018, MNRAS, 478, 155

Wilkins, S. M., et al. 2013, arXiv:1302.1387

Wilkins, S. M., et al. 2016, MNRAS, 458

Williams, C. C., et al. 2018, ApJS, 236, 33

Yang, G., et al. 2020, arXiv:2011.08192

Yang, G., et al. 2023, arXiv:2307.14509

Yung, L. Y. A., Somerville, R. S., Finkelstein, S. L., Wilkins, S. M., & Gardner, J. P. 2023, arXiv:2304.04348

Zahid, H. J., Yates, R. M., Kewley, L. J., & Kudritzki, R.-P. 2012, arXiv:1211.7062

Zhai, Z., Benson, A., Wang, Y., Yepes, G., & Chuang, C.-H. 2019, MNRAS, 490, 3667