

Doctoral Dissertation

博士論文

**A comprehensive study of Emission-line Galaxies  
at Cosmic Noon**

(銀河形成最盛期の輝線銀河の研究)

A Dissertation Submitted for the Degree of Doctor of  
Philosophy  
December 2024

令和6年12月博士(理学)申請

Department of Astronomy, Graduate School of Science  
The University of Tokyo  
東京大学大学院理学系研究科天文学専攻

**Nuo Chen**

陳諾



# Abstract

In this thesis, we present a comprehensive study of a large population of emission-line galaxies at  $2.05 < z < 2.5$  in the rest-frame optical regime, with a particular focus on low-mass galaxies with stellar mass less than  $10^9 M_{\odot}$ . The study encompasses the entire process, from sample selection to the derivation of various physical properties. The primary aim is to investigate a representative population of star-forming galaxies, which contain strong emission lines, during the epoch of Cosmic Noon. Among our parent sample, the low-mass population has not yet been studied in detail and could provide insights into the mechanisms driving galaxy evolution and shaping what galaxies are like today.

Many high-redshift galaxies are having strong rest-frame optical emission lines. In contrast to the traditional spectroscopic method, we introduce a new selection method to search for emission line galaxies based on their flux excess in broadband data relative to the best-fit stellar continuum. More than 1,000  $H\alpha$  emission line galaxies ( $H\alpha$  emitters, HAEs) at  $2.05 < z < 2.5$  have been found in the ZFOURGE survey, in which  $H\alpha$  emission line are detected in the observed  $K_s$  broad-band filter. Using the same method, we also extract the strong diagnostic emission lines for these individual HAEs:  $[\text{OIII}]\lambda\lambda 4959, 5007$ ,  $[\text{OII}]\lambda\lambda 3726, 3729$ . This measurements demonstrate good consistency with those obtained from spectroscopic surveys.

We investigate the relationship between the equivalent widths (EWs) of these emission lines and various galaxy properties, including stellar mass, stellar age, star formation rate (SFR), specific SFR (sSFR), ionization states (O32). A discrepancy between HAEs at  $z \sim 2.3$  and typical local star-forming galaxy observed in the SDSS has been identified, suggesting the evolution of gas-phase metallicity ( $Z$ ) and ionization parameters ( $U$ ) where high-redshift galaxies have lower metallicity and higher ionization parameter.

Notably, we have observed a significant number of low-mass HAEs exhibiting strongly elevated  $\text{SFR}(H\alpha)$  above the star formation main sequence (SFMS) and exceptionally high  $\text{EW}_{[\text{OIII}]}$ . Their galaxy properties are comparable to those of extreme objects, such as extreme O3 emitters (O3Es) and Ly $\alpha$  emitters (LAEs) at  $z \simeq 2 - 3$ . Considering that these characteristics may indicate potential strong Lyman continuum (LyC) leakage, higher redshift analogs of the low-mass HAEs could be significant contributors to the cosmic reionization.

Thus, we study the ionizing properties of these low-mass HAEs, especially the ionizing photon production efficiencies ( $\xi_{\text{ion}}$ ). The low-mass HAEs have a median value of  $\log(\xi_{\text{ion}}/\text{Hz erg}^{-1}) = 25.24_{-0.13}^{+0.10}$ , higher than that of main sequence galaxies by  $\sim 0.2$  dex at similar redshift, indicating that the low-mass HAEs are more efficient in producing ionizing photons. We further explore the dependence of  $\xi_{\text{ion}}$  on other

galaxy properties, such as the UV spectral slope ( $\beta_{\text{UV}}$ ), the UV magnitude ( $M_{\text{UV}}$ ), the EWs of H $\alpha$  and [OIII] emission lines. Galaxies with the bluer UV slopes, fainter UV luminosities and higher equivalent widths exhibit elevated  $\xi_{\text{ion}}$  by a factor of  $\sim 2$  compared to the median  $\xi_{\text{ion}}$  of our sample.

To investigate and resolve the rest-frame optical structures of high-redshift galaxies in remarkable detail, we utilize the JWST-JADES data, which is an extremely deep survey overlapping with the ZFOURGE-CDFS field. The data enable us to perform a rest-frame optical, spatially resolved analysis on a subsample of the HAEs. An [OIII] emission-line map of each HAE is created from the flux excess in the F150W filter, leading to the discovery of a population of kiloparsec-scale compact emission line regions (“Green Seeds”) with high EW. Interestingly, some of them have extremely large  $\text{EW}_{[\text{OIII}]} > 1000\text{\AA}$ , suggesting the possible LyC leakage from these emission line regions. Embedded within the host galaxy, many Green Seeds correspond to UV star-forming clumps and HII regions, indicating elevated starburst activity in them, with sSFR several times higher than that of the host galaxy. Based on theoretical frameworks, Green Seeds are expected to be formed through gravitational disk instability and/or galaxy mergers. Considering the stellar masses of Green Seeds, we speculate that high-mass Green Seeds may migrate toward the galactic center to build the central bulge, while low-mass Green Seeds are easily disrupted and short-lived. Besides, we propose that some Green Seeds could be the progenitors of globular clusters or ultracompact dwarf galaxies observed in the local universe.

This research on “Green Seeds” using data from the JADES survey has demonstrated the feasibility of isolating strong optical emission line regions from flux excesses at  $z \sim 2$ . Still, a more comprehensive study on these clumpy structures is required to understand their relationship to host galaxies and to determine whether their properties vary with redshift or environment. Moreover, the galaxy sample in this study is primarily derived from pre-JWST era, whereas lower mass galaxies at “Cosmic Noon”, with masses down to  $10^7 M_{\odot}$ , can now be identified from JWST observations. These new data will revolutionize our understanding of galaxy evolution at high redshifts, particularly in the crucial low-mass regime.

# Acknowledgement

I would first like to express the deepest gratitude to my supervisor, Dr. Kentaro Motohara for the insightful and instructive comments, remarks and engagement through the learning process of this doctoral dissertation. The door to his office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed this thesis to be my own work, but steered me in the right direction. Besides, he has given me patient and constant support and taught me how to become a qualified researcher during my Ph.D. course.

I would also like to thank Dr. Lee Spitler from Macquarie University who was involved in the ZFOURGE survey. Through the collaboration with him, he provided me a lot of advises on scientific research. I am gratefully indebted to his for very valuable comments on this thesis. Without his participation and input, this study could not have been successfully conducted. He also host me for my three-month visit to Macquarie University in 2022 in Australia. I have an exciting experience in Sydney and Siding Spring Observatory.

I would also like to express my gratitude to Dr. Matthew A. Malkan. During my Ph.D. course, I had the opportunity to visit UCLA twice, each time for about one month, where I was warmly hosted by him. Without his great help, the JWST-related studies could be fluently carried out.

I would also like to acknowledge Dr. Masahiro Konishi from the SWIMS team, who supported me a lot on the computer programming during the imaging processing. In virtue of his great reduction pipeline, I am able to reduce SWIMS images smoothly and perfectly.

I would like to express my special thanks to several experts in extragalactic astronomy, Dr. Tadayuki Kodama, Dr. Kenichi Tadaki, Dr. Rhythm Shimakawa, Dr. Kashino Daichi and Dr. Kimihiko Nakajima. They gave me many advises on this study and provided me with the related catalogs.

During my Ph.D. course, I have experienced pleasure time with all of the stuffs and student members of the SWIMS team, namely, Dr. Hidenori Takahashi, Dr. Shuhei Koyama, Natsuko Kato, Dr. Yasunori Terao, Dr. Kosuke Kushibiki, Hiroki Nakamura, Shogo Homan, Akino Yasuda, Jeung Yun, Tomoya Yukino, Misato Fujii, Kensho Tanaka. They have made these five years an invaluable period in my life.

Finally, I must express my very profound gratitude to my parents for providing me with unflinching support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.



# Contents

<b>1</b>	<b>INTRODUCTION</b>	<b>15</b>
1.1	Galaxy formation and evolution scenarios . . . . .	15
1.2	Cosmic Noon and Star-forming galaxies in this epoch . . . . .	18
1.2.1	Cosmic star-formation Histroy . . . . .	18
1.2.2	Star formation rate and its indicators . . . . .	18
1.2.3	The Star formation main sequence at $z \sim 2$ . . . . .	20
1.3	Emission line diagnostics and Extreme emission line galaxies . . . . .	21
1.3.1	Hydrogen Recombination Lines . . . . .	21
1.3.2	Forbidden lines . . . . .	22
1.3.3	Emission line diagnostics . . . . .	22
1.3.4	Extreme Emission Line Galaxies . . . . .	24
1.4	Detection of emission lines . . . . .	24
1.4.1	Spectroscopy . . . . .	24
1.4.2	Narrow-band imaging . . . . .	25
1.4.3	Broad/Medium band imaging with SED fitting . . . . .	26
1.5	Resolved study on galaxy properties . . . . .	27
1.6	Objectives and Structure of this dissertation . . . . .	28
<b>2</b>	<b>The Compilation of HAEs at <math>z \sim 2.3</math></b>	<b>31</b>
2.1	Observation and Data . . . . .	31
2.1.1	The ZFOURGE survey . . . . .	31
2.1.2	SWIMS medium K-band imaging . . . . .	31
2.2	Sample Selection . . . . .	32
2.3	SED fitting with emission line templates . . . . .	34
2.3.1	Stellar population models and Star formation history . . . . .	34
2.3.2	Nebular emission model . . . . .	34
2.3.3	Dust attenuation model . . . . .	35
2.4	Emission line measurement . . . . .	35
2.4.1	$H\alpha$ . . . . .	36
2.4.2	$[OIII]$ and $[OII]$ . . . . .	40
2.5	Rest-Frame Equivalent Widths . . . . .	41
2.6	Emission Line Ratios . . . . .	44
2.7	Luminosity Function . . . . .	44

<b>3</b>	<b>A population of EELGs at Cosmic Noon</b>	<b>49</b>
3.1	SFRs and Star Formation Main Sequence . . . . .	49
3.2	Bursty star formation in low-mass galaxies . . . . .	53
3.3	Equivalent widths in low-mass galaxies . . . . .	54
<b>4</b>	<b>Multiple Emission Line Analysis of the HAEs</b>	<b>57</b>
4.1	EW vs. Stellar properties . . . . .	57
4.1.1	[OIII] EWs . . . . .	57
4.1.2	[OII] EWs . . . . .	60
4.1.3	H $\alpha$ EWs . . . . .	64
4.2	EW vs. ISM properties . . . . .	64
4.3	The ionization parameters of low-mass HAEs . . . . .	67
<b>5</b>	<b>The Ionizing properties of the HAEs</b>	<b>73</b>
5.1	Measurement of $\xi_{\text{ion}}$ . . . . .	74
5.2	$\xi_{\text{ion}}$ evolution with galaxy properties . . . . .	74
5.2.1	$\xi_{\text{ion}}$ and stellar mass . . . . .	76
5.2.2	$\xi_{\text{ion}}$ and UV Properties . . . . .	76
5.2.3	$\xi_{\text{ion}}$ and nebular emission lines . . . . .	79
5.3	$\xi_{\text{ion}}$ Evolution with Redshift . . . . .	81
5.4	$\xi_{\text{ion}}$ in SPS models and Implications for Reionization . . . . .	84
<b>6</b>	<b>Resolved studies of the HAEs from JWST observations</b>	<b>89</b>
6.1	JWST observations in ZFOURGE field . . . . .	89
6.1.1	Astrometric correction . . . . .	89
6.1.2	The subsample of HAEs in JADES . . . . .	91
6.2	Resolved continuum and emission line map . . . . .	91
6.3	“Green Seeds” and “Red Seeds” . . . . .	96
6.4	The properties of Green Seeds . . . . .	100
6.4.1	Stellar mass . . . . .	100
6.4.2	Star formation rate . . . . .	101
6.4.3	Equivalent width . . . . .	103
6.5	Comparison between “Green Seeds” and other resolved structures . . . . .	105
6.5.1	Relevance between Green Seeds and HII regions . . . . .	105
6.5.2	Comparison with UV star-forming clumps . . . . .	106
6.5.3	Comparison with Local HII regions . . . . .	108
6.6	The origin and fate of Green Seeds . . . . .	111
6.6.1	Formation of Green Seeds . . . . .	111
6.6.2	Fate of Green Seeds . . . . .	115
<b>7</b>	<b>Conclusions</b>	<b>119</b>
7.1	Summary points . . . . .	119
7.2	Future Aspects . . . . .	122

<b>A</b>	<b>SWIMS Imaging in this study</b>	<b>125</b>
A.1	SWIMS imaging pipeline: SWSRED . . . . .	125
A.2	PSF matching on SWIMS reduced image . . . . .	126
A.3	Better constrained Photometric redshifts . . . . .	127
<b>B</b>	<b>SFR vs. <math>EW_{[\text{OIII}]}</math> of massive HAEs</b>	<b>129</b>
<b>C</b>	<b>The resolved structures of HAEs from JWST observations</b>	<b>131</b>
C.1	The Cutout images of remaining cross-matched HAEs . . . . .	131
C.2	The Continuum and Emission-line Map for the remaining HAEs with Green Seeds . . . . .	132
C.3	The remaining UV clumps in combination with Green Seeds . . . . .	132



# List of Tables

2.1	SWIMS medium K-band and observation in S18B . . . . .	32
2.2	Main criteria for selecting galaxies at $z \sim 2.3$ . . . . .	36
2.3	Parent sample and the number of emission line galaxies with different threshold . . . . .	39
2.4	The best-fit parameters of emission line luminosity functions at $z \sim 2.3$	46
4.1	Relationship between the Equivalent Widths and various Galaxy Properties . . . . .	58
5.1	Median $\xi_{\text{ion}}$ for HAEs at $z \sim 2.3$ separated into different bins of stellar masses, UV luminosities and UV-slopes. . . . .	78
5.2	Various SPS models and corresponding $\xi_{\text{ion}}$ . . . . .	86



# List of Figures

2.1	Rest-frame optical emission lines drop in Fourstar+SWIMS MB filters at $z \sim 2.3$ . . . . .	33
2.2	The SFMS of the parent ZFOURGE sample . . . . .	37
2.3	Four best-fit SED of HAEs . . . . .	38
2.4	H $\alpha$ Flux comparison between SED method and MOSDEF . . . . .	41
2.5	OIII Flux comparison between SED method and MOSDEF (3D-HST) . . . . .	42
2.6	OII Flux comparison between SED method and MOSDEF (3D-HST) . . . . .	43
2.7	The best-fit luminosity function of H $\alpha$ . . . . .	45
2.8	The best-fit luminosity function of [OIII] and [OII] . . . . .	47
3.1	The star formation main sequence of HAEs at $z \sim 2.3$ . . . . .	51
3.2	Comparison between SFR(H $\alpha$ ) and SFR(UV) of the HAEs . . . . .	52
3.3	The star formation main sequence of HAEs with mock galaxies . . . . .	53
3.4	The star formation history (SFH) of HAEs inferred from SED . . . . .	54
3.5	The ratio of SFR(H $\alpha$ ) and SFR(UV) as a function of stellar mass and $EW_{H\alpha}$ . . . . .	55
3.6	The distribution of H $\alpha$ and [OIII] EWs for HAEs . . . . .	56
4.1	Relationship between the $EW_{[OIII]}$ and stellar properties (stellar mass, age, SFR, sSFR) . . . . .	61
4.2	Relationship between the $EW_{[OII]}$ and stellar properties . . . . .	62
4.3	Relationship between the $EW_{H\alpha}$ and stellar properties . . . . .	63
4.4	Relationship between the emission line equivalent widths and O32 line ratio . . . . .	66
4.5	Residuals of $EW_{[OIII]}$ vs. residuals of SFR(H $\alpha$ ) . . . . .	68
4.6	Relation between [OIII]/[OII] and R <sub>23</sub> -index for a bunch of galaxies at $z \sim 2.3$ . . . . .	70
5.1	The ionizing photon production efficiency ( $\xi_{ion}$ ) of HAEs and its dependence on stellar mass . . . . .	75
5.2	$\xi_{ion}$ as a function of the UV slope and magnitude . . . . .	77
5.3	$\xi_{ion}$ as a function of the equivalent width of emission lines . . . . .	80
5.4	Evolution of estimated $\xi_{ion}$ with redshift from $1 < z < 9$ . . . . .	82
5.5	The relation between $\xi_{ion}$ and stellar mass at different redshifts. . . . .	86
5.6	The evolution of $\xi_{ion}$ with stellar age in various SPS models . . . . .	87
6.1	The footprints of the ZFOURGE survey and the JADES survey in the GOOD-South field . . . . .	90

6.2	JWST/NIRCam RGB cutout images with prominent green substructures . . . . .	92
6.3	The PSF-matching process of the JWST imaging . . . . .	93
6.4	The flowchart of our methodology to identify “Green Seeds” in a sample galaxy . . . . .	95
6.5	The classification of “Green Seeds” and “Red Seeds” . . . . .	98
6.6	Example SEDs of Green Seeds and Red Seeds . . . . .	99
6.7	The example of the stellar continuum map, OIII emission line map, and H $\alpha$ emission line map . . . . .	100
6.8	The stellar masses of “Green Seeds” and “Red Seeds” from SED fitting and the color-mass-light ratio . . . . .	102
6.9	The SFR and sSFR of “Green Seeds”, compared with their host galaxies	104
6.10	The distribution of [OIII] and H $\alpha$ EWs for Green Seeds . . . . .	105
6.11	Illustration of the UV clump-detection process for HAEs and example.	107
6.12	Multiple emission line analysis of “Green Seeds” and local HII regions	110
6.13	The profile shapes and non-parametric structure of the HAEs in our sample . . . . .	113
6.14	Distance of seeds to the galaxy center . . . . .	116
7.1	JWST MB filters and corresponding H $\alpha$ emission lines . . . . .	124
A.1	The flow chart of SWSRED . . . . .	126
A.2	SWIMS $K_2$ image before and after convolution . . . . .	127
A.3	An example of fitting $z_{\text{phot}}$ by EAZY before/after including SWIMS data . . . . .	128
B.1	Relationship between the $EW_{[\text{OIII}]}$ and SFR of the HAEs with stellar masses larger than $10^{9.5}M_{\odot}$ . . . . .	129
C.1	JWST/NIRCam RGB cutout images w/o prominent green substructures . . . . .	133
C.2	The stellar continuum map, [OIII] and H $\alpha$ emission line map of HAEs (rest) . . . . .	134
C.3	Continue of Figure C.2 . . . . .	135
C.4	The contrast maps and the UV star-forming clumps of HAEs (rest) .	136

# Chapter 1

## INTRODUCTION

The study of galaxies across cosmic time has been one of the most significant topics in astronomy, encompassing a broad range of subjects and physical processes. In this chapter, we begin by reviewing galaxy formation and evolution (§1.1) and the cosmic history of star formation along with star-forming galaxies (SFGs) during Cosmic Noon (§1.2). Next, we go through the astrophysics of interstellar medium (ISM) and discuss the interpretation of their emission-line spectra (§1.3). Following this, we highlight recent observations of strong emission-line galaxies, emphasizing the strengths and limitations of various observational methods (§1.4). We also point out the importance of resolved studies of emission lines within galaxies (§1.5). Finally, we outline the structure of this thesis and the cosmological parameters adopted in the final section (§1.6).

### 1.1 Galaxy formation and evolution scenarios

A galaxy is a massive, gravitationally bound system composed of stars, gas, dust and dark matter. Galaxies are fundamental building blocks of the universe, hosting processes that drive star formation, chemical enrichment, and cosmic evolution.

The last century witnessed remarkable progress in galaxy research. Edwin Hubble confirmed that galaxies are vast collections of stars beyond the Milky Way and developed the Hubble Sequence, categorizing galaxies into ellipticals, spirals and irregulars based on their morphology (Hubble 1936). In particular, Hubble proposed the iconic tuning fork diagram, where the main galaxy types are organized. In this classification, ellipticals are labeled from perfectly round (E0) to the most flattened (E7). Moving to the right, the tuning fork splits into two branches: spirals (S) and barred spirals (SB). This framework inspired subsequent studies of galaxy evolution along the sequence. Ellipticals are typically redder systems, composed of older stars, with little or no star formation and high stellar masses. In contrast, spirals are bluer, actively forming stars, contain larger fractions of cold gas, and host younger stellar populations (Roberts & Haynes 1994; Kennicutt 1998).

The 1970s marked another significant breakthrough in galaxy research with studies of galactic rotation curves, which uncovered the existence of large amounts of unseen mass. The groundbreaking discovery revealed that unseen mass, later termed dark matter, significantly influences galactic dynamics and hinted at its

role in galaxy formation (Rubin & Ford 1970; Rubin et al. 1980). In the 1980s, large redshift surveys like the CfA Redshift Survey (Huchra et al. 1983) mapped the three-dimensional (3D) spatial distribution of galaxies in the Universe, unveiling the largest structures in the universe: galaxy filaments (de Lapparent et al. 1986; Geller & Huchra 1989). These observations revealed that galaxies are not randomly distributed but align along a vast cosmic web interspersed with immense voids. A consistent picture of structure formation in the universe gradually emerged through the interplay of observational studies and theoretical models. Today, our understanding of the universe is rooted in the  $\Lambda$ -Cold Dark Matter ( $\Lambda$ CDM) model (e.g., Blumenthal et al. 1984; Carroll 2001; Peebles & Ratra 2003). The  $\Lambda$ CDM model provides the framework for understanding galaxy formation and evolution. Cosmological hydrodynamic simulations, built on this model, allow us to trace the evolutionary pathways of galaxies over cosmic time (e.g., Katz et al. 1992; Kauffmann et al. 1999; Spergel et al. 2003). In more detail, galaxies are thought to accrete gas from their surrounding environment, gradually converting it into stars. The cooling and condensation of neutral hydrogen, followed by its transformation into molecular hydrogen to fuel star formation, are fundamental processes driving galaxy evolution. Besides, galaxies primarily grow through the accretion of gas from the intergalactic medium (IGM), maintaining a dynamic balance among gas inflow, star formation, and the ejection or heating of gas driven by feedback processes.

Observational studies over the past century have predominantly focused on the nearby universe. On the other hand, observing high-redshift (high- $z$ ) galaxies, which is often referred to as "look-back" observations, could offer critical constraints on the aforementioned theoretical models. Since the 1990s, advancements in instrumentation have driven significant progress in our ability to observe the distant universe. The launch of the Hubble Space Telescope (HST), Spitzer space telescopes, along with the advent of ground-based 8–10 m-class telescopes such as Keck, Subaru, VLT, Gemini, has revolutionized observational astronomy. These developments have led to the discovery of high- $z$  galaxies through various techniques, including the "dropout" method for identifying Lyman-break galaxies at various redshifts (e.g., Steidel et al. 1996, 2003; Burgarella et al. 2011) and narrow-band (NB) imaging searching for detecting Lyman-alpha ( $\text{Ly}\alpha$ ) emitters (LAEs; Cowie & Hu 1998; Rhoads et al. 2000; Ouchi et al. 2003). Later, large observational campaigns have conducted deep imaging surveys in multiple sky fields, such as the Great Observatories Origins Deep Survey (GOODS; Giavalisco et al. 2004), the Hubble Ultra Deep Field (UDF; Beckwith et al. 2006), the COSMOS field (Scoville et al. 2007), the CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011) primarily by HST observations, and the Subaru Deep Field (SDF; Maihara et al. 2001), the UKIDSS Ultra-Deep Survey (UDS; Lawrence et al. 2007) by ground-based telescopes. These efforts have revealed a substantial population of high- $z$  galaxies. Notably, a remarkably luminous galaxy at  $z = 11.1$  was suggested using Hubble WFC3/IR slitless grism spectroscopy, making it the highest redshift object identified until 2021 (Oesch et al. 2016). Compared to local galaxies, these high- $z$  galaxies differ significantly in their morphology, star formation activity, molecular gas fraction and many other physical properties, reflecting the evolving nature of the universe. Analysis of rest-frame optical structures of galaxies reveals a high prevalence of morphologically peculiar galaxies at  $z > 2$ ,

as well as smaller and more compact structures (e.g., [Conselice et al. 2005, 2008; Conselice 2014](#)). Typical SFGs at  $z \sim 2$  were forming stars and assembling their stellar mass  $\sim 10\times$  faster than their local counterparts ([Speagle et al. 2014; Fang et al. 2018](#)), fueled by  $\sim 10\times$  larger cold molecular gas reservoirs ([Combes 2018; Tacconi et al. 2020](#)). In a resolved view, high- $z$  SFGs often contain large clumps of star formation within their disks (e.g., [Conselice et al. 2004; Elmegreen et al. 2007; Förster Schreiber et al. 2011a](#)). These results align with some cosmological galaxy formation models (e.g., [Steinmetz & Navarro 2002; Erb 2008; Bournaud et al. 2014](#)) while showing inconsistencies with others (e.g., [Maller et al. 2006; Hopkins et al. 2012](#)).

Excitingly, the launch of the 6.5m James Webb Space Telescope (JWST) in December 2021 has opened a brand-new era for studying galaxy formation and evolution at very high redshift. JWST was designed to peer farther into the Universe than any other telescopes, reaching  $\sim 100$  times more sensitive to faint, distant objects than HST. We first confirmed for the first time galaxies at  $z = 14.2$  by spectroscopy ([Carniani et al. 2024](#)). Using the dropout technique, large numbers of galaxy candidates at  $z > 6$  have been identified (e.g., [Castellano et al. 2022; Harikane et al. 2023; Labbé et al. 2023](#)). These groundbreaking JWST observations have revealed more bright galaxies in the early universe than previously predicted by theoretical models. This discrepancy raises a fundamental question: either the stellar mass densities observed are challenging to reconcile within the framework of standard  $\Lambda$ CDM models, or the astrophysical models for galaxy and star evolution is behaving very differently from those in the local universe.

Besides the number counts, detailed research on the properties of first-generation galaxies is also being actively pursued based on these observations. Around 300-800 Myr after the Big Bang ( $6 < z < 15$ ), the universe experienced a major phase transition from neutral Hydrogen to ionized plasma, known as the Epoch of Reionization (EoR). The ubiquitous neutral IGM was ionized by the enormous amounts of ionizing photons (also called LyC or Lyman continuum) from the first luminous sources at that epoch. Current understanding suggests that reionization happens at  $z > 6$  (e.g., [Pentericci et al. 2014; Schenker et al. 2014](#)) but several important aspects of the EoR still remain unclear, such as identifying the specific sources drive reionization ([Madau & Haardt 2015; Robertson et al. 2015](#)). With JWST data, it has become possible to directly identify the main sources driving cosmic reionization. For instance, [Atek et al. \(2024\)](#) suggested that the majority of photons responsible for cosmic reionization originated from dwarf galaxies based on the direct measurements of these galaxy properties. Additionally, JWST has measured spectral properties and provided critical insights into reionization and the evolution of stellar populations ([Robertson et al. 2023; Kashino et al. 2023](#)). Due to the limited wavelength coverage and image depth, HST has not fully captured the rest-frame optical and near-infrared (NIR) light of the galaxies at such high redshift. JWST has also detected mature structures like disks and evidence of chemical enrichment in galaxies during the EoR ([Tacchella et al. 2023](#)). These findings are prompting significant revisions to theoretical frameworks, particularly in understanding early star formation and the growth of structure during EoR.

Following the launch of JWST, significant efforts have been dedicated to study-

ing the aforementioned first-generation galaxies at very high redshifts. While, at intermediate redshifts ( $1 < z < 3$ ), although many efforts were made to investigate this redshift range prior to JWST, several unsolved issues remain. For instance, due to instrumental limitations, pre-JWST studies were mostly focusing on massive galaxies ( $M_* > 10^9 M_\odot$ ) at  $z \sim 2$ . As a result, the evolution and role of low-mass galaxies at  $z \sim 2$ , which are the most numerous galaxies in the universe, have not been explored in detail. These low-mass galaxies are believed to dominate the ionizing photon budget and play a crucial role in the metal enrichment of the IGM (e.g., Robertson et al. 2010; Lin et al. 2023). With the advent of JWST data, it is now possible to study these low-mass galaxies at  $z \sim 2$  in greater detail and investigate how they evolve with cosmic time.

## 1.2 Cosmic Noon and Star-forming galaxies in this epoch

### 1.2.1 Cosmic star-formation History

The cosmic history of star formation is one of the most fundamental aspects of observational cosmology. After the ‘‘Cosmic Dawn’’ ( $z > 10$ ) ended, galaxies continued to grow through gravity, which were attracting more dark matter, gas, and forming stars in them. Over the past two decades, numerous studies have measured the cosmic star-formation rate density (SFRD) across a wide range of redshifts. Assuming a universal initial mass function (IMF), Madau & Dickinson (2014) concluded an avalanche of observational data and led to a fairly robust outline of the evolution of the star formation activity of galaxies from  $z \sim 8$  to  $z \sim 0$ . It is concluded that after a rapid rise  $\propto (1+z)^{-2.9}$ , the cosmic SFRD peaked at the redshift  $z \sim 2$ , and subsequently declined as  $\propto (1+z)^{2.7}$  to  $z = 0$ . Within the redshift range of  $1 < z < 3$ , spanning just about 3.5 billion years, more than 50% of the present-day stellar mass was formed. As a result, the stage of the Universe is quite important because it sets how galaxies assemble their stellar mass and evolve to the present stage. SFGs at this redshift trace the prime formation epoch of massive disk and elliptical galaxies in the local universe. This pivotal epoch is frequently referred to as ‘‘Cosmic Noon’’, representing the most active period of galaxy formation and evolution.

### 1.2.2 Star formation rate and its indicators

Star formation rates (SFRs) in galaxies are among the most critical parameters defining galaxies formation and evolution. It is a measurement of the rate in which individual galaxies form new stars, in the unit of  $M_\odot \text{ yr}^{-1}$ . SFRs vary widely among galaxies and depend on factors like gas content, feedback processes, and surrounding environment. The related cosmic SFRD represents a global measurement of star formation activity over cosmic time, expressed as the total SFR per unit comoving volume of the universe, with the unit of  $M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$ . It is derived by summing up the SFRs of all galaxies within a given cosmic volume and redshift interval.

Accurately quantifying SFRs requires reliable diagnostic methods, or indicators. In the well-known framework of [Madau & Dickinson \(2014\)](#), three main types of SFR indicators have been utilized, as also summarized by ([Calzetti 2013](#)):

**A.** The Ultraviolet (UV) continuum emission ( $\sim 1300 - 3000\text{\AA}$ ) probes the direct stellar light emerging from short-lived massive stars, which have surface temperatures of  $> 10000\text{ K}$ . These stars emit a significant fraction of their radiation in the UV part of the spectrum, making it a tracer of recent star formation activity. After adopting a suitable initial mass function (IMF; e.g., [Salpeter 1955](#); [Chabrier 2003](#)) and flux calibration (e.g., [Kennicutt 1998](#)), UV continuum could be converted to a SFR(UV). In the rest-frame UV regime, extensive data from HST, spanning a wide wavelength range from observed frame UV to NIR, have enabled the construction of the UV luminosity function (LF) of galaxies from  $z \sim 0.5$  to  $z \sim 8$ , then converting to the cosmic SFRD (e.g., [Oesch et al. 2010](#); [Finkelstein et al. 2015](#)).

**B.** The Mid/Far-infrared (MIR/FIR) continuum emission ( $\sim 10 - 300\mu\text{m}$ ) from dusty galaxies can be another sensitive tracer of the recent star formation of young stellar populations. As the absorption cross section of the dust is strongly peaked at UV, the UV light from massive stars in star-forming regions is easily absorbed by surrounding dust, and reradiates it in the thermal IR, making the FIR continuum a tracer of star formation ([Leitherer & Heckman 1995](#); [Kennicutt 1998](#)). In the MIR/FIR regime, the Infrared Astronomical Satellite (IRAS) was the first space telescope to survey the all-sky in infrared light, using filters centered at 12, 25, 60, 100  $\mu\text{m}$ . It provided measurements of the infrared luminosity function (IRLF) in the local universe ([Sanders et al. 2003](#)). Subsequently, the infrared Space Observatory, the Spitzer Space Telescope, AKARI, and the Herschel Space Observatory extended these measurements, enabling the determination of the IRLF up to  $z \sim 4$  (e.g., [Caputi et al. 2007](#); [Rodighiero et al. 2010](#); [Gruppioni et al. 2013](#)). Recently, JWST/MIRI observations have advanced the calibration of SFR(IR) by providing deep and high-resolution data ([Kouroumpatzakis et al. 2023](#); [Ronayne et al. 2024](#)).

**C.** Emission lines are frequently used to measure SFRs because they directly trace regions where young, massive stars ionize the surrounding gas. The strength of these emission lines is closely related to photoionization rates, which are driven by the intense UV radiation emitted by O stars with lifetimes shorter than 20 Myr. Among these emission lines, the most reliable indicator is the hydrogen recombination of  $\text{H}\alpha$ . Additionally, certain forbidden metal lines, such as [OII], can serve as quantitative SFR tracers ([Kennicutt 1998](#); [Calzetti et al. 2004](#); [Kennicutt & Evans 2012](#)). However, these metal lines exhibit more complex dependencies on interstellar medium (ISM) conditions, including metallicity and excitation states. For the nebular emission line indicators such as  $\text{H}\alpha$ , both space-based telescopes and ground-based large telescopes have successfully measured the  $\text{H}\alpha$  luminosity functions. These studies have demonstrated the evolution of luminosity density, and consequently the SFRD up to  $z \sim 2.5$  (e.g., [Atek et al. 2010](#); [Sobral et al. 2013](#)).

It should be noticed that different SFR indicators operate on various timescales. SFR(UV) reflects the dominant UV-emitting stellar population, which consists of stars with lifetimes of  $\sim 100\text{ Myr}$ . As a result, SFR(UV) remains constant over timescales of about 100 Myr, assuming continuous star formation. Since dust absorbs UV light from young stars and re-emits it as IR radiation, the timescale of

SFR(IR) is thus similar to that of UV-based SFR, but accounts for obscured star formation. In contrast, SFR(H $\alpha$ ) is primarily influenced by stars with masses  $> 10M_{\odot}$ , which have shorter lifetimes of around 10 Myr. Consequently, SFR(H $\alpha$ ) represents a nearly instantaneous measure of star formation activity.

### 1.2.3 The Star formation main sequence at $z \sim 2$

Locally, a strong correlation exists between the SFR and stellar mass of galaxies. This relationship, first established using the vast numbers of galaxies provided by GALEX, HST, Spitzer, Herschel, is known as the star-forming main sequence (SFMS; Noeske et al. 2007; Elbaz et al. 2011). The SFMS is typically expressed as:

$$\log \text{SFR} = \alpha \log M_{*} - \beta, \quad (1.1)$$

where  $\alpha$  is the slope, usually found to be  $\sim 0.6 - 1.0$ , depending on the redshift and sample selection; and  $\beta$  is the normalization of the main sequence.

Subsequent studies of SFGs at Cosmic Noon revealed that this relationship persists at least up to  $z \sim 3$ , as demonstrated using SFR indicators such as UV continuum or H $\alpha$  luminosity (e.g., Daddi et al. 2007; Speagle et al. 2014; Whitaker et al. 2014). Additionally, it was found that the normalization of the SFMS evolves with redshift, with  $z \sim 2$  objects exhibiting  $\sim 10\times$  larger star formation rates than local ones of the same stellar mass. The scatter around the SFMS is another critical measurement, as it reflects the dominant mechanisms governing galaxy evolution along the SFMS. Tacchella et al. (2016) proposed that the observed small scatters (0.2 – 0.3 dex) can be explained by a “self-regulated” evolution, where star formation is balanced by gas inflows and depletion, while larger scatter being driven by bursty events such as major mergers or violent disk instabilities. However, separate studies at similar redshift have reported large quantitative discrepancies in the slope, normalization, and scatter of the SFMS. Shivaiei et al. (2015) suggested that these inconsistencies could arise from various observational and measurement factors, including the selection criteria for SFGs, the mass completeness of the sample, and the choice of SFR indicators.

The aforementioned studies mostly focus on the galaxies at  $z \sim 2$  with stellar mass larger than  $10^9 M_{\odot}$ , leaving the low-mass end relatively unexplored. While, recent advancements in ultra-deep imaging over the past few years have enabled investigations into the lower-mass regime of the SFMS at  $z \sim 2$ . For instance, Hayashi et al. (2016) used the NB2315 filter on Subaru/MORICS to detect  $\sim 100$  H $\alpha$  emitters (HAEs) with stellar masses down to  $10^8 M_{\odot}$ . Similarly, Terao et al. (2022) identified  $\sim 2,000$  HAEs based on  $K_s$ -band flux excess from the ZFOURGE survey (Straatman et al. 2016). Despite the mass incompleteness, both studies revealed a substantial population of low-mass ( $M_{*} < 10^9 M_{\odot}$ ) galaxies with specific star formation rates (sSFR = SFR /  $M_{*}$ ) that exceed the SFMS by an order of magnitude. Additionally, LAEs at  $z \sim 2$  are also found to lie slightly above the main-sequence at the low-mass end, based on their UV-derived SFRs (Hagen et al. 2016). These offsets suggest that these galaxies are undergoing starburst events rather than normal star-forming activity. Furthermore, they may indicate a larger scatter around the SFMS at lower stellar masses compared to more massive galaxies,

particularly when considering  $H\alpha$ -derived SFRs. This trend is also supported by results from cosmological hydrodynamical simulations (Sparre et al. 2017).

However, the reasons behind the unexpectedly high sSFRs in such low-mass galaxies remain physically unclear. Analogous low-mass starburst galaxies in the local universe, such as blue compact dwarf galaxies (BCDs; Kunth & Östlin 2000; Gil de Paz et al. 2003) and blueberry galaxies (Yang et al. 2017a), typically hold lower metallicities and higher ionization parameters compared to galaxies of similar stellar mass (e.g., Izotov et al. 2006; Janowiecki et al. 2017). The lower metallicity in these systems can be attributed to the recent inflow of pristine (metal-poor) gas, which provides fresh fuel for star formation and drives an increase in SFR. Higher ionization parameters, indicative of a greater production of ionizing photons, also correlate with more intense star formation activity. Moreover, when comparing  $SFR(H\alpha)$  and  $SFR(UV)$ , BCDs exhibit a larger  $H\alpha/UV$  ratio, suggesting bursty star formation histories (SFHs) in a shorter timescale. Although the physical conditions of low-mass galaxies at  $z \sim 2$  are still uncertain, these properties can be inferred through emission lines produced in their HII regions, which will be introduced in the next Section 1.3. Galaxy spectra contain a wealth of information on the fundamental physical processes occurring within them.

## 1.3 Emission line diagnostics and Extreme emission line galaxies

Stars are formed from gas, which continuously cycles in and out of galaxies between the IGM and the ISM. Since stars convert lighter elements into heavier ones through nucleosynthesis, it is expected that the gas inflowing from the IGM is dominated by light elements. On the other hand, the ISM serves as a record-keeper of the change in elements within a galaxy over time. The ISM comprises gas in ionic, atomic, and molecular forms, as well as dust and cosmic rays (Ferrière 2001). However, observational constraints for high- $z$  galaxies mean that most of our information about the ISM at these epochs comes from HII regions, which are composed of high-density ionized atomic hydrogen driven by O- and early B-type stars (Osterbrock 1989). The rest-frame UV-optical-NIR spectra of galaxies are characterized by a variety of prominent emission lines originating from these HII regions. These spectral lines serve as powerful tools for probing the physical and chemical conditions of galaxies. Compared to local universe, galaxies at higher redshift typically exhibit stronger emission lines with higher equivalent widths (EWs) (e.g., Erb et al. 2006; Kewley et al. 2013; Steidel et al. 2014). Below, we list some of the most significant strong emission lines.

### 1.3.1 Hydrogen Recombination Lines

In HII regions, neutral atoms are ionized by UV photons from hot “exciting” stars, becoming ions in a process called ionization. Simultaneously, electrons in the gas clouds recombine with these ions, emitting energy in the form of photons, a process known as recombination. Ionization and recombination occur in equilibrium

within HII regions.

The wavelengths of recombination lines depend on the transitions between energy levels in the atoms. For hydrogen, the primary component of the gas, these transitions span a wide range of wavelengths, from radio waves (caused by transitions between outer energy levels) to the infrared (e.g., the Paschen series), optical (e.g., the Balmer series), and ultraviolet (e.g., the Lyman series). Recombination to the ground state produces UV photons that can, in turn, ionize other hydrogen atoms. Among the observed hydrogen recombination lines, one of the most prominent is H $\alpha$   $\lambda$ 6563Å. Other strong recombination lines include H $\beta$   $\lambda$ 4861Å, P $\alpha$   $\lambda$ 1.875  $\mu$ m, and P $\beta$   $\lambda$ 1.282  $\mu$ m. The Ly $\alpha$  photons undergo a complex resonant scattering process, spreading the photons out over a wide spatial region and making it hard to directly measure the intrinsic emission line strength sometime. On the other hand, since Ly $\alpha$  photons are resonantly scattered by neutral hydrogen, it is a key probe of the neutral fraction of the IGM during EoR. For instance, in a highly neutral IGM, Ly $\alpha$  photons are absorbed. Besides, these hydrogen recombination lines are also widely used to trace star formation rates.

### 1.3.2 Forbidden lines

Forbidden lines from low-lying energy levels of metal ions, such as O $^+$ , O $^{++}$ , N $^+$ , are other prominent emission features from HII regions. Forbidden lines do not satisfy the selection rules for the most probable electric dipole transitions, but correspond to some “rule-breaking” transitions that can occur through less probable, slower pathways (Sparke & Gallagher 2007). These lines are produced in low-density ( $n_e < 10^4 \text{ cm}^{-3}$ ) environments, where ions can remain in excited states long enough for magnetic dipole or electric quadrupole transitions to occur. The intensity of a forbidden line is highly sensitive to the quantity of ions present, so it often gives us detailed information on the density and temperature of the HII regions.

Famous optical and infrared forbidden lines are [OIII] (5007Å,  $^1\text{D}_2 \rightarrow ^3\text{P}_2$ ; 4959Å,  $^1\text{D}_2 \rightarrow ^3\text{P}_1$ ), [OII] (3729Å,  $^2\text{D}_{5/2} \rightarrow ^4\text{S}_{3/2}$ ; 3726Å,  $^2\text{D}_{3/2} \rightarrow ^4\text{S}_{3/2}$ ), and [NII] (6583Å,  $^1\text{D}_2 \rightarrow ^3\text{P}_2$ ; 6548Å,  $^1\text{D}_2 \rightarrow ^3\text{P}_1$ ).

### 1.3.3 Emission line diagnostics

For studies of galaxy formation and evolution, understanding the physical conditions of the ISM in galaxies is essential. These spectral lines from ISM serve as powerful indicators of the physical and chemical conditions in galaxies, providing insights into their stellar population, star formation rate (SFR), chemical abundance, and ionization properties (e.g., Pagel et al. 1979, 1992; Kewley & Dopita 2002; Kewley et al. 2019)

The gas-phase metallicity is commonly expressed as the oxygen abundance relative to hydrogen, defined in units of  $12+\log(\text{O}/\text{H})$ . In a closed-box model of galaxy evolution, metallicity increase over time with successive generations of star formation. However, the closed-box model does not fully capture the complexities of real galaxies, as pristine gas inflows from IGM and outflows of high-metallicity gas are continuously occurring. Theoretically, as time progresses, galaxies undergo

chemical enrichment, leading to an increase in their mean metallicity. Simultaneously, stellar mass builds up through accretion processes. This results in the observed global correlation between the stellar masses ( $M_*$ ) and gas-phase oxygen abundances [ $12+\log(\text{O}/\text{H})$ ] of star-forming galaxies in both the local and high- $z$  universe, known as the mass–metallicity relation (MZR; e.g., [Lequeux et al. 1979](#); [Tremonti et al. 2004](#); [Erb et al. 2006](#)). Further research revealed that the scatter in the MZR could be minimized by incorporating the SFR into the relation, resulting in the “ $M_*$ -SFR- $Z$ ” relation, referred to as the “Fundamental Metallicity Relation” (FMR; e.g., [Mannucci et al. 2010](#); [Sanders et al. 2018](#); [Nakajima et al. 2023](#)). The FMR suggests that accretion of pristine gas from the IGM increases the SFR while simultaneously diluting the ISM’s metallicity.

Gas-phase oxygen abundances can be determined using a wide variety of emission-lines. The “direct” measurement method involves detecting auroral lines, mainly by  $[\text{OIII}]\lambda 4363$  ([Marino et al. 2013](#); [Curti et al. 2017](#); [Sanders et al. 2020](#)). However, these auroral lines are very weak that they are  $\sim 100\times$  fainter than other strong lines in the case of solar metallicity. Before the arrival of JWST data, the auroral lines are barely observed, especially in high- $z$  galaxies. To address this limitation, empirical calibrations are derived by fitting observed relationships between auroral metallicities and strong emission-line ratios. Several commonly used metallicity-sensitive emission-line ratios are employed to estimate oxygen abundances, including:  $([\text{OIII}]\lambda\lambda 4959, 5007 + [\text{OII}]\lambda\lambda 3726, 29)/\text{H}\beta$ ;  $([\text{OIII}]\lambda 5007/\text{H}\beta)/([\text{NII}]\lambda 6584/\text{H}\alpha)$  and  $[\text{NII}]\lambda 6584/\text{H}\alpha$ . Recently, several JWST/NIRSpec survey have been proposed to measure the auroral emission lines at high redshifts, resulting in significant detections of these weak lines (e.g., [Shapley et al. 2024](#)).

The abundance of ionization photons is another interesting physical property of galaxies. The ionization parameter, defined as  $U = n_{\gamma,i}/n_H$ , is the ratio of the ionizing photon density ( $n_{\gamma,i}$ ) to the hydrogen density ( $n_H$ ). This parameter serves as a key metric for quantifying the abundance of ionizing photons produced by massive stars and determines the ionization state within a galaxy ([Yeh & Matzner 2012](#)). In starburst environments in the local universe, such as M82, the maximum ionization parameter is around  $\log U \sim -2.3$  ([Förster Schreiber et al. 2001](#); [Smith et al. 2006](#)). On the other hand, the ionization parameter of galaxies evolves with redshift, that high- $z$  galaxies typically hold higher  $\log U$  than their local counterparts, independent of stellar mass (e.g., [Kewley et al. 2015](#); [Kaasinen et al. 2017, 2018](#)).

Globally, the ionization parameter in galaxies is anti-correlated with gas-phase metallicity, such that low-metallicity galaxies tend to have larger ionization parameters ([Dopita & Evans 1986](#)). This relationship may arise because stellar winds in high-metallicity environments have greater metal opacity, absorbing a larger fraction of ionizing photons and leaving fewer to ionize the surrounding HII regions. The ionization parameter could be measured directly using the emission-line ratio of  $[\text{OIII}]\lambda 5007/[\text{OII}]\lambda\lambda 3726, 29$  (hereafter O32) and  $([\text{SIII}]\lambda 9069 + [\text{SIII}]\lambda 9531)/[\text{SII}]\lambda\lambda 6717, 31$  (hereafter S32).

[Nakajima & Ouchi \(2014\)](#) has used the O32 versus R23-index diagram to examine the metallicity and ionization parameters of Ly $\alpha$  emitters at  $z \sim 2$ . Their study revealed lower metallicity and higher ionization parameters in these emitters compared to typical LBGs at similar redshifts. Building upon these findings, it is

currently speculated that galaxies showing strong emission lines, such as [OIII], may play a significant role in the cosmic reionization process (e.g., Nakajima et al. 2016; Yang et al. 2017b; Jaskot et al. 2019; Tang et al. 2019; Onodera et al. 2020).

### 1.3.4 Extreme Emission Line Galaxies

In spectroscopic surveys, low-mass galaxies undergoing vigorous starbursts across galaxy-wide scales are often recognized by their high-excitation emission lines with unusually high EWs (e.g., van der Wel et al. 2011; Atek et al. 2011; Nakajima et al. 2016; Forrest et al. 2017; Tang et al. 2019). These high emission line EWs are most commonly observed in the [OIII] $\lambda\lambda$ 4959, 5007 and H $\alpha$  lines, giving rise to the designation “Extreme Emission Line Galaxies” (EELGs, typically defined as EW larger than several hundred Å). Their intense nebular emission lines are often driven by ionizing photons produced by massive and short-lived O- and B- stars, while the underlying stellar continuum reflects contributions from longer-lived, less massive stars (Eldridge & Stanway 2022).

Over the past two decades, the advent of large sky surveys, such as the Sloan Digital Sky Survey (SDSS; York et al. 2000; Kauffmann et al. 2003; Brinchmann et al. 2004), the COSMOS survey (Scoville et al. 2007), have enabled systematic searches of relatively large EELG samples at various redshifts. These studies reveal that EELGs are quite rare in the local universe but become increasingly common at higher redshifts (Izotov et al. 2011; Smit et al. 2014). EELGs provide a unique opportunity to study the extreme conditions in galaxies of the early universe, while also offering critical insights into the process of cosmic reionization as they are currently considered the primary drivers of cosmic reionization at  $6 < z < 15$  (Robertson et al. 2015; Finkelstein et al. 2019). Although EELGs are a critical sub-sample of extragalactic objects, a thorough understanding of their properties remains elusive, for example a multiple emission line analysis of them. To achieve this, large and representative samples of EELGs must be assembled and studied in detail.

## 1.4 Detection of emission lines

Given the significance of emission line diagnostics, several extensive sky surveys have endeavored to extract these emission lines from large sets of observational data. Examples include the Sloan Digital Sky Survey (SDSS; York et al. 2000; Kauffmann et al. 2003) and the MOSFIRE Deep Evolution Field survey (MOSDEF; Kriek et al. 2015). Traditionally, spectroscopy or narrow-band imaging are usually applied for finding emitters. However, these traditional observations usually consume a large amount of time and suffer from selection bias and constructing large samples.

### 1.4.1 Spectroscopy

Multi-object spectroscopy is the most widely used method to obtain full spectra of galaxies and isolate emission lines from the stellar continuum. A key advantage of spectroscopy is the ability to measure emission line strengths with great accuracy.

In order to acquire rest-frame optical spectroscopy that covers all strong emission lines in the rest-frame wavelength range of 3,700 to 6,800 Å for  $z \sim 2$  galaxies, multi-object spectrographs operating in the J, H, and K bands with high resolving power are used. Examples include Keck/MOSFIRE (McLean et al. 2012) and Gemini/Flamingos2 (Eikenberry et al. 2012). These advances, high-throughput and sensitive instruments have provided a wealth of valuable spectroscopic data in recent years.

When conducting multi-object spectroscopic observations, one or more masks with long slits are often carefully prepared in advance, known as “long-slit spectroscopy”. The mask with slits ensures that the dispersed spectrum of each source is captured on the detector array without overlapping with spectra from other sources. However, due to the limited widths of the slits (typically  $< 1''$ ), some light from the outskirts of galaxies often falls outside the slit apertures. This slit loss can account for up to 20% of the total light, when the seeing is comparable to the slit width. In addition to slit loss, spectroscopic targets often encounter selection biases and limitations in search volume, particularly for high- $z$  observations. Each mask can accommodate the simultaneous observation of only about 30 objects, making it challenging to construct large samples within a limited observation time. Moreover, exposure times for each filter can span several hours, especially for high- $z$  galaxies, which are generally fainter than the local ones. For instance, one of the largest high- $z$  spectroscopic survey, MOSDEF (Kriek et al. 2015), required 24 nights of observations to obtain rest-frame optical spectra for  $\sim 600$  galaxies. The MOSDEF sample also exhibited selection biases toward massive galaxies ( $M_* > 10^9 M_\odot$ ), potentially limiting a comprehensive understanding of the broader galaxy population.

On the other hand, “fiber spectroscopy” on large-aperture telescopes enables large-scale galaxy spectroscopic surveys. For example, the Fiber Multi-object Spectrograph (FMOS; Kimura et al. 2010) on the Subaru Telescope allows the observation of up to 400 fibers distributed across a 30'-diameter field of view (FOV). The FMOS-COSMOS survey successfully obtained the spectroscopic data of nearly 2,000 galaxies (Silverman et al. 2015; Kashino et al. 2019). However, each fiber integrates light over a small area, so it cannot resolve spatial details within extended objects. Also, the relatively low sensitivity of fiber spectroscopy makes it challenging to observe fainter galaxy populations, resulting in a sample that is even more biased toward massive galaxies.

## 1.4.2 Narrow-band imaging

Photometric observations with NB filters, which have a ratio of central wavelength to bandwidth  $\lambda/\Delta\lambda \sim 100$ , allow the derivation of emission lines. These NB filters are well designed with bandpasses tailored to specific emission lines at specific redshifts, making them ideal for imaging clusters of galaxies where member galaxies share very similar redshifts. For instance, MAHALO-Subaru (Kodama et al. 2013; Shimakawa et al. 2018a,b) observe several separate proto-clusters at different redshifts using different narrow-band filters on Subaru/MOIRCS (Suzuki et al. 2008).

With NB filters, galaxies exhibiting strong emission lines can be easily identified.

The strengths of these emission lines can be determined by measuring the color excess between NB and a corresponding broad-band (BB;  $\lambda/\Delta\lambda \sim 5$ ) filter with a similar central wavelength. Typically, emission lines with observed-frame equivalent widths larger than  $50\text{\AA}$  in the  $K_s$  band produce a color excess of  $\sim 0.2$  magnitude in NB filters, allowing emission line strengths to be derived with sufficient accuracy.

However, the narrow redshift windows of NB imaging surveys impose significant limitations on their search volume. Constructing large samples within a reasonable observation time is challenging. For example, Sobral et al. (2013) compiled a sample of  $\sim 800$   $H\alpha$  emitters at  $z \sim 2.23$  from more than 100 hours exposure-time data in the COSMOS and UKIDSS-UDS fields. Besides, multiple emission lines can fall within the same filter, leading to blending. For instance,  $H\alpha$  and  $[\text{NII}]$  doublets are so close in wavelength that even low-resolution spectroscopy cannot separate them, resulting in contamination in the NB filter and uncertainty on  $H\alpha$  flux measurements.

### 1.4.3 Broad/Medium band imaging with SED fitting

Both spectroscopy and NB imaging face challenges in constructing large samples of high- $z$  galaxies. Encouragingly, recent studies have demonstrated the feasibility of efficiently identifying galaxies with strong emission lines using only BB photometry (e.g., Stark et al. 2013; Saito et al. 2020; Onodera et al. 2020; Terao et al. 2022). These studies have revealed that galaxies with high EWs of emission lines exhibit noticeable flux excess in BB photometry. Advances in spectral energy distribution (SED) fitting techniques have made it possible to extract emission line fluxes from BB photometry. By incorporating emission line templates, i.e., stellar population synthesis models augmented with nebular emission line fluxes, SED fitting accounts for the flux boost caused by emission lines. This approach allows for estimating the underlying stellar continuum, which is lower than the observed flux, and subsequently deriving emission line strengths from the flux excess. Moreover, SED fitting with emission line templates could not only derive the emission line strengths, but also improve the accuracy of derived galaxy properties, such as stellar masses, dust attenuation and ages. For example, Onodera et al. (2020) used the COSMOS2015 catalog (Laigle et al. 2016) to select extreme  $[\text{OIII}]$  emitters (O3Es) at  $3 < z < 3.7$ , based on flux excess observed in the UltraVISTA- $K_s$  filter.

However, this method has limitations. BB filters, with their larger bandwidths compared to NB filters, are even more susceptible to contamination from additional emission lines. In the broad-band  $K_s$  filter at  $z \sim 2$ , emission lines beyond  $[\text{NII}]$ , such as  $[\text{SII}]$  doublets, may also contribute to the observed flux. This necessitates corrections based on the ratio of  $H\alpha$  to the total emission line strength to accurately estimate the true  $H\alpha$  flux. The related line ratios are often estimated from relative line strengths table in the local universe (e.g., Inoue 2011), but carry uncertainties due to galaxy diversity in stellar mass, SFR and metallicity (e.g., Faisst et al. 2018; Reddy et al. 2018; Topping et al. 2021).

These aforementioned studies solely relied on a single broad-band photometric filter to search for emission line galaxies, without extracting emission line measurements from other photometric filters. On the other hand, only with the photomet-

ric data from broad-band filter may not be enough to derive accurate information and sometimes extreme emission lines even cause the overestimates of the continuum levels (Terao et al. 2022). These two issues can be addressed by incorporating medium-band filters (e.g., van Dokkum et al. 2009) during the analysis. MB photometry employs narrower bandwidths ( $\Delta\lambda/\lambda \sim 15$ ) compared to BB filters, allowing for more precise wavelength sampling of emission lines. The ZFOURGE survey (Straatman et al. 2016) utilized the FourStar imager (Persson et al. 2013) to acquire NIR medium-band photometry. Combined with other optical and infrared photometry, the ZFOURGE catalog offers a powerful tool for estimating multiple emission lines and conducting emission line diagnostics. By building composite spectral energy distributions (SEDs), Forrest et al. (2018) categorized a significantly large galaxy sample at  $1 < z < 4$  and characterized  $\sim 150$  EELGs. One notable advantage of employing photometry for multiple emission lines analysis is the ability to construct larger and more unbiased samples, which enhances the statistical significance of the results and allows for more robust conclusions.

## 1.5 Resolved study on galaxy properties

Resolved studies of galaxy properties are fundamental for understanding the spatially dependent processes that govern galaxy evolution, allowing us to place new constraints on galaxy formation and evolution: their mass assembly histories, evolution modes, chemical enrichment and earliest quenching mechanisms. In pre-JWST era, lower-redshift studies have been able to resolve galaxies and their components at  $z < 2$ , but at higher redshifts, resolved analyses have typically only been possible in lensed systems (e.g., Zitrin et al. 2011; Vanzella et al. 2017).

With the unprecedentedly deep and high-resolution data from JWST, which also extend to longer wavelengths than HST, we can now resolve the rest-frame optical structures of high-redshift galaxies in remarkable detail. The JWST Early Release Observations (EROs) characterized the morphologies of galaxies at Cosmic Noon. For example, recent studies report a much higher fraction of elongated or disk-like morphologies in galaxies at  $z \sim 2$  compared to previous HST observations (Ferreira et al. 2022, 2023; Kartaltepe et al. 2023; Jacobs et al. 2023). The improved sensitivity of JWST, particularly in the infrared, allows the detection of low surface brightness structures that were too faint (or unable) to be captured by HST.

As was mentioned in Section 1.3, emission lines are powerful indicators of the physical and chemical conditions within galaxies, including their stellar populations, metallicities. Over the past decade, spatially resolved analyses of emission lines at  $z \sim 2$  have mainly been conducted through ground-based integral field unit (IFU) and observation surveys (e.g., Förster Schreiber et al. 2009; Wisnioski et al. 2019). These observations have revealed key properties of individual galaxies at  $z \sim 2$ , such as metallicity gradients (Wuyts et al. 2016), electron density distribution (Davies et al. 2021). Inevitably, due to limitations in their depth and resolution, these studies were restricted to galaxies with stellar mass larger than  $10^{9.5} M_{\odot}$ . On the other hand, the EELGs tend to have lower stellar masses compared to other samples. Therefore, to achieve a comprehensive and unbiased analysis of spatially resolved emission lines at Cosmic Noon, it is essential to include these EELGs. With

the arrival of high-resolution data from JWST, it has become feasible to spatially resolve galaxy emission lines at high redshifts (e.g., [Giménez-Arteaga et al. 2023, 2024](#); [Shen et al. 2024](#)). These studies successfully extracted emission line regions on a kiloparsec (kpc) scale with high EWs of several thousand angstroms, revealing intense starbursts.

## 1.6 Objectives and Structure of this dissertation

The objective of this dissertation is to comprehensively investigate the properties of a population of low-mass EELGs at Cosmic Noon, whose physical characteristics remain unclear so far. As mentioned earlier, the strong emission lines from HII regions carry crucial information about their host galaxies. This study aims to construct a large sample of star-forming galaxies with multiple emission line flux measurements and to analyze their properties statistically.

To achieve this, we present a systematic search for HAEs at  $2.05 < z < 2.5$  using the photometric catalog from the FourStar galaxy evolution survey (ZFOURGE; [Straatman et al. 2016](#)). By applying SED fitting with emission line templates, we construct a large sample of HAEs, enabling a statistical understanding of their physical properties. This method not only derives  $H\alpha$  emission line strengths but also extracts [OIII] and [OII] emission line strengths from flux excesses in the J and H medium-band ZFOURGE data. This approach allows us to probe more detailed physical characteristics, such as metallicity and ionization parameters, particularly for low-mass galaxies through multiple emission line analyses.

Given that previous studies suggested that low-mass EELGs are believed to dominate the ionizing photon budget and play an important role during the EoR, we further investigate their ionization properties by measuring the ionizing photon production efficiency,  $\xi_{\text{ion}}$ , of individual galaxies.

Furthermore, we conduct a detailed spatially resolved analysis of a subsample of the parent HAEs at  $2.05 < z < 2.35$  using the JWST/NIRCam photometric data from the first and second data releases of the JWST Advanced Deep Extragalactic Survey (JADES; [Eisenstein et al. 2023a](#); [Rieke et al. 2023](#); [Eisenstein et al. 2023b](#)). From the F150W imaging, we identify a large number of resolved strong [OIII] emission line regions (“Green Seeds”) within the galaxy structures. Using a specialized algorithm, we extract and analyze the properties of these regions to gain deeper insights into their role within their host galaxies.

The outline of this paper is as follows. In Chapter 2, we first describe the observations and data used in this study. This includes a brief introduction to the multi-wavelength photometric catalog from the ZFOURGE survey, supplemented by medium K-band imaging from SWIMS. This chapter also details the sample selection process, the SED fitting with the emission line templates, and the basic measurements of emission line fluxes, including  $H\alpha$ , [OIII], [OII]. In Chapter 3 and Chapter 4, we derive the main physical properties of the selected HAEs, such as stellar mass, SFR, and EWs of each emission line. We then carry out a multiple emission line analysis of the HAEs and compare our sample with other analogous objects from other studies. In Chapter 5, we measure the  $\xi_{\text{ion}}$  of these HAEs and examine its relationship with various galaxy properties. This chapter also explores

the observational and modeled results of  $\xi_{\text{ion}}$  and discusses the implications of low-mass galaxies for their contribution to cosmic reionization. In Chapter 6, we describe the methodology used for extracting the resolved strong emission line regions and introduce the newly discovered Green Seeds in the HAE samples. A comprehensive analysis of the properties of these Green Seeds is presented, comparing them with other spatially resolved structures. We also discuss the potential triggers for the formation and evolution of these Green Seeds and their possible roles in host galaxies. Finally we summarize our results and conclusions, and propose future observations to further investigate the properties of HAEs at Cosmic Noon in Chapter 6.

We adopt the following abbreviations for strong emission-line ratios:

$$\text{O32} = [\text{OIII}]\lambda 5007 / [\text{OII}]\lambda\lambda 3726, 29, \quad (1.2)$$

$$\text{R23} = ([\text{OIII}]\lambda\lambda 4959, 5007 + [\text{OII}]\lambda\lambda 3726, 29) / \text{H}\beta. \quad (1.3)$$

Throughout this thesis, we adopt the AB magnitude system (Oke & Gunn 1983), assume a Chabrier(2003) initial mass function (IMF) and a  $\Lambda$ CDM cosmology with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.3$ , and  $\Omega_\Lambda = 0.7$ .



# Chapter 2

## The Compilation of HAEs at $z \sim 2.3$

### 2.1 Observation and Data

#### 2.1.1 The ZFOURGE survey

We use the photometric catalog from the FourStar galaxy evolution survey (ZFOURGE, [Straatman et al. 2016](#)). ZFOURGE is a 45-night photometric observation survey with the FourStar near-infrared camera ([Persson et al. 2013](#)) on 6.5-meter Magellan telescope. The observation targets at three legacy fields: GOODS-S ([Giacconi et al. 2002](#)), COSMOS ([Scoville et al. 2007](#)) and UDS ([Lawrence et al. 2007](#)) with a total coverage of  $\sim 450 \text{ arcmin}^2$  (128, 135, 189  $\text{arcmin}^2$  in GOODS-S, COSMOS, UDS, respectively).

The unique characteristic of ZFOURGE is that it has five near-infrared medium-band (MB) filters:  $J_1$ ,  $J_2$ ,  $J_3$ ,  $H_s$ ,  $H_l$ , covering a similar wavelength range as the broad-band (BB) filters  $J$ ,  $H$ , and a ultra-deep  $K_s$  map. ZFOURGE catalog also includes multiwavelength public data. In all, the GOODS-S, COSMOS, and UDS fields have 40, 37, and 26 photometric filters with the 80% completeness of 26.0, 25.5, and 25.8 magnitudes in the  $K_s$  images, respectively.

Such a large number of photometric filters makes it possible to accurately derive the photometric redshift (hereafter  $z_{\text{phot}}$ ) of galaxies. [Nanayakkara et al. \(2016\)](#) measure the spectroscopic redshift (hereafter  $z_{\text{spec}}$ ) of  $\sim 200$  galaxies at  $1.5 < z < 2.5$  in ZFOURGE-COSMOS and UDS field, and confirm that the primary  $z_{\text{phot}}$  for star-forming galaxies (SFGs) from ZFOURGE catalog has a very good accuracy that  $\Delta z / (1 + z_{\text{spec}}) < 3\%$ , where  $\Delta z = |z_{\text{spec}} - z_{\text{phot}}|$ .

#### 2.1.2 SWIMS medium K-band imaging

SWIMS (Simultaneous-color Wide-field Infrared Multi-object Spectrograph, [Konishi et al. 2012](#); [Motohara et al. 2016](#)) is the first-generation near-infrared instrument for the University of Tokyo Atacama Observatory (TAO) 6.5m telescope ([Yoshii et al. 2010](#)). It has medium K-band filters (detailed in Table 1), which can provide more detailed information on the  $\text{H}\alpha$  emission line at  $z \sim 2$ . During its com-

**Table 2.1.** SWIMS medium K-band and observation in S18B

Filter	$\lambda$ ( $\mu\text{m}$ )	Depth <sup>a</sup> ( $5\sigma$ , AB mag)	FWHM ( $''$ )
$K_1$	1.95 – 2.09	23.7	1 $''$ 0
$K_2$	2.10 – 2.24	23.8	0 $''$ 6

**Notes.** <sup>a</sup> When calculating image depths (limiting magnitudes), we follow the same method as in [Straatman et al. \(2016\)](#) and directly measure the fluxes of circular apertures with 0 $''$ 6 diameter (same as ZFOURGE) placed at 5000 random positions on the final reduced images.

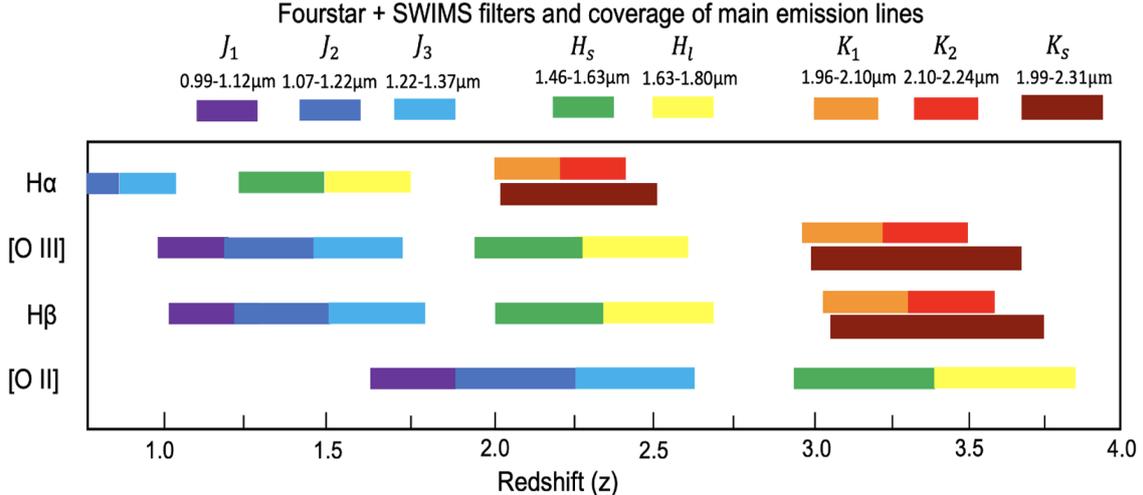
missioning observation at the Subaru Telescope in S18B, an area of approximately 20 arcmin<sup>2</sup> within the ZFOURGE-COSMOS Field has been observed, which contributes to nearly 1/6 of the total coverage of the ZFOURGE-COSMOS catalog. The total integration time is  $\sim 2$  hours for the  $K_1$  filter and  $\sim 1.5$  hours for the  $K_2$  filter, as outlined in Table 1.

The SWIMS data are reduced by a custom Python-3 pipeline, named “SWSRED”, which has demonstrated good stability and performance ([Konishi et al. 2020](#)). To incorporate the SWIMS sources into the ZFOURGE catalog, we follow the same PSF matching method employed by the ZFOURGE survey (see Section 3.1 of [Straatman et al. 2016](#) for details). Among the objects at  $z \sim 2.3$  in the ZFOURGE-COSMOS field, nearly 1/6 of them have detections in both the  $K_1$  and  $K_2$  filters and their medium K-band photometry is then merged into the ZFOURGE catalog. The SEDs are fitted as described below in Section 2.3. The total integration times of the  $K_1/K_2$  images are comparably shorter than those of the ZFOURGE ultra-deep  $K_s$  images. Also, the coverage of SWIMS observation is much smaller than total ZFOURGE survey. Thus, we continue with the extraction of the H $\alpha$  emission line from the ZFOURGE  $K_s$  filter in Section 2.4.

## 2.2 Sample Selection

We construct a parent sample of galaxies with H $\alpha$  emission lines falling within the ZFOURGE  $K_s$  filter. Based on the transmission curve of the  $K_s$  filter, we determine that the H $\alpha$  emission line at  $2.05 < z < 2.5$  is fully shifted into the  $K_s$  filter as also shown in Figure 2.1.

In the ZFOURGE catalog,  $z_{\text{phot}}$  were obtained using the photometric reshift code, Easy and Accurate Z from Yale (EAZY; [Brammer et al. 2008](#)). Additionally, we merged the SWIMS  $K_1/K_2$  fluxes into the ZFOURGE catalog. Among the 1541 objects at  $2.05 < z < 2.5$  in ZFOURGE-COSMOS field, 207 objects ( $\sim 13\%$ ) have both  $K_1$  and  $K_2$  detection. For objects that possess supplementary medium K-band data, we conducted a reiteration of the EAZY code and updated the  $z_{\text{phot}}$  values with the newly generated outputs. The incorporation of the additional SWIMS MB data contributes to an enhanced constraint on the photometric redshifts. We quantify the errors in the photometric redshifts, the  $\sigma_z = |z_{\text{phot}} - z_{\text{spec}}|/(1 + z_{\text{spec}})$  as ZFOURGE was done. After including our SWIMS MB data,  $\sigma_z$  drop from 0.03 to 0.02 in the ZFOURGE-COSMOS field (see details in Appendix A).



**Figure 2.1.** Combinations of several strong emission lines, including  $H\alpha$ ,  $[O\text{III}]$ ,  $H\beta$ ,  $[O\text{II}]$  in the observed-frame and the MB (BB) filters in which these emission lines drop. At  $z \sim 2.3$ ,  $H\alpha$  would fall into the K-band filters. Simultaneously,  $H\beta$  and  $[O\text{III}]$  could be observed in the H medium-band filters and  $[O\text{II}]$  in the J medium-band filters.

After updating the photometric redshifts in the ZFOURGE catalog, we compare them with the spectroscopic redshifts ( $z_{\text{spec}}$ ) from the MOSDEF survey (Kriek et al. 2015) and the grism redshifts ( $z_{\text{grism}}$ ) from the 3D-HST data release (Brammer et al. 2012; Momcheva et al. 2016). From the MOSDEF catalog, we have identified around 110 galaxies with high-quality spectroscopic redshift measurements at  $z_{\text{spec}} \sim 2.3$ . These galaxies have also been cross-matched to the ZFOURGE-COSMOS catalog, allowing for a combined analysis of both datasets. The three legacy fields of ZFOURGE, namely COSMOS, UDS, and GOODS-S, are all covered by the 3D-HST survey. We have selected galaxies with high-quality  $z_{\text{grism}}$  values (where  $z_{\text{best}_s} \leq 2$ , please refer to Momcheva et al. (2016) for more details) and cross-matched nearly 400 galaxies at  $z_{\text{grism}} \sim 2.3$  across the three fields. The majority of galaxies demonstrate a difference  $\Delta z / (1 + z) < 0.05$ , with only 12 galaxies ( $< 3\%$ ) being outliers that have  $\Delta z / (1 + z) > 0.15$ . To fully utilize the spectroscopic (grism) data, we proceed to replace the  $z_{\text{phot}}$  values in the ZFOURGE catalog with corresponding  $z_{\text{spec}}$  or  $z_{\text{grism}}$  values obtained from the MOSDEF and 3D-HST Emission-Line Catalogs. In cases where both  $z_{\text{spec}}$  and  $z_{\text{grism}}$  are present,  $z_{\text{spec}}$  takes precedence.

The ZFOURGE catalog employs a flag called `use`, which serves to eliminate various objects such as stars, objects in close proximity to stars, low signal-to-noise ratio (S/N) objects, and objects with low exposure time. We use this flag to eliminate contaminants that a standard selection of galaxies can be obtained by choosing sources with `use=1`. Additionally, the ZFOURGE catalog includes a list of AGN hosts that were identified through X-ray, IR, and radio selection methods, as outlined in Cowley et al. (2016). We exclude these AGNs in the following fitting and analyzing of galaxy properties. Applying these selection criteria, a total of 3754 galaxies at  $2.05 < z < 2.5$  are retained, with 1307, 1235, and 1212 galaxies in the GOODS-S, COSMOS, and UDS fields, respectively.

## 2.3 SED fitting with emission line templates

In this study, we perform SED fitting to obtain primary galaxy properties using the 2020.0 version of Code for Investigating GALaxy Emission (CIGALE; [Burgarella et al. 2005](#); [Noll et al. 2009](#); [Boquien et al. 2019](#)). We utilize the photometric data covering from  $0.3 - 8 \mu\text{m}$  from ZFOURGE catalog and SWIMS.

The emission lines templates in CIGALE are computed based on user-defined gas-phase metallicities (Z) and ionization parameters (U), allowing for adjustable emission line templates during the SED fitting process. This is an update of SED fitting carried out in [Terao et al. \(2022\)](#) who have employed the same dataset but performed SED fitting using the Fitting and Assessment of Synthetic Templates (FAST; [Kriek et al. 2009](#)). In addition, they have used a fixed emission-line template to be added into the spectrum, differing from ours where multiple parameter choices are available for emission line templates. Besides, [Terao et al. \(2022\)](#) solely considered the best fitting result from the template with the smallest  $\chi^2$  value in their analysis, whereas we adopt a Bayesian-like approach, assigning weights to all models based on their  $\chi^2$  values. This Bayesian fitting approach allows for obtaining a better estimate of the physical properties, such as stellar mass, with less uncertainties.

### 2.3.1 Stellar population models and Star formation history

We utilize composite stellar population models generated from BC03 ([Bruzual & Charlot 2003](#)) with a Chabrier IMF ([Chabrier 2003](#)). The metallicity Z of stellar population are permitted to be 0.004, 0.008 and 0.02. Next, we adopt a delayed- $\tau$  model to represent the star formation history (SFH) in a functional as follows:

$$\text{SFR}(t) \propto \frac{t}{\tau^2} \times \exp(-t/\tau) \quad \text{for } 0 \leq t \leq t_0. \quad (2.1)$$

The delayed- $\tau$  model offers a smooth SFH, characterized by an increasing SFR from the onset of star-formation until it reaches its peak at  $\tau$ . Subsequently, the SFR gradually decreases. This model is considered to be a more representative SFH for SFGs compared to a constant SFH or an exponentially declining SFH (e.g., [Cohn et al. 2018](#); [Onodera et al. 2020](#)).

When establishing grids for our fitting process, we set the stellar population age ( $t_0$ ) within the range of  $\log(t_0/\text{yr}) = 7-10$ , with steps of 0.1 dex. The upper limit of  $t_0$  is assumed not to exceed the age of the universe at  $z \sim 2$ . The e-folding time ( $\tau$ ) ranges within  $\log(\tau/\text{yr}) = 8-10$ , with steps of 0.1 dex.

### 2.3.2 Nebular emission model

CIGALE models the emission of ionized gas in HII regions of the galaxy by using the nebular templates based on [Inoue \(2011\)](#) and implemented through CLOUDY 13.0 ([Ferland et al. 1998, 2013](#)). These nebular templates provide the relative intensities of 124 lines emitted by HII regions. The templates are parameterized according to a given ionization parameter U, and gas-phase metallicity Z (which is assumed to be the same as the stellar one), along with a fixed electron density  $n_e = 100 \text{ cm}^{-3}$ .

In our fitting process, we consider an adjustable ionization parameter with values of  $\log U = -4, -3, -2, -1$ . A higher-resolution grid of  $\log U$  have no effect on the results. Besides, Lyman continuum (LyC) photons are assumed to be completely absorbed by neutral hydrogen, i.e.,  $f_{esc} = 0$ , and there is no LyC absorption by dust.

### 2.3.3 Dust attenuation model

In this study, we fit the stellar continuum using the Calzetti curve supported by CIGALE. Also, stellar continuum and nebular emission usually suffer different dust extinction because HII regions possess a distinct distribution of dust or dust with different properties (Calzetti et al. 1994; Charlot & Fall 2000). To address this issue, an approach is to assume that each component is subject to a different dust attenuation curve. For starburst galaxies, the Milky Way curve (Cardelli et al. 1989) is commonly adopted for the nebular line emission, even at higher redshift (e.g., Reddy et al. 2020). Additionally, the extinction of the stellar continuum  $E(B-V)_{star}$ , and the extinction in the ionized gas  $E(B-V)_{neb}$  are typically different. We parameterize the difference of color excesses by a factor  $f$  such that:

$$E(B - V)_{neb} = \frac{E(B - V)_{star}}{f}, \quad (f < 1). \quad (2.2)$$

While  $f$ -factor still suffers from significant uncertainty at high redshift (e.g., Kashino et al. 2013; Price et al. 2014; Reddy et al. 2020), Saito et al. (2020) proposed a simple redshift evolution for the  $f$ -factor as  $f = 0.44 + 0.2z$ . Thus, we adopt the Milky Way curve of Cardelli et al. (1989) with an  $f$ -factor of 0.8 as the dust attenuation curves for the nebular emission.

## 2.4 Emission line measurement

To obtain dependable emission line fluxes from the best-fit model and subsequently select emitters, we follow Terao et al. (2022) and employ the concept of “flux excess” ( $F_{excess}$ ), in units of  $10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2}$ . This value is calculated as the difference between the total observed flux and the flux of the stellar continuum derived from SED within a broad/medium-band filter of a bandwidth ( $\Delta\lambda$ ). The computation is as follows:

$$F_{excess} (\text{erg s}^{-1} \text{ cm}^{-2}) = f_{obs} \times \Delta\lambda - \int_{\lambda_1}^{\lambda_2} f_{cont} d\lambda, \quad (2.3)$$

where  $F_{excess}$  represents the total flux of all emission lines within a specific filter, theoretically being zero when no emission line falls within the filter,  $\lambda_1(\lambda_2)$  the cut-off wavelength,  $\Delta\lambda = \lambda_2 - \lambda_1$  the bandwidth of the filter,  $f_{obs}$  the observed flux density of the filter that the target emission line is located, and  $f_{cont}$  the stellar continuum obtained from the best-fit SED model for each galaxy. The best-fit SED of CIGALE also returns the fluxes of the strong emission lines. By modeling different production rate of the Lyman continuum photons ( $N_{lyc}$ ) in a galaxy, Terao et al. (2022) have found that the derived  $H\alpha$  line flux from the emission line template strongly depends

**Table 2.2.** Main criteria for selecting galaxies at  $z \sim 2.3$

Emission (continuum)	ZFOURGE Filter <sup>a</sup>	Flux excess (S/N)
H $\alpha$	$K_s$	$3\sigma$
[OIII]	$H_s$ or $H_l$	$2\sigma$
[OII]	$J_2$ or $J_3$	$2\sigma$

**Notes.** <sup>a</sup> The delicate design of  $J$  and  $H$ -band medium filters make [OIII] and [OII] emission lines drop into  $H_s$  and  $J_2$  filters at  $z < 2.258$ , and into  $H_l$  and  $J_3$  filters at  $z > 2.258$ , simultaneously. See also in Figure 2.1.

on the model assumption. On the other hand, the stellar continuum flux density from the SED fitting has been more robust against the model assumption. Besides, [OIII] and [OII] emission line fluxes are strongly depend on the ISM properties (Z, U) assumed in the SED model, which cannot fully capture the true strengths of the emission lines.

One remaining issue in our method is the potential overlap of multiple emission lines within the same broad/medium-band filter due to its large bandwidth. It becomes challenging to isolate individual emission lines using only the broad/medium-band (BB/MB) data. To address this issue, we make an assumption and define a “purity ratio”, denoted as  $r_{EL}$ , which represents the target emission line’s contribution relative to the combined strengths of all the emission lines present in the same filter.

$$F_{EL} (\text{erg s}^{-1} \text{cm}^{-2}) = r_{EL} \times F_{excess}, \quad (2.4)$$

where  $F_{EL}$  is the final derived observed emission line fluxes and  $F_{excess}$  is the flux excesses from Equation (6).

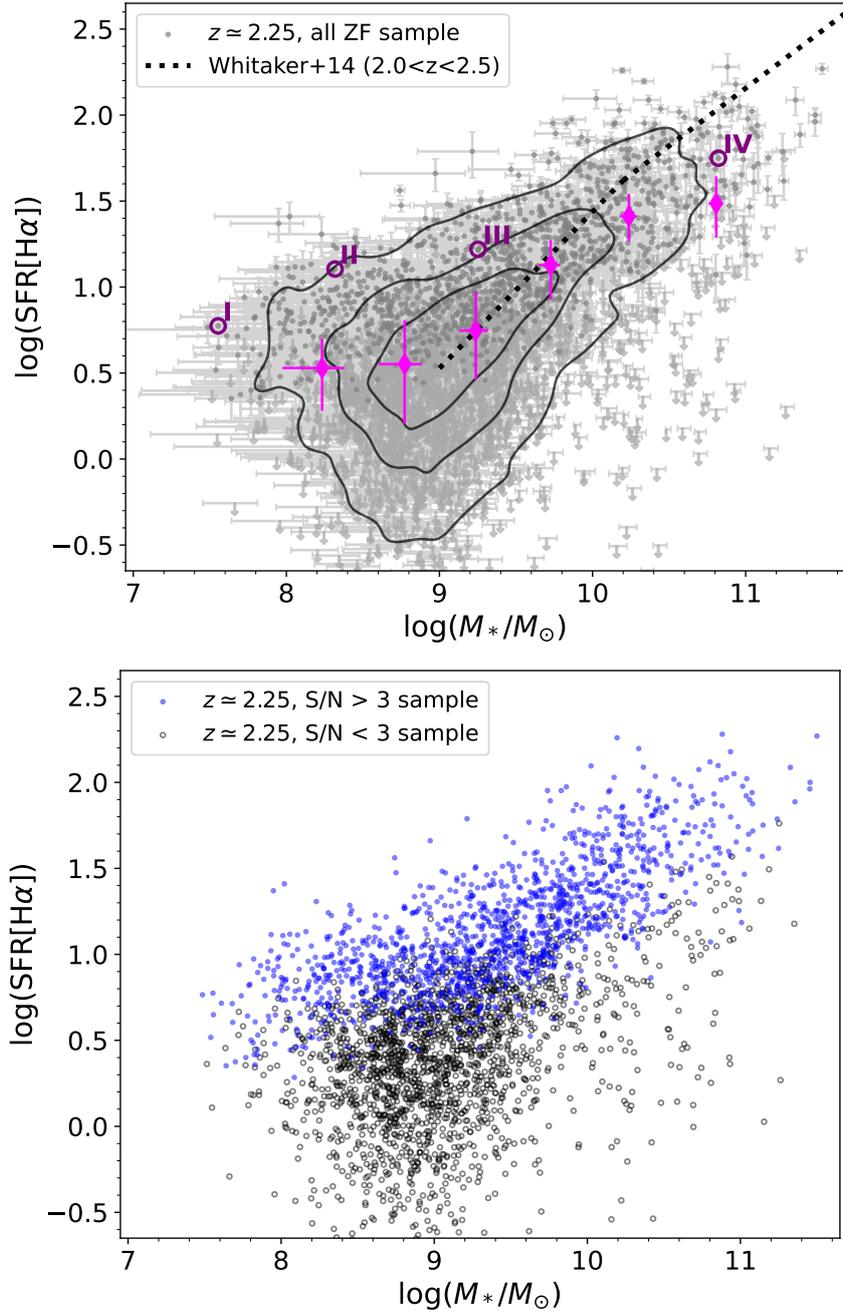
In our analysis, the uncertainties in emission line flux measurements arise from two primary sources: the observed flux errors ( $\sigma_{obs}$ ) and the uncertainties in the estimation of the stellar continuum ( $\sigma_{cont}$ ). Here,  $\sigma_{obs}$  are directly taken from the ZFOURGE catalog, while  $\sigma_{cont}$  are derived from Bayesian sampling of the model flux error by CIGALE. The total uncertainty in the emission line flux ( $\sigma_{EL}$ ) is computed using the formula:

$$\sigma_{EL} = r_{EL} \times \sqrt{\sigma_{obs}^2 + \sigma_{cont}^2}. \quad (2.5)$$

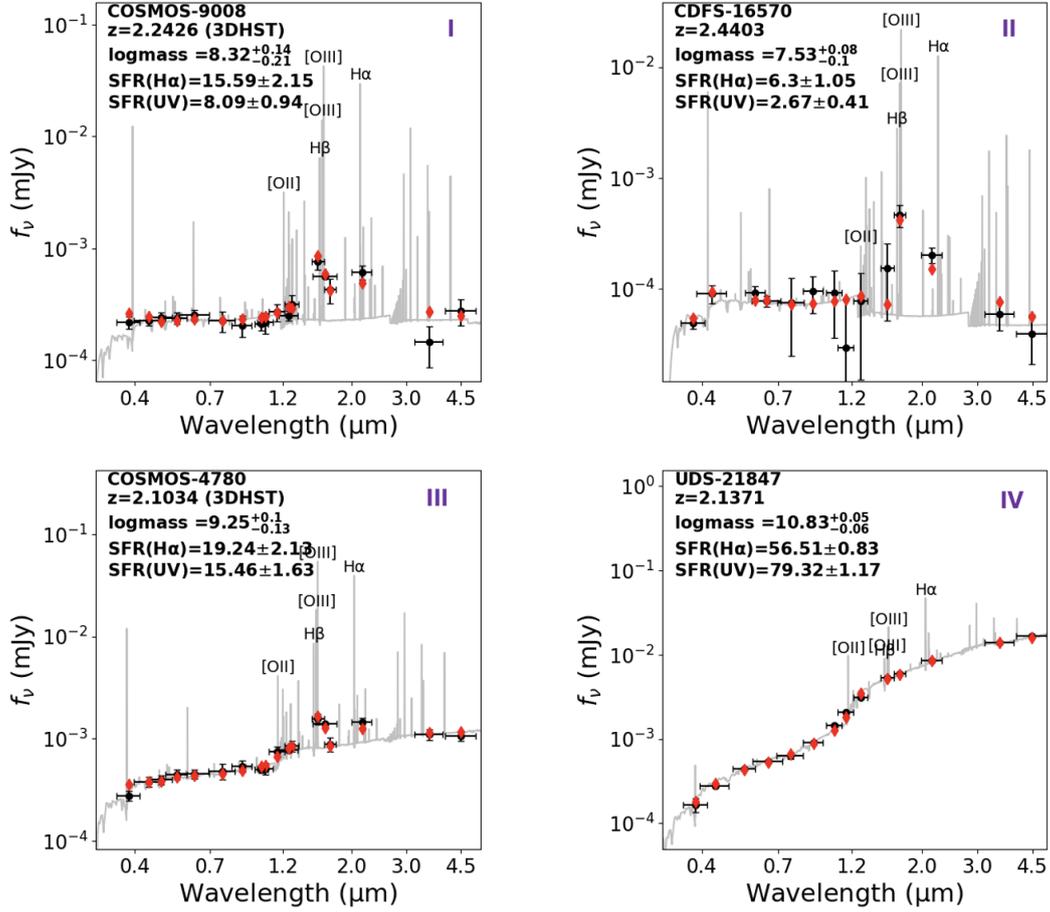
In Table 2.2, we present the comprehensive information and criteria for each emission line used in our analysis.

### 2.4.1 H $\alpha$

In the  $K_s$  filter, the main contaminants include [NII] $\lambda\lambda 6548, 84$  and [SII] $\lambda\lambda 6717, 31$ . We refer to the MOSDEF Emission-Line Catalog (Kriek et al. 2015) to estimate typical emission line ratios at  $2.05 < z < 2.5$ . Initially, we remove galaxies that show non-detection of [NII], [SII] and  $[\text{NII}]\lambda 6584 / \text{H}\alpha > 0.5$  in the MOSDEF catalog. The later criterion is introduced because such strong [NII] emission is unlikely to be associated with star formation and could indicate the presence of AGN hosts (BPT diagram; Kauffmann et al. 2003). This selection yields a sample of 453 objects.



**Figure 2.2.** Upper: The star formation rate ( $\text{H}\alpha$ ) as a function of stellar mass, in the ZFOURGE fields.  $\text{SFR}(\text{H}\alpha)$  is derived from the calibration in [Kennicutt & Evans \(2012\)](#). Grey circles show all the galaxies selected from flux excesses in  $K_s$  photometry. Each galaxy is given the error of stellar mass from SED fitting and the error of  $\text{SFR}(\text{H}\alpha)$  from  $K_s$  photometry. Those  $3\sigma$  upper limits for the  $\text{H}\alpha$ -undetected sample have downward arrows. Magenta diamonds are median values from 6 mass bins with the median uncertainty on them. The best-fit  $M_*$ -SFR relation from [Whitaker et al. \(2014\)](#) are also shown as black dotted lines. Bottom: Same as the upper panel, but only show the individual galaxies without errors. The parent sample are clearly separated by an  $S/N$  threshold of 3, that blue circles are those with sufficient flux excesses.



**Figure 2.3.** Four example SEDs from the three ZFOURGE fields are given in the small panels. Each is corresponding to the purple circle in Figure 2.2. Among them, COSMOS-9008 have redshift measurement from MOSDEF (Kriek et al. 2015) and COSMOS-4780 have redshift measurement from 3D-HST (Momcheva et al. 2016). Black circles are observed fluxes in several photometric filter sets and red diamonds are the best-fit SEDs convolved by the filter transmission curves. The grey spectrum is the best-fit SED based on CIGALE. Note that several optical median-band filter sets we used in the SED fitting are not shown here.

**Table 2.3.** Parent sample and the number of emission line galaxies with different threshold

S/N threshold	# of HAEs	# of [OIII] emitters w/ H $\alpha$ <sup>a</sup>	# of [OII] emitters (w/ H $\alpha$ ) <sup>b</sup>
Parent sample <sup>c</sup>	3754	3754	3754
S/N > 2	2029	1003 (859)	1061 (824)
S/N > 3	1318	541	588
S/N > 5	562	233	212
S/N > 10	110	48	26

**Notes.** <sup>a</sup> The galaxies contain [OIII] emission line with S/N > 2 and H $\alpha$  emission lines with S/N > 3, corresponding to the main criteria for selecting emission line galaxies as in Table 2.2.

<sup>b</sup> The galaxies contain [OII] with S/N > 2 and H $\alpha$  with S/N > 3.

<sup>c</sup> The galaxies with use=1 in the ZFOURGE catalog (AGN excluded). In this stage, we do not apply any S/N cutoff for emission lines.

From this sample, we obtain the line ratios,  $H\alpha/(H\alpha+[NII]+[SII])$  and take their average to obtain  $0.67 \pm 0.10$ . Finally, we adopt  $r_{H\alpha} = 0.7$ . To further validate this assumption, we also compute the emission line ratio from the best-fit SED of the entire HAE sample, and the resulting median value yields  $r_{H\alpha} \simeq 0.78$ . Thus, we believe that adopting  $r_{H\alpha} = 0.7$  introduces only minor systematic errors. From the model spectrum, we could also obtain the purity ratio of each galaxy. However, the purity ratios from best-fit model spectra show almost no correlation to the spectroscopic measurements.

[SII] falls outside the  $K_s$ -band for 121 (9%) galaxies at  $2.45 < z < 2.5$ , for which the purity ratios  $r_{H\alpha}$  become larger. However, because of the  $\sim 2\%$  uncertainty of  $z_{\text{phot}}$  in our study, we cannot identify which galaxies are exactly falling within the redshift range above. Thus, we apply a constant  $r_{H\alpha}$  regardless of the  $z_{\text{phot}}$  of the galaxies.

To identify H $\alpha$  emitters in the  $K_s$ -band filter, we implement a selection process based on flux excesses and photometric errors. We select candidates of HAEs by requiring the flux excesses in the  $K_s$ -band exceed three times the photometric errors ( $> 3\sigma$ ), that is,

$$F_{\text{excess}, K_s} > 3 \times \Delta f_{K_s} \times \Delta \lambda. \quad (2.6)$$

This criterion yield a sample of 1,318 H $\alpha$  emitters (422, 463, 433 in GOODS-S, COSMOS, UDS, respectively) at  $z_{\text{med}} = 2.25$ . In Figure 2.2, we display the complete sample of galaxies from the ZFOURGE catalog on the  $M_*$ -SFR diagram, including galaxies with flux excesses below  $3\sigma$ . Since broad-band flux excess is much harder to be detected compared to narrow-band, an S/N > 3 cutoff is already sufficient to capture many galaxies with high equivalent widths. In Table 2.3, we provide a count of the number of emission line galaxies at various S/N threshold. Among them, some sample with very extremely intense emission lines exhibit significant S/N exceeding 10 in BB or MB filters.

Additionally, we show four examples of galaxies with their corresponding best-fit SEDs. Notably, the proximity between the observed flux and the model flux reinforces the robustness of our fitting methodology. The two low-mass examples

show a distinct characteristic on the rest-frame optical to near-IR bands, displaying a flat continuum. This feature signifies the presence of young stellar populations. Also, these two galaxies exhibit a noticeable excess flux in the  $K_s$ -band, which is boosted by the strong  $H\alpha$  emission line.

In order to validate the reliability of the emission lines obtained from the SED fitting results, we perform a comparison between the observed  $H\alpha$  fluxes ( $F_{H\alpha}$ ) in our work and the slit-loss-corrected fluxes obtained from the MOSDEF Emission Line Catalog ( $F_{\text{spec}}$ , Kriek et al. 2015). Reddy et al. (2015) has introduced the slit-loss corrections of the MOSDEF survey by modeling the HST light profile of each galaxy, resulting in a slit-loss within 18%. This comparison is conducted for a total of 78 galaxies in the ZFOURGE-COSMOS field, all of which have  $H\alpha$  detections with an  $S/N > 3$ , according to both the MOSDEF catalog and our method. Figure 2.4 presents the comparison of  $H\alpha$  fluxes between our method and the MOSDEF catalog. We find that 63 out of the 78 (81%) of  $H\alpha$  emitters exhibit consistent flux values within a factor of 2, demonstrating agreement between the two datasets. This analysis further strengthens the confidence in the emission line measurements derived from our SED fitting approach.

## 2.4.2 [OIII] and [OII]

For galaxies at  $z \sim 2.3$ , [OIII] emission lines would drop in either the  $H_s/H_l$  filter. To ensure accurate measurements of [OIII] line fluxes, we assume that the total flux excesses are contaminated by  $H\beta$ . Building upon this, we adopt a Case-B recombination with  $T_e = 10,000K$  and  $n_e = 100 \text{ cm}^{-3}$ . This allows us to derive the intrinsic  $H\beta$  fluxes from the intrinsic  $H\alpha$  fluxes using the following relation,

$$F_{H\beta,int} = \frac{F_{H\alpha,int}}{2.86}. \quad (2.7)$$

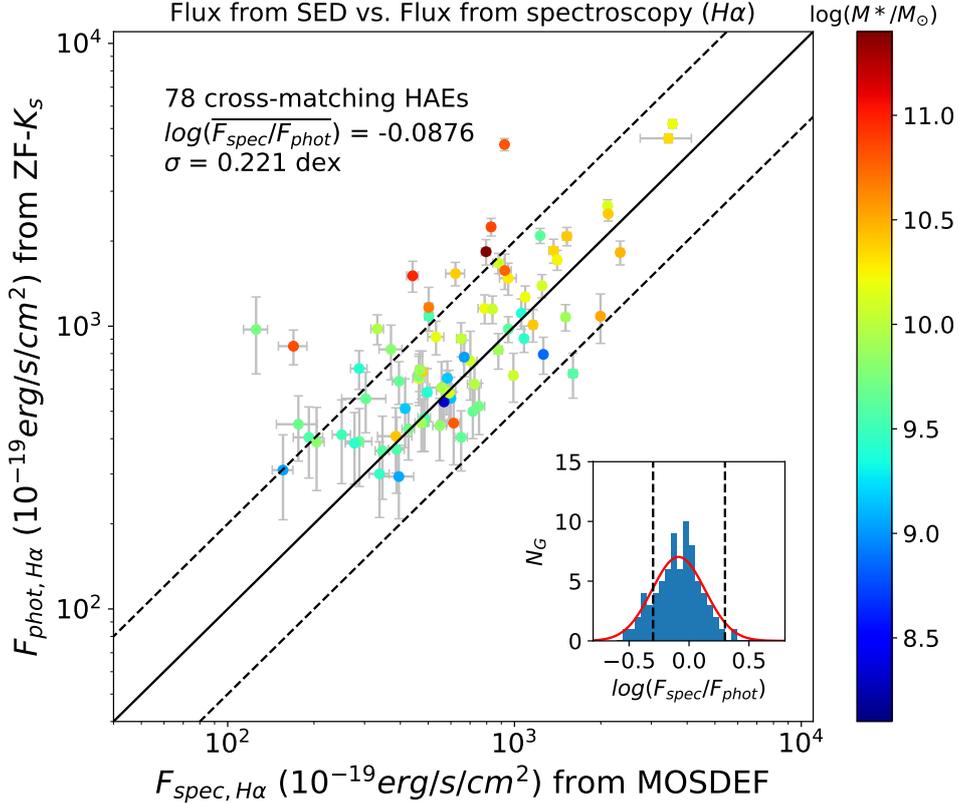
Here,  $F_{H\alpha,int}$  and  $F_{H\beta,int}$  are corrected for dust extinction from the observed fluxes  $F_{H\alpha,obs}$  and  $F_{H\beta,obs}$ . By subtracting the observed flux  $F_{H\beta,obs}$  from the total flux excesses, we can obtain the [OIII] emission line fluxes.

$H\beta$  falls outside the  $H$ -band medium filter for 221 (16%) galaxies at  $2.27 < z < 2.33$ . Again, due to the  $z_{\text{phot}}$  uncertainties, we do not make adjustment for the  $H\beta$  contamination fraction depending on the redshift of the galaxies.

We expect [OII] emission lines at  $z \sim 2.3$  to be detected in either the  $J_2/J_3$  filters. We estimate the contamination in a similar manner as for  $H\alpha$ . The main sources of contamination include [NeIII] $\lambda\lambda 3870, 3969$  and Balmer lines such as  $H\epsilon \lambda 3970$ . Likewise, we refer to the MOSDEF catalog to obtain the average purity ratio and set  $r_{[\text{OII}]} = 0.7$ . Again, the purity ratios from best-fit model spectra show no correlation to the spectroscopic measurements. The contamination lines falls outside the  $J$ -band medium filter for galaxies at  $2.45 < z < 2.5$ , which is same to the case of  $H\alpha$ . We still apply a constant  $r_{[\text{OII}]}$  when deriving the [OII] emission.

To demonstrate the reliability of the emission line fluxes, we also compare the [OIII] and [OII] emission line fluxes with those obtained from the MOSDEF catalog and 3D-HST catalog. Figure 2.5 and Figure 2.6 present these comparisons, revealing a consistency with the spectroscopic measurements.

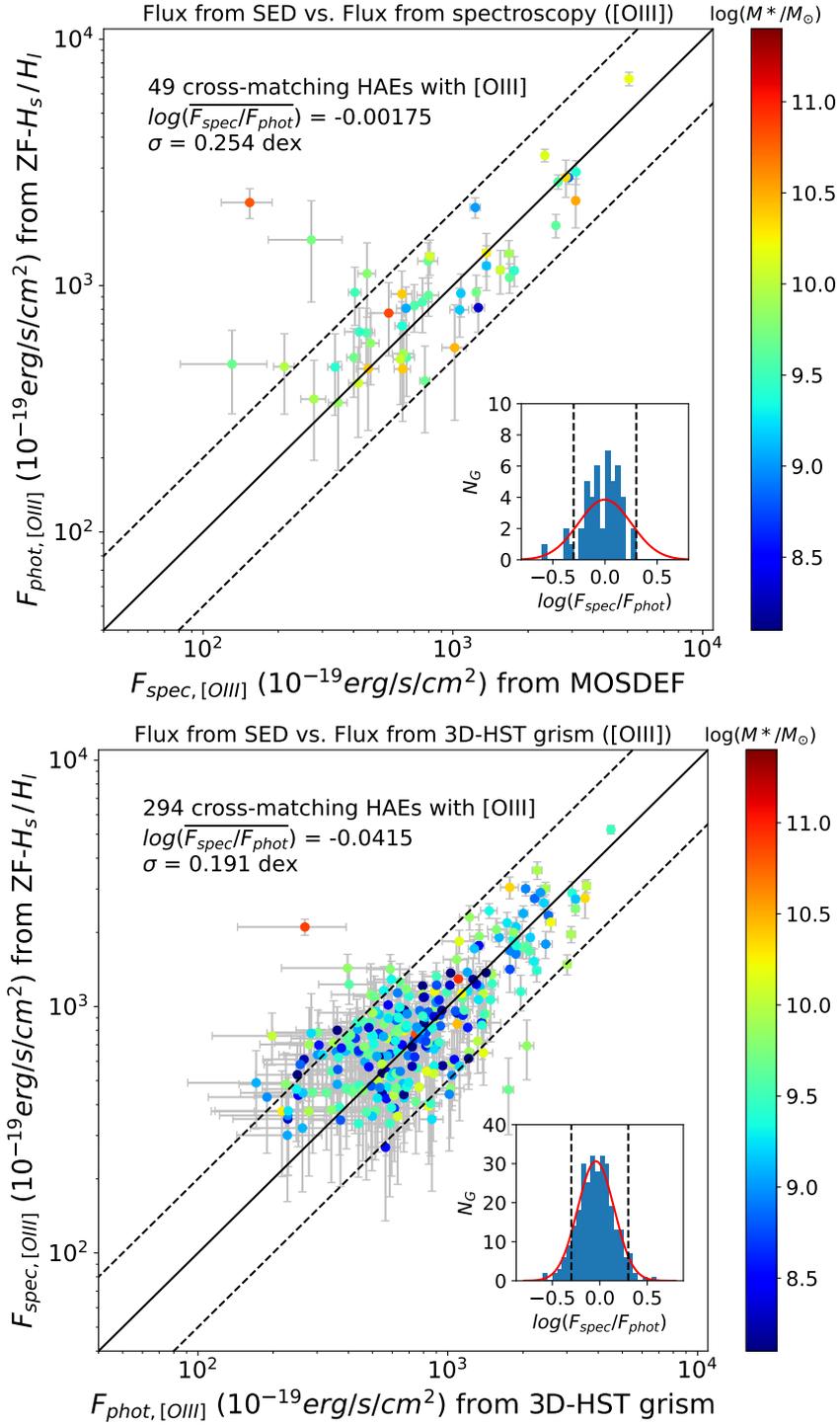
In conclusion, our derivation of the  $H\alpha$ , [OIII], [OII] fluxes demonstrate no significant systematic biases (within 0.1 dex offset) and small scatter within 0.3 dex with those obtained from spectroscopic surveys, particularly for galaxies with relatively low masses down to the limits of the MOSDEF and 3D-HST survey. This suggests that the emission line fluxes estimated based on the flux excesses are robust and reliable.



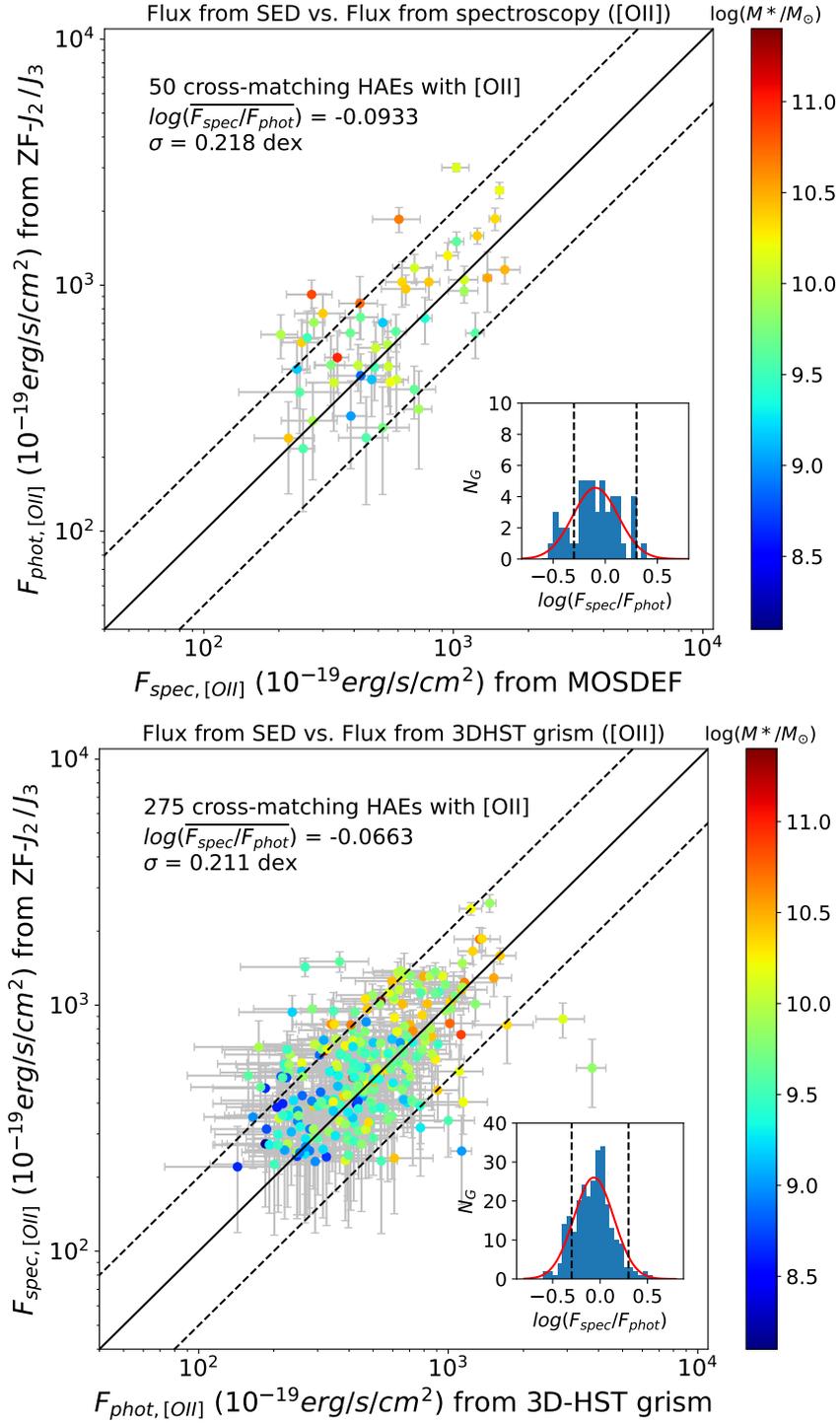
**Figure 2.4.** Comparison between the observed  $H\alpha$  fluxes derived from SED fitting and those from the MOSDEF spectroscopic emission-line catalog (Kriek et al. 2015). Black dashed line indicates an agreement with a factor of 2. The error bars on the  $y$ -axis are the flux errors from the MOSDEF catalog, while those on the  $x$ -axis is from our method. Galaxies are separated into histograms on the lower right panel according to their residual from a 1:1 line. Here, the flux ratios ( $F_{spec}/F_{phot}$ ) are scaled to  $\log_{10}$ -space with steps of 0.05 and black dashed lines (a factor of 2) are added. The mean difference and scatter on this one-to-one relation are also added. The color gradient of dots shows the stellar mass of individual galaxy. The overall estimation of emission line fluxes agrees well with the spectroscopic measurements, indicating the robustness of our method.

## 2.5 Rest-Frame Equivalent Widths

Rest-frame  $EW_{H\alpha}$ ,  $EW_{[OIII]}$  and  $EW_{[OII]}$  are calculated by dividing the line flux by the continuum flux density at a certain rest-frame wavelength. The continuum flux density is determined using the best-fit SED model through the following process. Firstly, we exclude the flux density points that contain both continuum



**Figure 2.5.** Upper: Comparison between the observed [OIII] $\lambda\lambda 4959, 5007$  fluxes derived from SED fitting and those from the MOSDEF spectroscopic emission-line catalog (Kriek et al. 2015). Bottom: Same for the observed [OIII] fluxes but comparing with the grism spectra catalog from the 3D-HST survey (Momcheva et al. 2016). Plot details as in Figure 2.4. The estimates agree, suggests our method works, even at low masses.



**Figure 2.6.** Upper: Comparison between the observed [OII] fluxes derived from SED fitting and those from the MOSDEF Emission-Line Catalog (Kriek et al. 2015). Bottom: Same for the observed [OII] fluxes but comparing with the grism spectra catalog from the 3D-HST survey (Momcheva et al. 2016). Plot details as in Figure 2.4. Again, agreement indicates our method is robust.

and emission from the model. Next, we fit the continuum flux density points to a power-law slope,  $f_\nu \propto \lambda^\alpha$  within the wavelength windows of  $(\lambda_0 - 100) \times (1 + z)$  to  $(\lambda_0 + 100) \times (1 + z)$ , where  $\lambda_0$  represents the rest-frame wavelength of the emission line in Angstrom. Finally, the fitted continuum at  $\lambda_0 \times (1 + z)$  is taken as the desired continuum flux density. Note that in following chapters,  $EW_{[\text{OIII}]}$  refers to the combined equivalent widths of the  $[\text{OIII}]\lambda\lambda 4959, 5007$  doublet.

## 2.6 Emission Line Ratios

Based on our measurements, the available emission line diagnostics in our study are O32 and R23. Note that, we derive the intrinsic  $\text{H}\beta$  fluxes from the intrinsic  $\text{H}\alpha$  fluxes, as explained in Section 2.4.2. To obtain the intrinsic O32 and R23 values, the related emission lines are corrected for dust attenuation,  $E(B - V)_{neb}$ , which is obtained from the Bayesian SED fitting result.

Note that, the definition of  $[\text{OIII}]$  in these two diagnostics is slightly different; O32 uses  $[\text{OIII}]\lambda 5007$ , while R23 uses  $[\text{OIII}]\lambda\lambda 4959, 5007$ . We use the line ratio of  $[\text{OIII}]\lambda 5007 : [\text{OIII}]\lambda 4959 = 2.97 : 1$  for the conversion.

We have compiled a catalog of  $\text{H}\alpha$  emitters for the three ZFOURGE fields. The catalog includes a total of 1,318 HAEs with redshifts ranging from 2.05 to 2.5. Each entry in the catalog provides information such as coordinates, observed emission line fluxes, flux uncertainties, and SED-derived properties including stellar mass, stellar age, and dust attenuation. The identification of individual sources in the catalog is based on their unique ID, which corresponds to the ZFOURGE catalog. Within the full sample of HAEs, 859 sources have a detection of  $[\text{OIII}]$  emission lines, while 824 sources have a detection of  $[\text{OII}]$  emission lines. Additionally, there are 626 HAEs that have both  $[\text{OIII}]$  and  $[\text{OII}]$  emission lines detected. This catalog provides a comprehensive dataset for studying HAEs, along with multiple emission lines, at  $z \sim 2.3$ .

## 2.7 Luminosity Function

The luminosity function characterizes the number density of galaxies as a function of their luminosity. It tells us how common or rare galaxies of a given brightness are in the universe. In order to derive the luminosity function of the HAEs, we first calculate the volume and number density of our survey. Considering the redshift range  $2.05 < z < 2.5$  and the total coverage of ZFOURGE survey (see section 2.1), we obtain that our  $\text{H}\alpha$  survey probes a (comoving) volume of  $\Delta V = 6.8 \times 10^5 \text{ Mpc}^3$ . Then number density in each luminosity bin is calculated as follows,

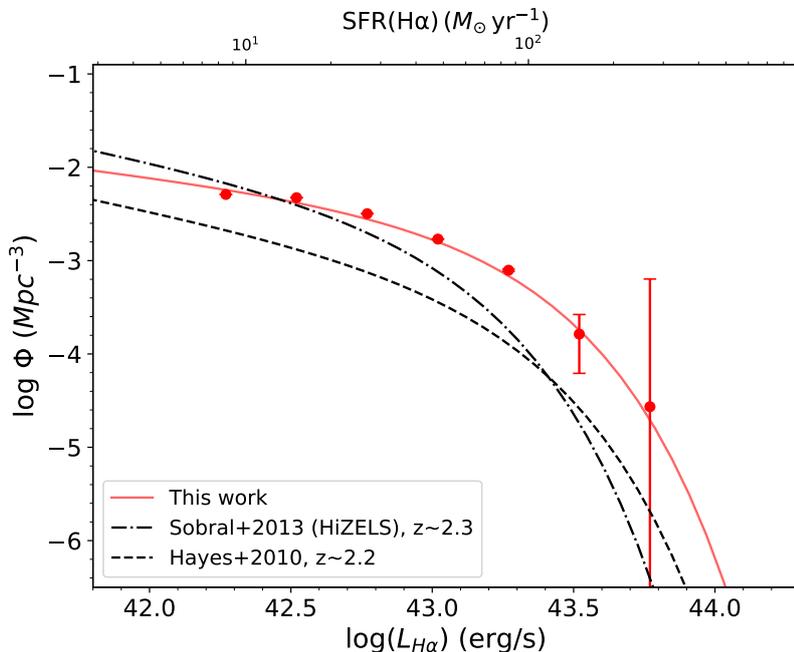
$$\phi^*(\log(L_c)) = \frac{1}{\Delta(\log L)} \sum_{|\log \frac{L_i}{L_c}| < \frac{\Delta(\log L)}{2}} \frac{1}{\Delta V}, \quad (2.8)$$

where  $\log(L_c)$  is the central luminosity in each bin,  $\Delta(\log L)$  is the step of each bin and  $\log(L_i)$  is the  $\text{H}\alpha$  luminosity of each galaxy in  $\log_{10}$  space. In this study, we set

the bin width to  $\Delta(\log L) = 0.25$ , with  $\log(L_c)$  ranging from 42.0 to 44.0, where the sample achieves more than 90% completeness for galaxies with  $L_{\text{H}\alpha} > 10^{42} \text{ erg s}^{-1}$ . Based on these prerequisites, we determine the best-fit H $\alpha$  luminosity function of our HAEs at  $z_{\text{med}} = 2.25$ . The luminosity function is fitted using Schechter functions (Schechter 1976), which are defined by three parameters,  $\alpha$ ,  $\phi^*$  and  $L^*$ . In the  $\log_{10}$  space, the Schechter function is given by,

$$\phi(L) dL = \ln 10 \phi^* \left( \frac{L}{L^*} \right)^\alpha e^{-(L/L^*)} \left( \frac{L}{L^*} \right) d\log L, \quad (2.9)$$

where  $\phi^*$  is the normalization density,  $L^*$  is a characteristic galaxy luminosity where the power law form of the function cuts off, and  $\alpha$  is the power law slope at low luminosity. The Schechter function is fitted to each bin of  $\log(L_c)$ , and the best-fit results for the H $\alpha$  luminosity function are presented in Figure 2.7 and Table 2.4.



**Figure 2.7.** The H $\alpha$  luminosity function of our sample. The red dots and curve illustrates the best-fit luminosity function when removing the AGNs. The best-fit curve from Sobral et al. (2013) and Hayes et al. (2010a) are also included for comparison, which are shown by the black dashed-dotted curve and the black dashed curve, respectively. The parameters of the Schechter function for each best-fit curve with are summarized in Table 2.4.

We also include two previous studies for comparison. Sobral et al. (2013) derived the H $\alpha$  luminosity function from SFGs in the High-redshift(Z) Emission Line Survey (HiZELS, A narrow-band imaging survey; Geach et al. 2008; Sobral et al. 2009). Similarly, Hayes et al. (2010a) used the VLT to obtain extremely deep narrow-band (NB2090) and broad-band ( $K_s$ ) imaging observations in GOODS-S field.

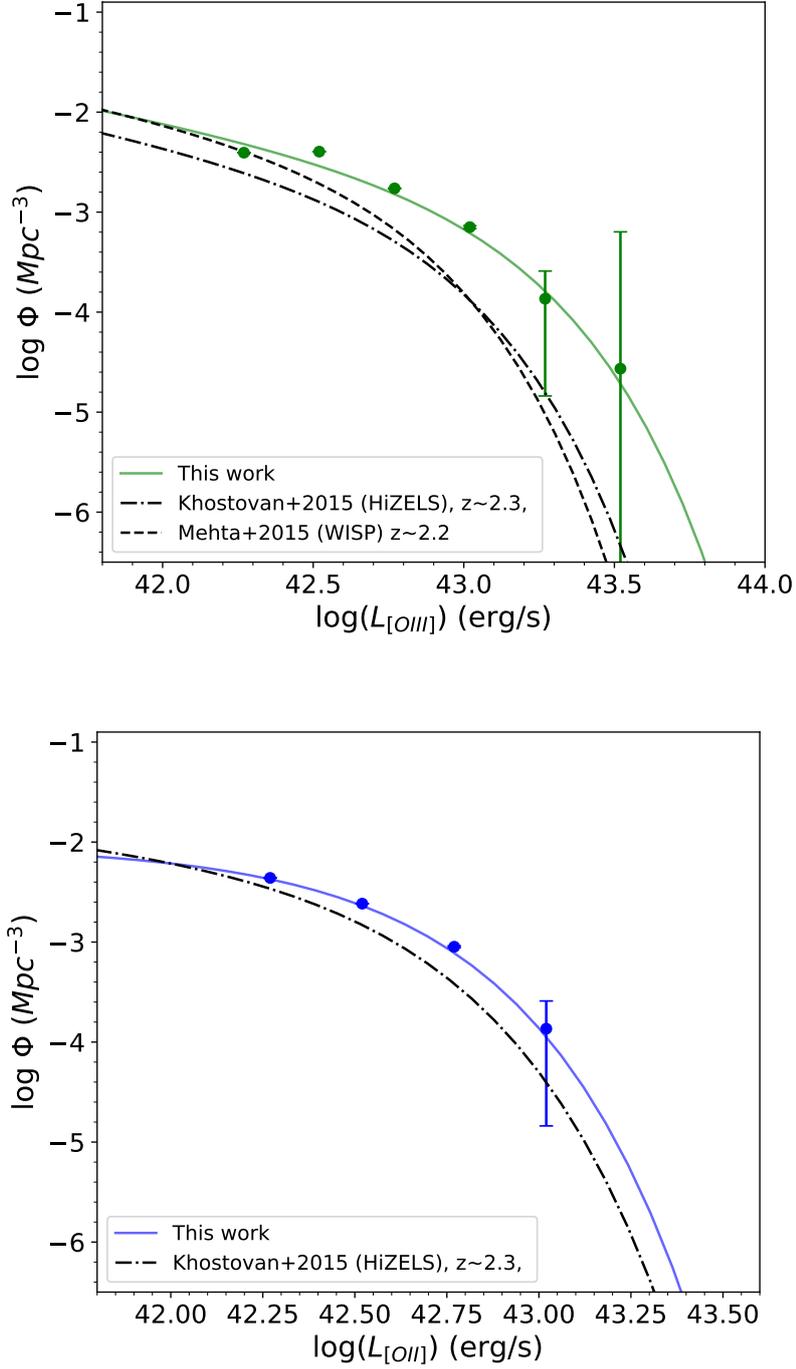
Interestingly, the bright end of the H $\alpha$  luminosity function in our study exceeds that of previous works, suggesting the presence of more H $\alpha$ -luminous galaxies than previously reported. But still, these luminous galaxies remain rare, highlighting the

need for larger surveys and spectroscopic follow-ups to further investigate this unique population. On the other hand, the ZFOURGE survey has a smaller coverage and total volume compared to other wide-field surveys, such as HiZELS. This limitation may introduce a larger uncertainty in the normalization density  $\phi^*$  of the luminosity function. Hayes et al. (2010a) fitted the H $\alpha$  luminosity function in GOODS-S field and obtain a set of Schechter parameters with similar  $L^*$ , but a lower  $\phi^*$  than our work. It is reasonable to infer that differences in the luminosity function are primarily driven by survey volume bias. The ZFOURGE field may contain a relatively larger number of brighter H $\alpha$  emitters, leading to an excess at the bright end of the luminosity function.

We perform the same fitting for the [OIII] and [OII] luminosity functions, comparing our results with those from Khostovan et al. (2015) and Mehta et al. (2015) in Figure 2.8. Similarly, we find a larger number of bright [OIII] ([OII]) emitters compared to previous studies. While, the [OII] luminosity function shows better agreement with the literature.

**Table 2.4.** The best-fit parameters of emission line luminosity functions at  $z \sim 2.3$

The emission line	$\log L^*$ (erg s $^{-1}$ )	$\log \phi^*$ (Mpc $^{-3}$ )	$\alpha$
H $\alpha$	43.10	-2.84	-1.35
[OIII]	42.91	-2.94	-1.57
[OII]	42.44	-2.47	-1.30



**Figure 2.8.** The [OIII] ([OII]) luminosity function of our sample. The green (blue) dots and curve illustrates the best-fit luminosity function. The best-fit curve from [Khostovan et al. \(2015\)](#) and [Mehta et al. \(2015\)](#) are also included for comparison, which are shown by the black dashed-dotted curve and the black dashed curve, respectively. The parameters of the Schechter function for each best-fit curve with are summarized in [Table 2.4](#).



# Chapter 3

## A population of EELGs at Cosmic Noon

In the previous chapter, we introduced the selection of HAEs based on the excess in the observed  $K_s$  broadband flux relative to the stellar continuum estimated from the best-fit SED. Interestingly, we observe a large number of HAEs scattered above the SFMS with a median offset  $\Delta\text{MS}_{\text{med}} \sim 0.3$  dex, below the mass completeness limit ( $10^9 M_\odot$ ). A similar population has also been reported in [Hayashi et al. \(2016\)](#) and [Terao et al. \(2022\)](#), characterized by high, starburst-like star forming activities. Our analysis further reveals that high equivalent widths in [OIII] and H $\alpha$  are prevalent in these low-mass galaxies, suggesting a new population of extreme emission line galaxies at Cosmic Noon.

### 3.1 SFRs and Star Formation Main Sequence

In this work, we quantify the SFRs of galaxies by two indicators: H $\alpha$  and FUV ( $1500\text{\AA}$ ,  $L_{1500}$ ) luminosities. For the UV luminosity measurement, galaxies are required to be detected in ZFOURGE  $B$  and  $V$  filter, which yields a cut of 2.5% of total HAEs. We compute the observed UV continuum ( $L_{1500, \text{uncor}}$ ) and the UV slope ( $\beta_{\text{UV}}$ ) over a rest-frame wavelength range of  $1400 - 2800\text{\AA}$  by performing a multi-band fitting to broad-band photometry with the relation of  $f_\lambda \propto \lambda^{\beta_{\text{UV}}}$ . For ZFOURGE-COSMOS field, the fitting includes broad-band photometry of  $B, G, V, R, R_p, I, Z, Z_p$ . For ZFOURGE-UDS and ZFOURGE-GOODS-S field, the fitting includes broad-band photometry of  $B, V, R, I, Z$ . The amount of dust attenuation is measured through the Bayesian SED fitting, providing the dust extinction  $E(B - V)_{\text{star}}$ ,  $E(B - V)_{\text{neb}}$ . We parameterize the difference between  $E(B - V)_{\text{star}}$  and  $E(B - V)_{\text{neb}}$  by a factor  $f = 0.8$  ([Saito et al. 2020](#)). Once we obtain the intrinsic UV luminosity ( $L_{1500, \text{cor}}$ ) and intrinsic H $\alpha$  luminosity ( $L_{\text{H}\alpha, \text{cor}}$ ), the star formation rates are converted by using the calibration of [Kennicutt & Evans \(2012\)](#) with a correction to the [Chabrier \(2003\)](#) IMF.

In Figure 3.1, we present the the SFMS of H $\alpha$ , UV for the 1318 HAEs from the catalog. The low-mass ( $< 10^9 M_\odot$ ) HAEs are denoted as blue circles, while the median SFR in six mass bins is represented by large open circles. The mass bins are defined as follows:  $\log(M_*/M_\odot) < 8.5$  for the first bin,  $\log(M_*/M_\odot) >$

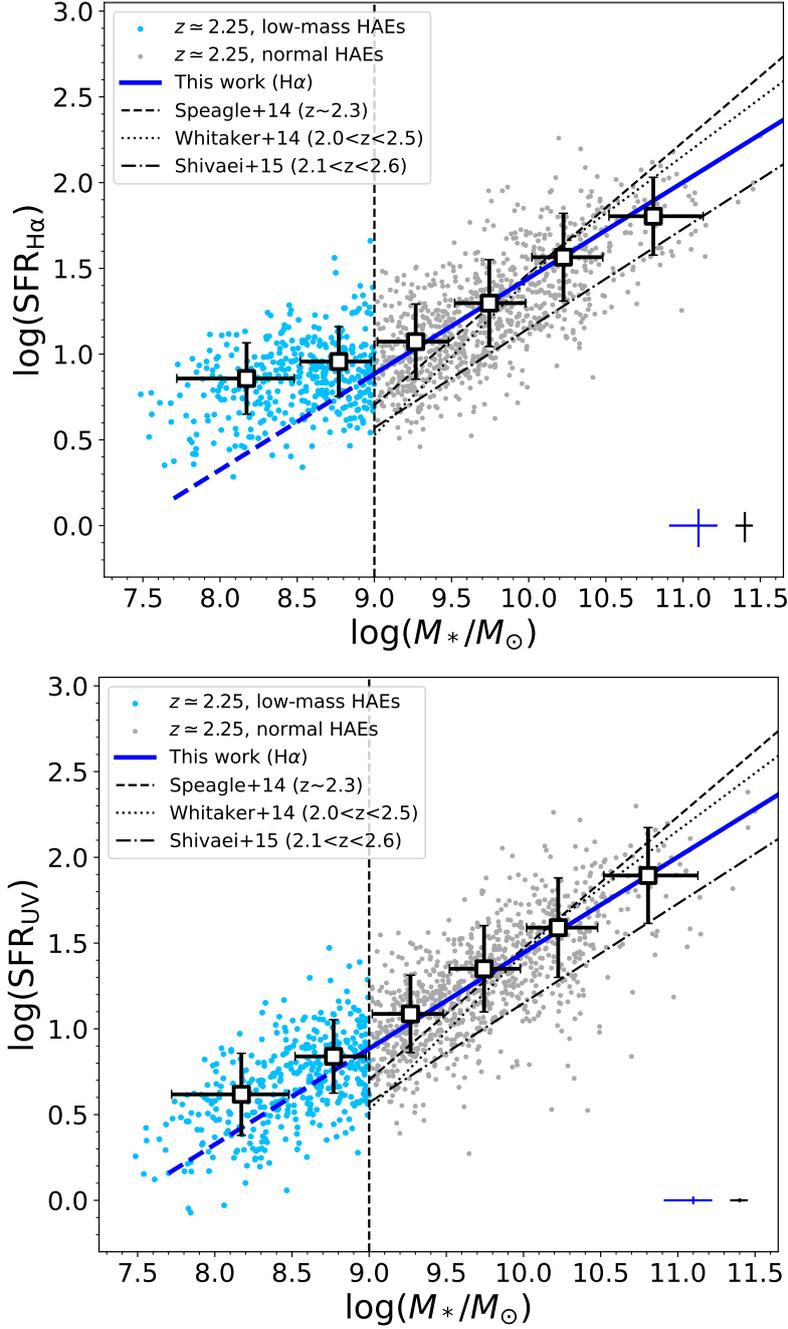
10.5 for the last bin, and the rest are divided into 0.5 dex widths. We apply the linear least squares regression to the SFR and stellar mass data points, above the mass completeness  $\log(M_*/M_\odot) = 9.0$  (Straatman et al. 2016). The best-fit linear correlation between SFRs and stellar masses, with the 68% confidence interval on slope and intercept, is given by,

$$\begin{aligned}\log \text{SFR}(\text{H}\alpha) &= (0.56 \pm 0.03) \times \log M_* - (4.15 \pm 0.28); \\ \log \text{SFR}(\text{UV}) &= (0.60 \pm 0.04) \times \log M_* - (4.47 \pm 0.36).\end{aligned}\tag{3.1}$$

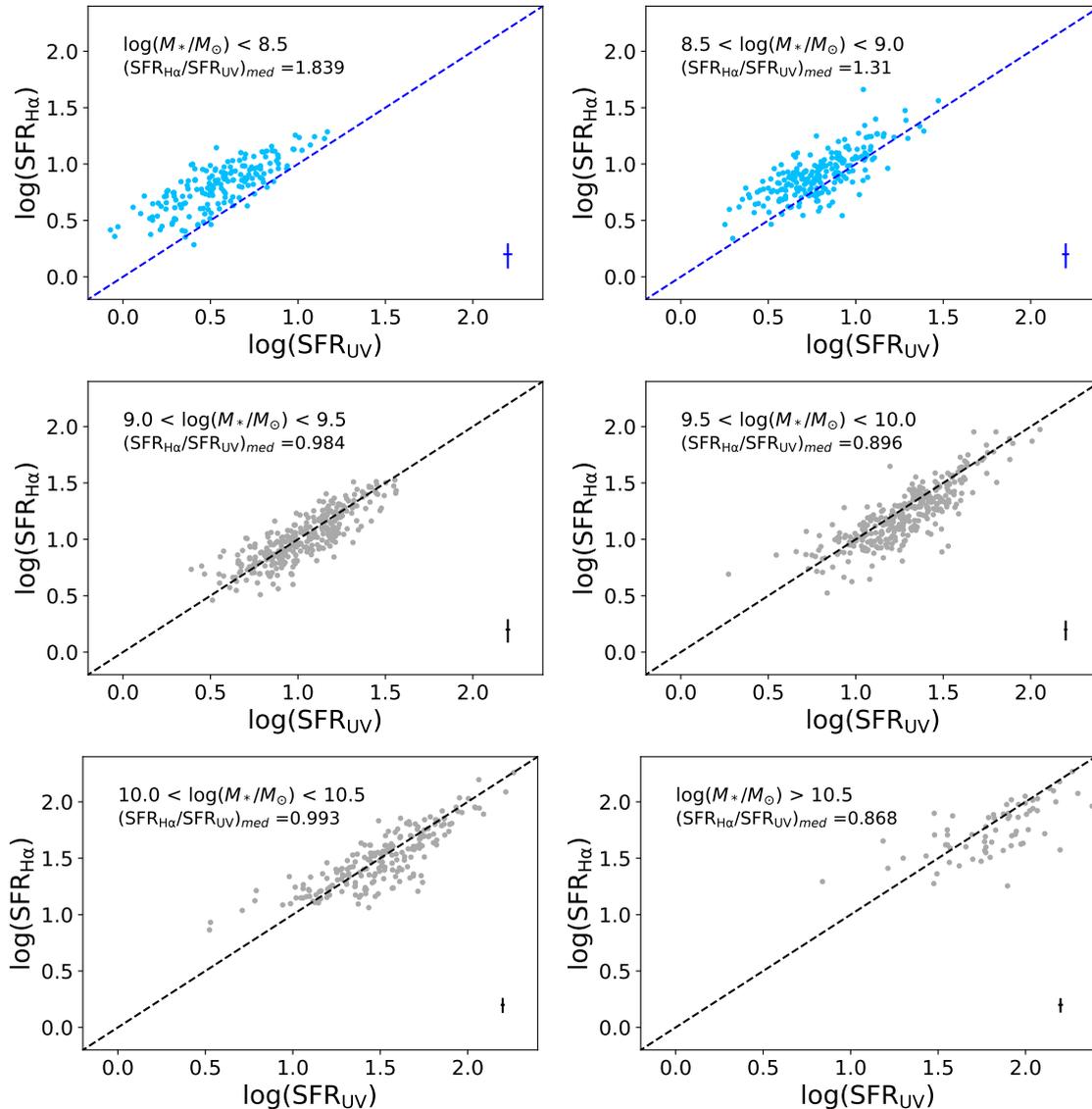
The slope in Equation 3.1 is different from that in Whitaker et al. (2014), which has a slope of 0.91 at the low-mass end. Shivaei et al. (2015) has reported that the slope of the SFMS can be influenced by various observational and measurement factors. These factors can contribute to the discrepancies in the slopes reported in different studies, as shown in Figure 3.1. In our study, the sample biases primarily arise from the selection criterion of HAEs. In Figure 3.1, all galaxies are included without the  $3\sigma$  requirement, where we observe a good agreement between the data points and the extrapolated SFMS from Whitaker et al. (2014). Additionally, if we apply a  $2\sigma$  criteria for the HAEs selection (in Table 2.2), the resulting SFMS slope is found to be  $0.69 \pm 0.03$ . These observations suggest that the selection criterion and its associated sample biases have a notable impact on the derived slope of the SFMS, emphasizing the need to carefully account for such effects when interpreting and comparing results across different studies.

We extrapolate the SFR –  $M_*$  relation from Equation 3.1 into low-mass domain ( $< 10^9 M_\odot$ ) in Figure 3.1. In the bottom panel of Figure 3.1, the low-mass HAEs are found to be closer to the SFMS of UV, showing only a slight elevation in SFR(UV) compared to the SFMS by an average of 0.05 dex. However, a significant fraction of low-mass HAEs lie above the SFMS of H $\alpha$ , exhibiting an average SFR(H $\alpha$ ) higher than the SFMS by 0.25 dex. To further illustrate this trend, we plot the SFR(H $\alpha$ ) vs. SFR(UV) ratio for our sample in six mass bins in Figure 3.2. In the high-mass domain ( $> 10^9 M_\odot$ ), the ratio of these two SFR indicators is mainly close or below unity. In contrast, for galaxies with  $\log(M_*/M_\odot) < 9.0$ , the ratio is clearly above unity. We divide the HAEs in our study into two populations: 401 low-mass HAEs with  $\log(M_*/M_\odot) < 9.0$  and 917 high-mass HAEs with  $\log(M_*/M_\odot) > 9.0$ .

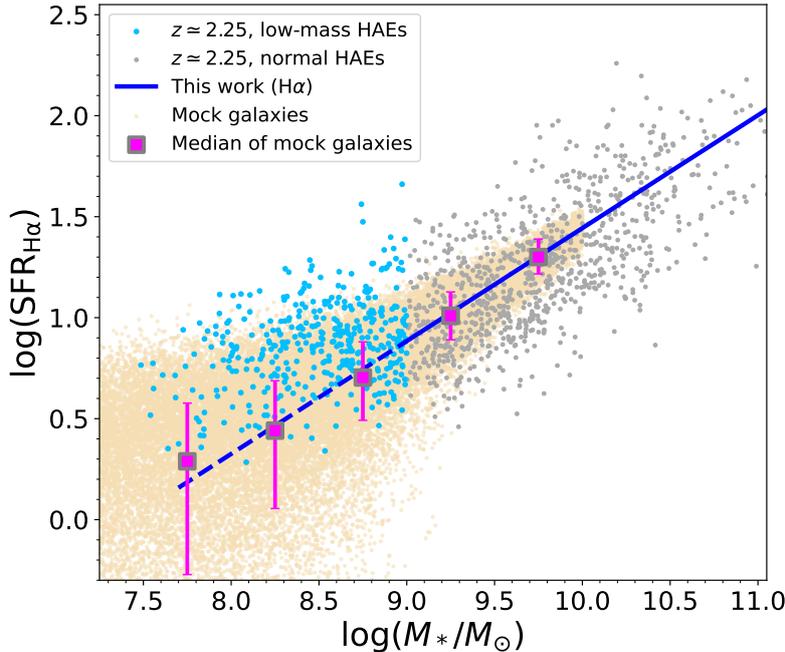
It is important to note that only galaxies with strong H $\alpha$  emission lines can be detected in the low-mass end due to the sensitivity limit of the  $K_s$  band in ZFOURGE. Before investigating their SFHs, we assess the impact of uncertainties in the emission line fluxes. Since faint galaxies are more significantly affected by observational uncertainties, it is possible that the observed high sSFRs in low-mass galaxies could be influenced by errors in their H $\alpha$  flux measurements. To evaluate this effect, we generate 50,000 mock galaxies with stellar masses in the range  $\log(M_*/M_\odot) = 7.0 - 10.0$ . Their SFRs are computed using the first equation in 3.1. Then we introduce fluctuations in the SFRs by incorporating observational uncertainties from the Ks-band flux errors and stellar continuum, following Equation 2.5. In Figure 3.3, we present these mock galaxies as yellow dots and derive the intrinsic scatter of the mock galaxies across different stellar mass bins, represented by magenta bars in Figure 3.3. Based on the statistical analysis of the mock galaxies, we suggest that the observed excess in the SFMS exceeding  $> 0.5$  dex cannot



**Figure 3.1.** Based on the flux excesses in  $K_s$  photometry, we define H $\alpha$  emitters (HAEs) if they show  $> 3\sigma$  H $\alpha$  detection in  $K_s$  photometry. It is obvious that those low-mass HAEs ( $< 10^9 M_\odot$ ) tend to scatter above the SFMS(H $\alpha$ ) but not obvious in SFMS(UV). Upper: The star formation main sequence (SFMS) of 1318 HAEs at  $z_{\text{med}} = 2.25$  in the ZFOURGE fields based on H $\alpha$  emission line. Blue circles are 401 low-mass HAEs with  $\log(M_*/M_\odot) < 9$ , while grey circles show the other HAEs in our catalog. Open squares are median stacks in six mass bins, while the error bars on them represent the scatter in each mass bin. Blue solid line is the best linear fit to the galaxies with  $\log(M_*/M_\odot) > 9.0$ , which is extrapolated to lower mass with blue dashed line. The best-fit SFMS from Whitaker et al. (2014), Speagle et al. (2014) and Shivaei et al. (2015) are also shown with black dotted, dashed, and dot-dashed lines, respectively. The error bars on the bottom-right corner represent the median uncertainty of low-mass HAEs (blue) and high-mass HAEs (black). Bottom: Same as the upper panel, but SFR of each galaxy are calculated from the UV continuum. SFR(H $\alpha$ ) and SFR(UV) are corrected for dust attenuation using the Cardelli/Calzetti curve.



**Figure 3.2.** Relation between the SFR( $\text{H}\alpha$ ) and SFR(UV) of the HAEs in our sample. SFR( $\text{H}\alpha$ ) and SFR(UV) are corrected for dust attenuation using the Cardelli/Calzetti curve. The color reflects different mass bins adopted in our study, as in Figure 3.1. More clearly, we find the low-mass HAEs ( $< 10^9 M_\odot$ ) have the SFR( $\text{H}\alpha$ )/SFR(UV) ratio close to a factor of 2, indicating their characteristic galaxy properties.



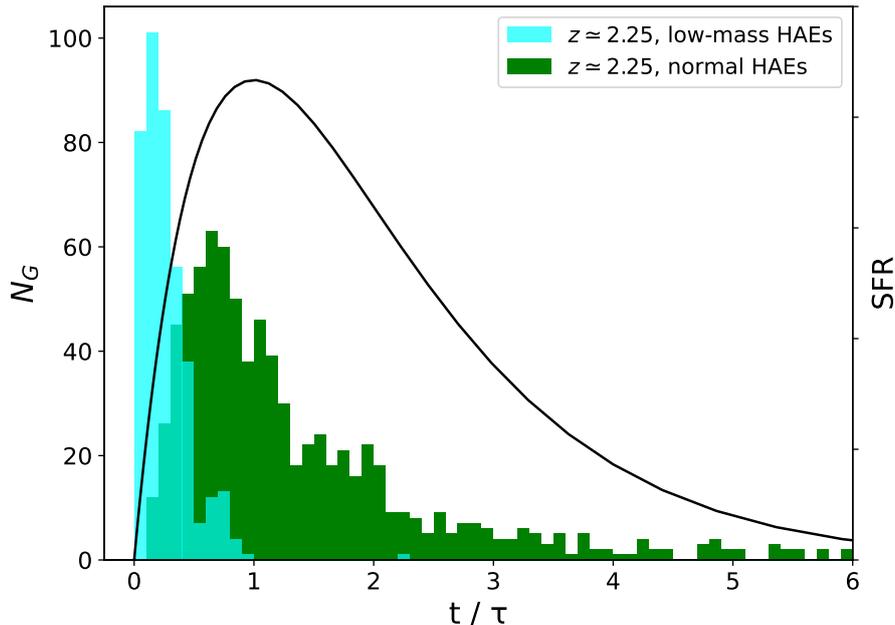
**Figure 3.3.** SFMS plot same as the upper panel of Figure 3.1. We generate 50,000 mock galaxies shown as gray dots, with the scatter in certain stellar mass bins as magenta bars. The HAEs that largely scatter above the SFMS ( $> 0.5$  dex) are unlikely to be caused by uncertainties in the emission line flux measurements.

be merely attributed to uncertainties in the emission line flux measurements. This supports that these galaxies with very high sSFRs exhibit intrinsic differences in their physical properties, rather than being artifacts of observational uncertainties.

## 3.2 Bursty star formation in low-mass galaxies

The comparison of  $\text{SFR}(\text{H}\alpha)$  and  $\text{SFR}(\text{UV})$  directly visualizes the burstiness of star formation activity. Previous studies (e.g., Weisz et al. 2012; Domínguez et al. 2015; Emami et al. 2019) have explored the time evolution of the  $\text{H}\alpha$ -to-UV ratios in variety types of star formation history (SFH) models. Generally, SFR indicators have different timescales, with  $\text{SFR}(\text{UV})$  having a timescale of  $\sim 100$  Myr and  $\text{SFR}(\text{H}\alpha)$  having a timescale of  $\sim 10$  Myr. Emami et al. (2019) has found that in galaxies undergoing rising star formation, the luminosity ratio between  $\text{H}\alpha$  and UV is expected to be higher than unity, and  $\text{SFR}(\text{H}\alpha)$  will reside above the SFMS by up to an order of magnitude. High-resolution hydrodynamical simulations by Sparre et al. (2017) have also shown that the scatter on the SFMS is larger for the  $\text{H}\alpha$ -derived SFR because it is more sensitive to short bursts compared to the UV-based indicator.

We explore the SFH of the HAEs in our sample in Figure 3.4. In this study, we apply the delayed- $\tau$  models to model the SFH of galaxies as Equation 2.1. The distribution of the  $t/\tau$  ratio of the HAEs reveals that the low-mass HAEs have a much smaller  $t/\tau$ , indicating that they are in their early stage of star-forming with rising SFRs. On the other hand, the distribution of high-mass HAEs shows more



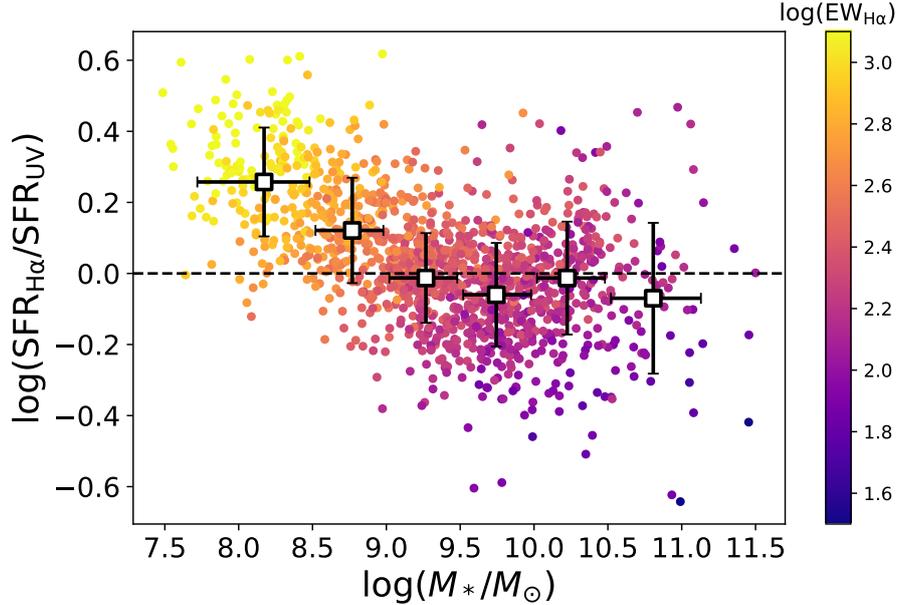
**Figure 3.4.** The star formation history (SFH) of the H&AEs in our sample. SFH of the low-mass H&AEs ( $< 10^9 M_\odot$ ) and high-mass H&AEs ( $> 10^9 M_\odot$ ) are distributed as a histogram of  $t/\tau$  in cyan column and green column. For reference, a model delayed- $\tau$  SFH is shown as black solid line. The SFH suggests that rising star-forming activities are occurring in these low-mass H&AEs.

variability, indicating a range of SFHs in these galaxies. The rising star formation in low-mass galaxies indicates the abundance of young stellar population ( $\leq 100$  Myr) in these systems, which can contribute to intense  $H\alpha$  emission lines.

### 3.3 Equivalent widths in low-mass galaxies

In Figure 3.5, we further plot the ratio of  $SFR(H\alpha)$  and  $SFR(UV)$  as a function of stellar mass, with an additional dimension of the equivalent width of  $H\alpha$  ( $EW_{H\alpha}$ ). It is not surprising that galaxies with high  $EW_{H\alpha}$  exceeding  $1000 \text{ \AA}$  exhibit the highest  $SFR(H\alpha)/SFR(UV)$  ratios. In comparison, galaxies with  $EW_{H\alpha} < 100 \text{ \AA}$  tend to align with the SFMS. To provide a clearer view, we directly present the histogram of  $EW_{H\alpha}$  (and  $EW_{[OIII]}$ ) for H&AEs in Figure 3.6, separating the parent H&AEs into low-mass and high-mass populations. The results reveal a distinct trend that low-mass H&AEs have a much higher median  $EW$  in  $H\alpha$  ( $[OIII]$ ) compared to their high-mass counterparts. Quantitatively, the low-mass H&AEs have a median  $EW_{H\alpha,med} \simeq 624 \text{ \AA}$  and  $EW_{[OIII],med} \simeq 695 \text{ \AA}$ , whereas the high-mass H&AEs have  $EW_{H\alpha,med} \simeq 194 \text{ \AA}$  and  $EW_{[OIII],med} \simeq 152 \text{ \AA}$ . Notably, most of the low-mass H&AEs exhibit  $EW$ s exceeding  $200 \text{ \AA}$ , representing a large population of EELGs at  $z \sim 2.3$  found in this study.

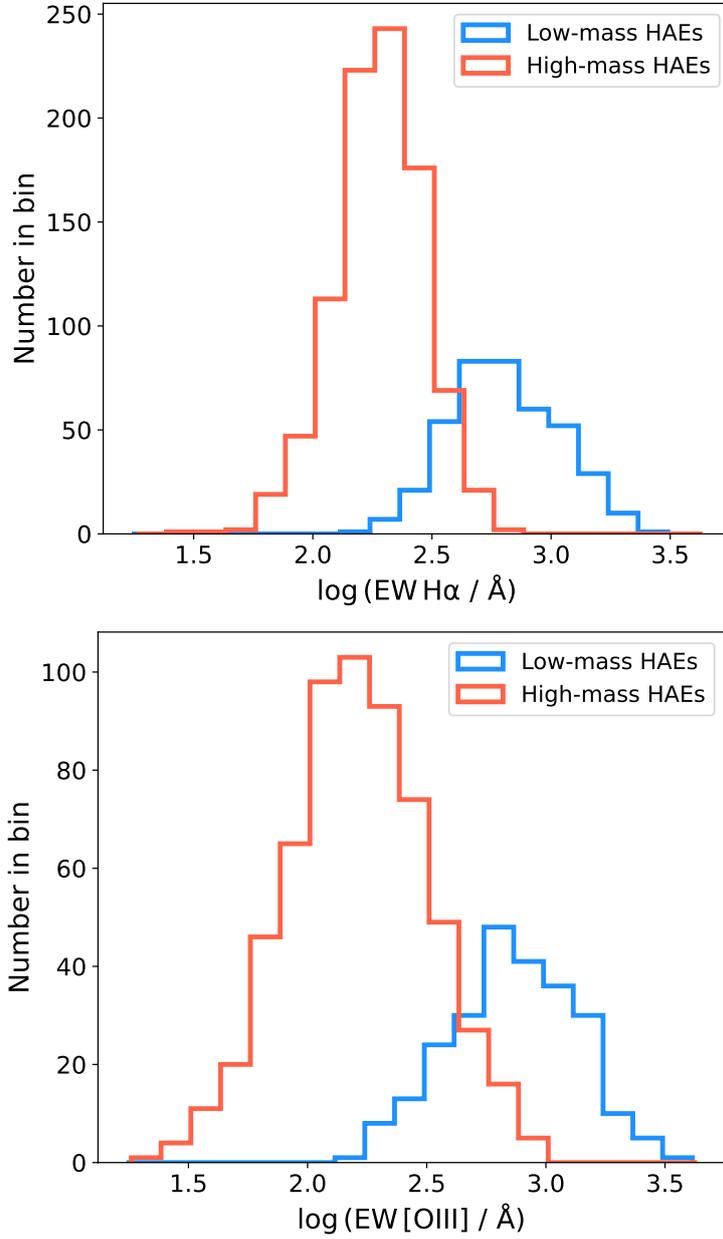
As also shown in Figure 3.5, galaxies with the highest  $EW_{H\alpha}$  ( $> 1000 \text{ \AA}$ ) have the lowest stellar masses in our sample with a mean value of  $\log(M_*/M_\odot) = 8.09$ . Stefanon et al. (2022) conducted a stacking analysis of  $z \sim 8$  galaxies and found that the stacked galaxy exhibits extremely strong emission line, reaching up to



**Figure 3.5.** The ratio of SFR( $\text{H}\alpha$ ) and SFR(UV) as a function of stellar mass. The color gradient of dots shows the  $\text{EW}_{\text{H}\alpha}$  of individual galaxy. Open squares are median stacks in 6 mass bins with the error bar representing the scatter in each mass bin. Galaxies with high  $\text{EW}_{\text{H}\alpha}$  more than  $1000\text{\AA}$  possess the highest SFR( $\text{H}\alpha$ )/SFR(UV) and the lowest stellar masses in our sample.

$\text{EW}_{\text{H}\alpha} \sim 2000\text{\AA}$ . Moreover, the stellar mass estimated from their stacked photometry is  $\log(M_*/M_\odot) = 8.12$ , closely aligning with our results. These similarities suggest that the low-mass HAEs in our study could also serve as analogs of galaxies during the EoR. This connection highlights the potential of studying these low-mass HAEs at Cosmic Noon to gain insights into the physical conditions of galaxies at the very early universe.

Overall, we regard these low-mass HAEs as a large population of EELGs at Cosmic Noon, characterized by their strong emission lines (some with both  $\text{EW}_{\text{H}\alpha} > 200\text{\AA}$  and  $\text{EW}_{[\text{OIII}]}$   $> 200\text{\AA}$ ), and high sSFRs. The discovery of such galaxies reinforces the idea that low-mass systems are not only significant contributors to the ionizing photon budget but also play a crucial role in the assembly of stellar mass across cosmic time. Moreover, the large sample size provides a unique opportunity to statistically study the extreme physical conditions within these galaxies, offering the insights into the processes driving their evolution.



**Figure 3.6.** The distribution of  $\text{H}\alpha$  and  $[\text{OIII}]$  EWs for the parent HAEs are shown as histograms, color-coded in blue and orange based on their stellar masses. The analysis of  $\text{EW}_{[\text{OIII}]}$  includes only those HAEs with  $[\text{OIII}]$  flux excesses having  $S/N > 2$ , and contains a sample of 859 objects.

# Chapter 4

## Multiple Emission Line Analysis of the HAEs

In this chapter, we investigate the relationship between rest-frame equivalent widths and various properties of the HAEs at  $z \sim 2.3$ , especially focusing on those low-mass HAEs. Specifically, we examine their dependence on stellar mass, stellar age, SFR, and sSFR. Additionally, we explore the correlations between equivalent widths and the available emission line index, O32. For comparison, we also consider other galaxy samples from different studies. This includes star-forming and starburst ( $EW_{H\alpha} > 50\text{\AA}$ ) galaxies from the SDSS DR7 MPA/JHU catalog (Kauffmann et al. 2003; Brinchmann et al. 2004), extreme O3Es at  $z \sim 2.2$  (Tang et al. 2019), galaxies from the MOSDEF survey in the mid-redshift windows at  $z \sim 2.3$  (Kriek et al. 2015; Reddy et al. 2018). By comparing the properties and trends among these different samples, we aim to gain a comprehensive understanding of the relationships of these attributes.

Given the substantial sample size, we perform a best-fit linear correlation analysis and report the Spearman’s rank correlation coefficient,  $r_s$ , of these variables in Table 4.1. The Spearman statistical test underscores the robustness of the correlation between equivalent widths ( $\log[EW/\text{\AA}]$ ) and the physical parameters of galaxies.

### 4.1 EW vs. Stellar properties

#### 4.1.1 [OIII] EWs

The [OIII] equivalent widths as a function of physical parameters (stellar mass, age, SFR, sSFR) are shown in Figure 4.1. It should be noted that not all HAEs with [OIII] lines have detections of [OII] lines. This could be due to uncertainties in flux measurements near the Balmer and 4000Å break, where the rest-frame wavelength of the [OII] line is located. Also, galaxies at  $z \sim 2$  often show intrinsic [OIII]/[OII] line ratios greater than unity (e.g., Shapley et al. 2015; Sanders et al. 2016). The weaker [OII] line may not reach the detection limit of our medium-band flux excess. For those HAEs with [OIII] detections but without [OII] detections, we mark them as open circles in Figure 4.1 to differentiate them from HAEs with both detections.

The upper-left panel of Figure 4.1 highlights a clear trend between  $\log(EW_{[\text{OIII}]})$

**Table 4.1.** Relationship between the Equivalent Widths and various Galaxy Properties

Line <sup>a</sup>	Attribute <sup>a</sup>	$N_{gal}$ <sup>b</sup>	Intercept <sup>c</sup>	Slope <sup>c</sup>	$r_s$ <sup>d</sup>
[OIII]	$M_*/M_\odot$	859 (All sample)	$7.321 \pm 0.098$	$-0.527 \pm 0.010$	-0.864
		626 (w/ [OII])	$7.411 \pm 0.142$	$-0.536 \pm 0.015$	-0.818
		233 (w/o [OII])	$7.339 \pm 0.173$	$-0.530 \pm 0.020$	-0.886
	Age/yr	859 (All sample)	$9.322 \pm 0.198$	$-0.819 \pm 0.023$	-0.762
		626 (w/ [OII])	$8.991 \pm 0.283$	$-0.786 \pm 0.033$	-0.685
		233 (w/o [OII])	$8.703 \pm 0.231$	$-0.730 \pm 0.028$	-0.883
	SFR(H $\alpha$ )/ $M_\odot$ yr <sup>-1</sup>	859 (All sample)	$3.027 \pm 0.045$	$-0.550 \pm 0.038$	-0.442
		626 (w/ [OII])	$2.782 \pm 0.054$	$-0.395 \pm 0.042$	-0.352
		233 (w/o [OII])	$3.170 \pm 0.101$	$-0.548 \pm 0.106$	-0.202
	sSFR(H $\alpha$ )/yr <sup>-1</sup>	859 (All sample)	$8.147 \pm 0.116$	$0.701 \pm 0.014$	0.860
		626 (w/ [OII])	$8.162 \pm 0.165$	$0.704 \pm 0.020$	0.806
		233 (w/o [OII])	$7.630 \pm 0.172$	$0.629 \pm 0.022$	0.887
O32	626 (w/ [OII])	$2.347 \pm 0.010$	$1.003 \pm 0.035$	0.749	
[OII]	$M_*/M_\odot$	824 (All sample)	$3.826 \pm 0.098$	$-0.187 \pm 0.010$	-0.510
		626 (w/ [OIII])	$3.965 \pm 0.109$	$-0.204 \pm 0.012$	-0.533
		198 (w/o [OIII])	$3.876 \pm 0.216$	$-0.187 \pm 0.022$	-0.482
	Age/Myr	824 (All sample)	$4.081 \pm 0.167$	$-0.240 \pm 0.019$	-0.400
		626 (w/ [OIII])	$4.551 \pm 0.183$	$-0.297 \pm 0.021$	-0.465
		198 (w/o [OIII])	$3.978 \pm 0.416$	$-0.221 \pm 0.047$	-0.306
	SFR(H $\alpha$ )/ $M_\odot$ yr <sup>-1</sup>	824 (All sample)	$2.210 \pm 0.028$	$-0.151 \pm 0.022$	-0.237
		626 (w/ [OIII])	$2.180 \pm 0.030$	$-0.132 \pm 0.024$	-0.193
		198 (w/o [OIII])	$2.315 \pm 0.066$	$-0.215 \pm 0.050$	-0.363
	sSFR(H $\alpha$ )/yr <sup>-1</sup>	824 (All sample)	$4.202 \pm 0.114$	$0.260 \pm 0.014$	0.524
		626 (w/ [OIII])	$4.328 \pm 0.123$	$0.278 \pm 0.015$	0.574
		198 (w/o [OIII])	$4.424 \pm 0.266$	$0.279 \pm 0.031$	0.449
O32	626 (w/ [OIII])	$2.014 \pm 0.008$	$-0.092 \pm 0.029$	-0.093	
H $\alpha$	$M_*/M_\odot$	1318 (All sample)	$5.867 \pm 0.057$	$-0.365 \pm 0.006$	-0.843
		626 (w/ [OIII], [OII])	$5.477 \pm 0.106$	$-0.326 \pm 0.011$	-0.694
	Age/Myr	1318 (All sample)	$7.253 \pm 0.131$	$-0.564 \pm 0.015$	-0.707
		626 (w/ [OIII], [OII])	$6.806 \pm 0.184$	$-0.522 \pm 0.022$	-0.629
	SFR(H $\alpha$ )/ $M_\odot$ yr <sup>-1</sup>	1318 (All sample)	$2.803 \pm 0.027$	$-0.321 \pm 0.023$	-0.357
		626 (w/ [OIII], [OII])	$2.436 \pm 0.037$	$-0.056 \pm 0.030$	-0.048
	sSFR(H $\alpha$ )/yr <sup>-1</sup>	1318 (All sample)	$7.060 \pm 0.041$	$0.560 \pm 0.005$	0.955
		626 (w/ [OIII], [OII])	$6.826 \pm 0.060$	$0.536 \pm 0.007$	0.937
	O32	626 (w/ [OIII], [OII])	$2.383 \pm 0.009$	$0.324 \pm 0.033$	0.365

**Notes.** <sup>a</sup> All attributes are calculated as the log scale to exhibit the correlation with  $\log(EW/\text{\AA})$  for the line listed leftmost. <sup>b</sup> The HAEs in our study are separated into sub-sample based on whether they have detection of other lines (w/) or not (w/o). <sup>c</sup> Intercept and slope are obtained from the best-fit linear relationship between the galaxy properties and the equivalent width. <sup>d</sup> The Spearman's rank correlation coefficient of the galaxy properties and the equivalent width.

and  $M_*$ , confirming a trend that has also been observed in previous studies such as Reddy et al. (2018) and Tang et al. (2019). This trend is interpreted as an anti-correlation between equivalent widths and the stellar continuum flux density. In galaxies with higher equivalent widths, the contribution of the stellar continuum to the total flux is reduced. Since the stellar continuum is closely related to stellar mass, it is not surprising that our sample demonstrates such an anti-correlation. When comparing our best-fit results with those from MOSDEF (Reddy et al. 2018), we observe a very similar slope and intercept. On the other hand, our study successfully extends this relation to the lower mass domain, around  $\sim 10^8 M_\odot$ . This new finding highlights the prevalence of high [OIII] equivalent widths in low-mass galaxies at high redshift, which was not previously well-documented.

It is also found that the stellar age obtained from SED fitting has an anti-correlation to  $\log(\text{EW}_{[\text{OIII}]})$  in the upper-right panel of Figure 4.1. This observation is consistent with tests conducted on photoionization models, which have shown a strong correlation between  $\text{EW}_{[\text{OIII}]}$  and stellar age in starburst events, i.e., simple stellar populations (Stasińska & Leitherer 1996). According to these tests,  $\text{EW}_{[\text{OIII}]}$  can decrease by more than two magnitudes within a time span of  $10^7$  yr, even more than recombination lines. Therefore,  $\text{EW}_{[\text{OIII}]}$  can serve as proxies for the ratio of the current rate of star formation and the past integrated SFR (Reddy et al. 2018). The HAEs with the largest  $\text{EW}_{[\text{OIII}]}$  in our sample are likely in the early stage of star formation, with rapidly rising SFRs. This rapid increase in SFRs is accompanied by higher  $\text{EW}_{[\text{OIII}]}$ , as we have observed. On the other hand, the star formation history (SFH) of older ( $> 10^{8.5}$  yr) HAEs is more complex, resulting in a wider distribution of  $\text{EW}_{[\text{OIII}]}$  values and the gradual breakdown of the linear relation as the stellar age increases.

In the bottom-left panel of Figure 4.1, we examine the relationship between SFR and  $\log(\text{EW}_{[\text{OIII}]})$ . While previous studies by Reddy et al. (2018) and Tang et al. (2019) did not find strong variations in SFR with  $\log(\text{EW}_{[\text{OIII}]})$ , our results suggest a possible anti-correlation between two variables. This discrepancy may be attributed to sample selection biases across different studies. Reddy et al. (2018) focused on the MOSDEF sample, which primarily consists of galaxies with stellar masses larger than  $\sim 10^{9.5} M_\odot$ . If we restrict our analysis to galaxies in our sample with  $M_* > 10^{9.5} M_\odot$ , the best-fit result yields a slope of  $-0.06$ , which is consistent with the one by Reddy et al. (2018) (see Appendix B). This result is significantly different from the slope obtained when considering the full sample ( $-0.55$ ). The SFMS( $\text{H}\alpha$ ) in Figure 3.1 further supports this explanation, as lower-mass HAEs tend to scatter above the SFMS and exhibit higher  $\text{EW}_{[\text{OIII}]}$  values compared to main-sequence galaxies with similar SFR. This discrepancy contributes to the steeper slope observed when fitting the full sample. Tang et al. (2019) specifically selected extremely O3Es with  $\text{EW}_{[\text{OIII}]} > 200 \text{ \AA}$ , which excludes galaxies with lower EW even in the high-mass domain. Also, the sample size in Tang et al. (2019) is smaller compared to ours, potentially leading to biases in their analysis. Although our results demonstrate a discrepancy, the relationship between  $\log(\text{EW}_{[\text{OIII}]})$  and SFR is generally less significant compared to other parameter combinations, supported by the Spearman’s rank correlation coefficient.

The bottom-right panel of Figure 4.1 shows the correlation between sSFR and

$\log(\text{EW}_{[\text{OIII}]})$ . We observe a strong correlation where HAEs with higher sSFRs tend to have higher  $\text{EW}_{[\text{OIII}]}$  values. The  $[\text{OIII}]$  luminosity is known to be an indicator of SFR (Maschietto et al. 2008), and the  $[\text{OIII}]$ -calibrated SFR is consistent with the UV-measured SFR for high-redshift emitters (e.g., Suzuki et al. 2015). On the other hand, the continuum luminosity scales with  $M_*$ . Thus, it is not surprising to observe a good correlation between  $\text{EW}_{[\text{OIII}]}$  and sSFR. Our results further support the notion that sSFR can serve as a useful indicator for  $\text{EW}_{[\text{OIII}]}$  and can potentially be applied even at higher redshifts. In addition, our sample is located in the extrapolated regions of the sequence derived from the SDSS sample. The discrepancy in the locations of these samples is likely attributed to differences in galaxy properties. At high redshift, the molecular gas fraction become larger (Geach et al. 2011). The increasing fraction of molecular gas along with the redshift leads to the evolution of SFMS (e.g., Speagle et al. 2014; Whitaker et al. 2014), that high-redshift galaxies are having larger sSFRs. Because the  $\text{EW}_{[\text{OIII}]}$  also have dependence on sSFR, it will follow the same trend and our sample is also evolving along the vertical axis of the panel.

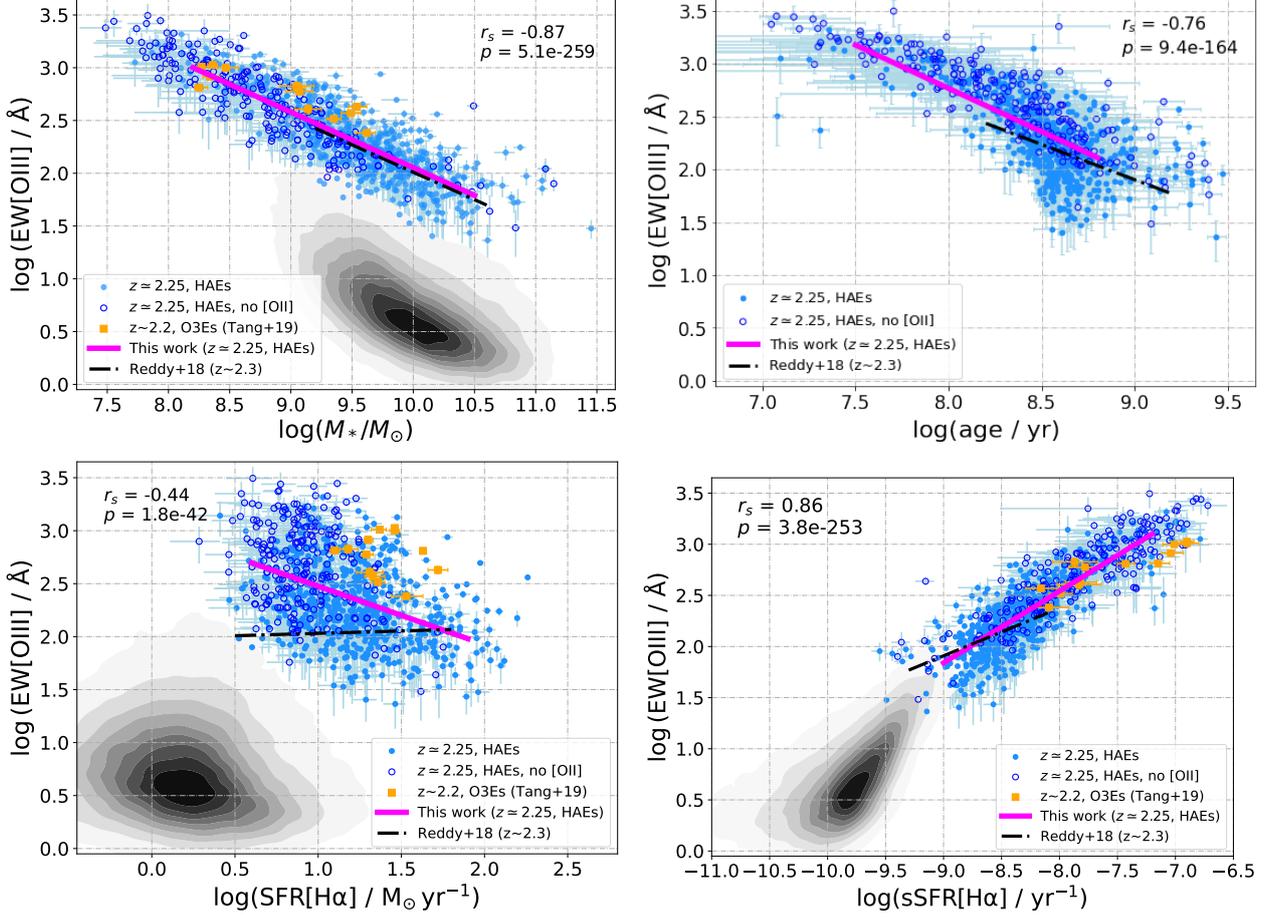
#### 4.1.2 $[\text{OII}]$ EWs

In Figure 4.2 and Table 4.1, we present the  $[\text{OII}]$  equivalent widths as a function of various physical parameters such as stellar mass, age, SFR, and sSFR. Similar to the  $[\text{OIII}]$  equivalent widths, some HAEs have  $[\text{OII}]$  detections but no  $[\text{OIII}]$  detections, and we distinguish them by marking them as open circles in Figure 4.2.

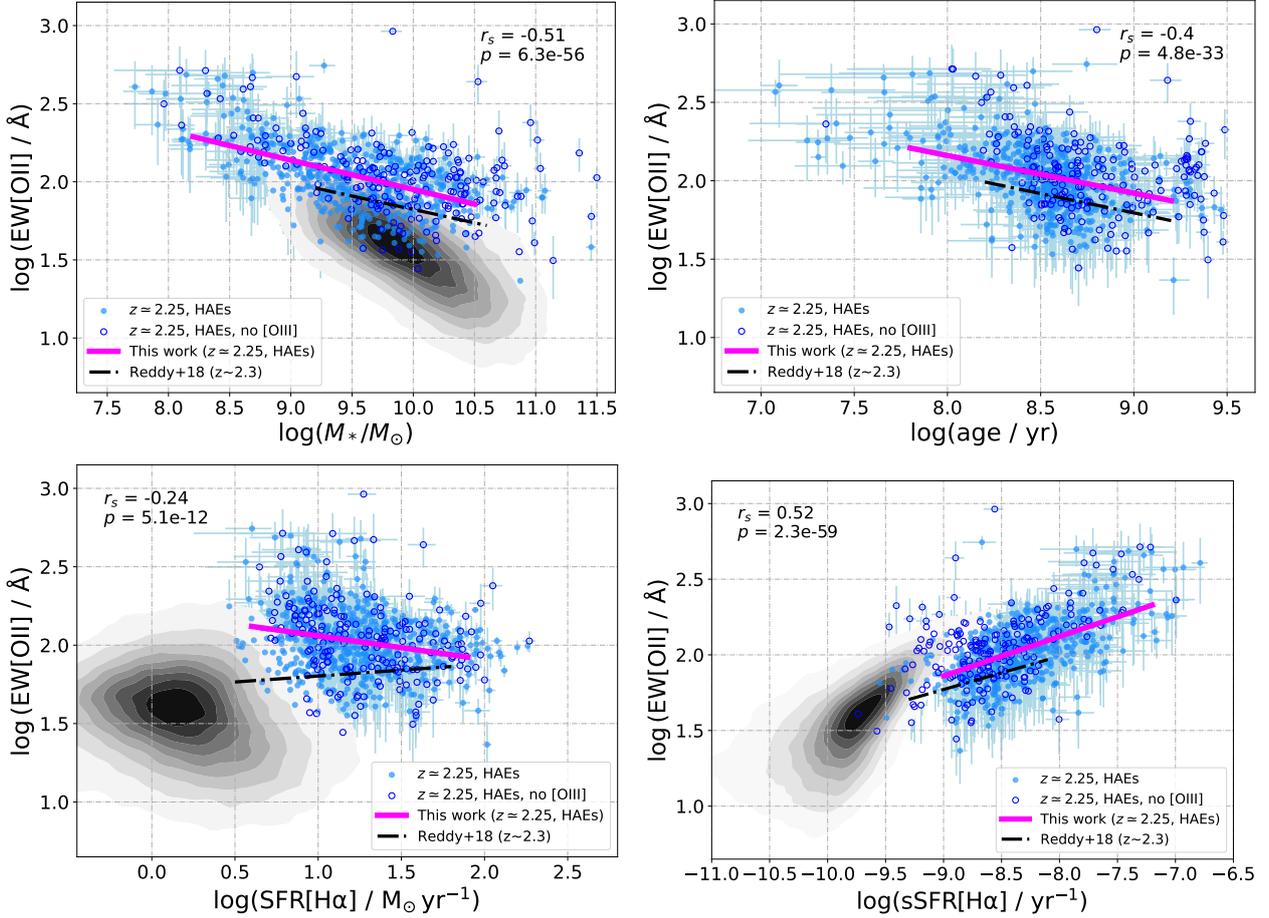
While the  $[\text{OII}]$  equivalent widths do show correlations with stellar mass, stellar age, and sSFR, these relationships are generally weaker compared to the  $[\text{OIII}]$  equivalent widths, as indicated by the Spearman’s rank correlation coefficient. This observation is consistent with the findings of Reddy et al. (2018), who also reported the least significant correlations for the  $[\text{OII}]$  equivalent width among various emission lines, including  $[\text{OIII}]$ ,  $\text{H}\beta$ ,  $\text{H}\alpha$ .

The weaker dependence of  $[\text{OII}]$  equivalent widths on these attributes can be attributed to intrinsic differences between  $[\text{OII}]$  and  $[\text{OIII}]$ . In extreme interstellar medium (ISM) environments, neutral oxygen atoms are more likely to be excited to doubly ionized oxygen ( $\text{O}^{++}$ ) rather than singly ionized oxygen ( $\text{O}^+$ ). Consequently, stronger  $[\text{OIII}]$  emission lines are more commonly observed at higher redshifts. This leads to a significant number of low-mass galaxies exhibiting extremely high  $[\text{OIII}]/[\text{OII}]$  ratios due to lower metallicity and higher ionization parameters (e.g. Cardamone et al. 2009; Erb et al. 2010; Richard et al. 2011; Nakajima et al. 2020). Consistent with this inference, in the upper-left panel of Figure 4.1, galaxies with  $[\text{OIII}]$  but without  $[\text{OII}]$  emission are predominantly low-mass galaxies ( $< 10^9 M_\odot$ ), while in Figure 4.2, galaxies with  $[\text{OII}]$  but without  $[\text{OIII}]$  emission are more concentrated in the massive galaxy regime ( $> 10^{9.5} M_\odot$ ).

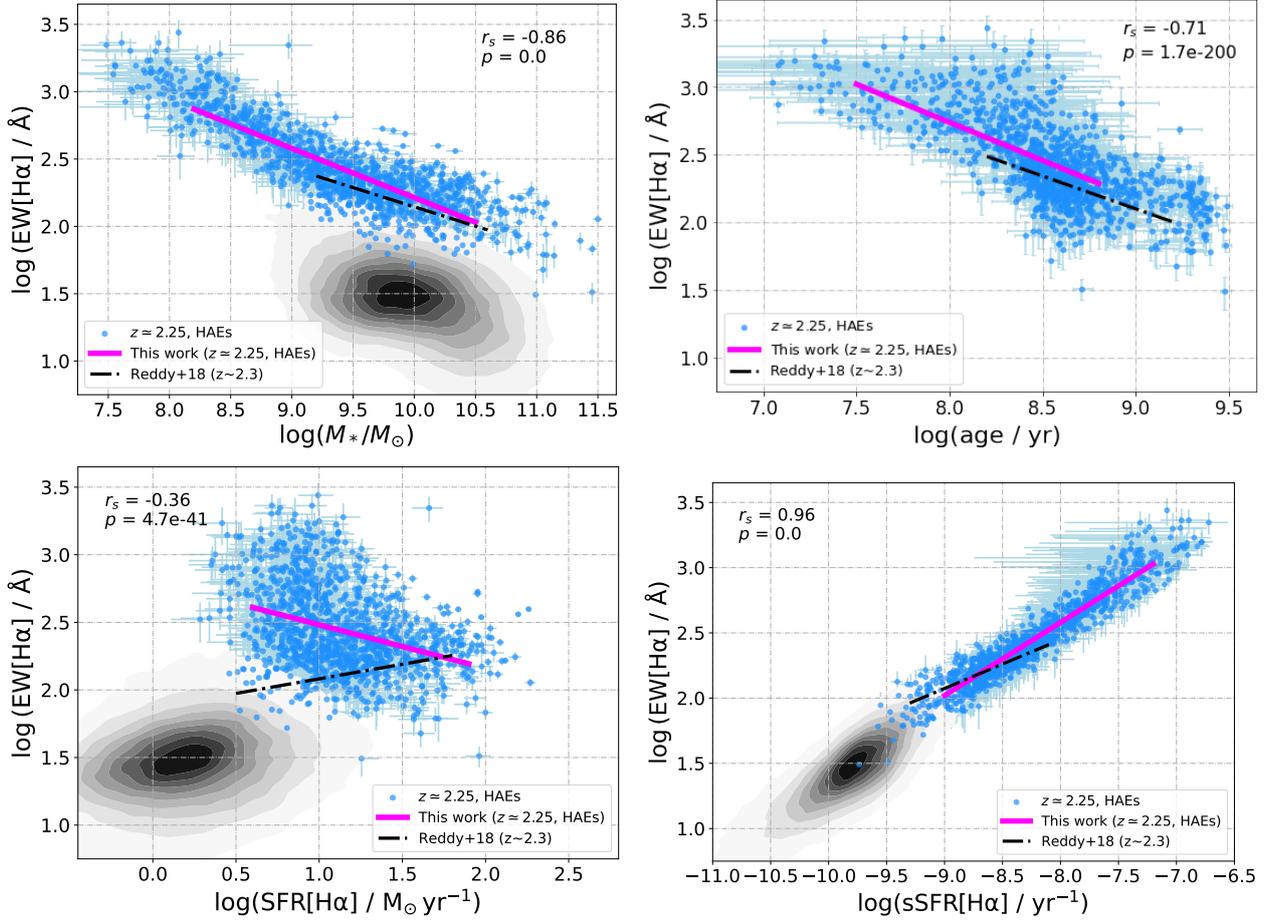
Nonetheless, when comparing our sample with the SDSS sample, we find that the  $[\text{OII}]$  equivalent widths of galaxies at  $z \sim 2.3$  are relatively larger than those of local galaxies. This suggests that the  $[\text{OII}]$  equivalent widths still exhibit some dependence on the star-formation activities in galaxies, despite the weaker correlations with other physical parameters.



**Figure 4.1.** Relationship between the  $\text{EW}_{[\text{OIII}]}$  and stellar properties (stellar mass, age, SFR, sSFR) of the 859 HAEs at  $z \sim 2.3$  in our study. The stellar mass and stellar age of each sample is derived from the Bayesian result of CIGALE. The upper left, the upper right, the lower left, and the lower right panel shows the stellar mass, the stellar age, the SFR, the sSFR, versus  $\text{EW}_{[\text{OIII}]}$ , respectively. In each panel, the magenta solid line tracks the best-fit linear correlation of the full sample. Those HAEs with both [OIII] and [OII] detection are marked as blue solid circles, while HAEs with only [OIII] detection are presented by blue open circles. For comparison, the SDSS star-forming/star-burst galaxies are represented by contours, with the densest region depicted in black. Besides, extreme O3Es at  $z \sim 2.2$  from (Tang et al. 2019) are marked as orange square in these panels. The red dashed-dotted line present the relationship for  $z \sim 2.3$  massive galaxies from MOSDEF (Reddy et al. 2018).  $\text{EW}_{[\text{OIII}]}$  of our sample exhibits a strong correlation to stellar mass, age and sSFR.



**Figure 4.2.** Relationship between the  $EW_{[\text{OII}]}$  and stellar properties (stellar mass, age, SFR, sSFR) of the 824 HAEs with [OII] at  $z \sim 2.3$  in our study. Outlines as in Figure 4.1. Here, those HAEs with both [OIII] and [OII] detection are marked as blue solid circles, while HAEs with only [OII] detection are presented by blue open circles.



**Figure 4.3.** Relationship between the  $\text{EW}_{\text{H}\alpha}$  and stellar properties (stellar mass, age, SFR, sSFR) of the 1318 HAEs at  $z \sim 2.3$  in our study. Outlines as in Figure 4.1.

### 4.1.3 H $\alpha$ EWs

The H $\alpha$  equivalent widths as a function of aforementioned stellar properties are also shown in Figure 4.3 and Table 4.1. The equivalent widths of recombination lines also work as proxies for stellar mass and stellar age in literature (e.g., Reddy et al. 2018; Faisst et al. 2019; Atek et al. 2022). Typically, galaxies with strong [OIII] are also characterized by strong recombination lines. Thus, it is not surprising that the correlations between the H $\alpha$  equivalent width and these physical parameters are found to be similar to those observed for the [OIII] equivalent width.

The H $\alpha$  recombination line directly reflects the ongoing star formation rate (SFR), while stellar mass is closely related to the stellar continuum. Therefore, it is not surprising that  $\log(\text{EW}_{\text{H}\alpha})$  exhibits the strongest dependence on sSFR, as sSFR is a measurement of the current star formation activity relative to the stellar mass. Interestingly, when examining the correlation between the equivalent width of these emission lines and sSFR for the SDSS sample, i.e., the bottom-right panel of Figure 4.1, 4.2, 4.3, we find that the slope of the best-fit linear correlation between these two attributes remains unchanged only for H $\alpha$ . On the other hand, for our HAEs sample, the slope of the correlation between sSFR and equivalent widths of [OIII] and [OII] are different to that of the SDSS sample. This finding probably contradicts the statement by Reddy et al. (2018) that the sSFR-EW relationship is largely redshift-invariant for all emission lines. While our result indicates that this relation is likely to be redshift-invariant only for H $\alpha$  emission.

Overall, our sample indicates that both the H $\alpha$  and [OIII] equivalent widths are sensitive to stellar mass, stellar age, and specific SFRs. Galaxies with larger H $\alpha$  or [OIII] equivalent widths tend to have lower stellar masses, younger stellar populations, and higher sSFRs. In contrast, the [OII] equivalent widths show a much weaker dependence on these parameters. Note that, the HAEs in our study represents a less biased sample compared to previous works, which may have biased towards high-mass emitters. Although different selection biases exist, the overall trend does not change significantly.

## 4.2 EW vs. ISM properties

The results of Section 4.1 demonstrate the response of the equivalent widths of [OIII], [OII], H $\alpha$  to stellar properties. Here, we further discuss their dependence on ISM properties. The ISM properties include gas-phase metallicity, ionization parameters, electron density, which can be indicated by various line indices. Commonly used metallicity-sensitive line indices include N2 ( $[\text{NII}]/\text{H}\alpha$ ), O3N2 ( $([\text{OIII}]/\text{H}\beta)/([\text{NII}]/\text{H}\alpha)$ ), R23 Kewley & Dopita (2002); Kobulnicky & Kewley (2004); Pettini & Pagel (2004). On the other hand, O32 and O3 ( $[\text{OIII}]/\text{H}\beta$ ) are often used as ionization-sensitive line indices. The line ratios of [OII] and [SII] doublets can serve as electron density diagnostics. In our work, we have access to the O32 and R23 line indices. It is important to note that the R23 index does not vary monotonically with gas-phase metallicity, but instead follows an evolutionary path combined with the O32 index, as indicated by photoionization models (e.g.,

Kewley & Dopita 2002; Ferland et al. 2013).

In the upper panel of Figure 4.4, we show the dependence of  $\log(\text{EW}_{[\text{OIII}]})$  on the  $\log(\text{O32})$  index for individual galaxies. For those HAEs without  $[\text{OII}]$  detections, we assign a lower limit to their O32 values using a  $2\sigma$  upper-limit flux for  $[\text{OII}]$ . These lower limits are indicated by rightward arrows, and they are not included in the calculation of the best-fit linear correlation presented in Table 4.1.

The O32 index serves as a direct indicator of the ionization state of the ISM. We observe a clear trend where O32 increases with increasing  $\text{EW}_{[\text{OIII}]}$ , suggesting a harder ionizing radiation field in these high  $\text{EW}_{[\text{OIII}]}$  galaxies. Additionally, O32 exhibits a secondary dependence on gas-phase metallicity, showing an anti-correlation with metallicity (e.g., Kewley & Dopita 2002; Bian et al. 2018). This could also imply that these high  $\text{EW}_{[\text{OIII}]}$  galaxies are likely to have lower metallicities.

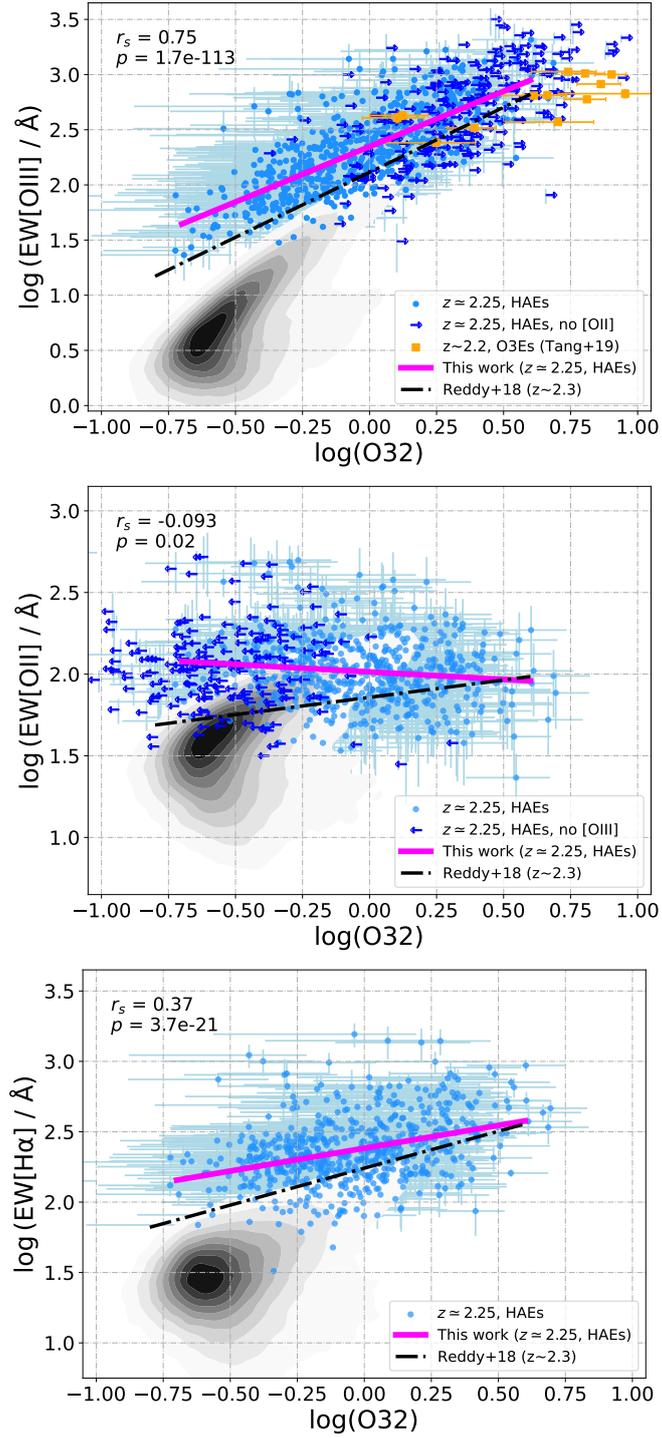
For HAEs without  $[\text{OII}]$  detections, we find a median lower limit of  $\text{O32} > 2.2$ , which is significantly higher than the median O32 value for the rest of the sample, which is close to unity. Among these non-detections, the subset with the highest lower limits for O32 ( $\text{O32} > 5$ ) all have  $\text{EW}_{[\text{OIII}]}$  values exceeding  $500\text{\AA}$ . These galaxies with the highest O32 values, indicative of high ionization parameters in the ISM, are suggested to be powered by extremely young and massive stellar populations.

The middle panel of Figure 4.4 shows how  $\log(\text{EW}_{[\text{OII}]})$  varies with the  $\log(\text{O32})$  index for our sample. Similarly to the previous panel, we assign a  $2\sigma$  upper-limit flux for  $[\text{OIII}]$  to the objects without  $[\text{OIII}]$  detections, indicating their upper limits for the O32 values (indicated by leftward arrows). These objects are not included in the fitting of the linear correlation presented in Table 4.1.

In contrast to the strong correlation observed between  $\text{EW}_{[\text{OIII}]}$  and O32, we find that  $\log(\text{EW}_{[\text{OII}]})$  is nearly independent of  $\log(\text{O32})$  in our sample. As mentioned in Section 4.1.2, the  $[\text{OII}]$  equivalent width is not sensitive to various galaxy properties due to the lack of  $\text{O}^+$  ions in extreme ISM environments. This explanation is also applicable to the result observed here, where the O32 index directly indicates the ionization state of the ISM. On the other hand, the result from Reddy et al. (2018) indicated a correlation between  $[\text{OII}]$  and O32, although the dependence was less significant than that of  $[\text{OIII}]$ . Based on the distribution of data points in this panel, we found that the discrepancy between the two results is mainly driven by objects with lower O32 ratios in our sample, as they exhibit higher  $[\text{OII}]$  equivalent widths compared to the MOSDEF sample in Reddy et al. (2018). We suggest that this discrepancy may be explained by a selection bias towards strong  $[\text{OII}]$  emitters in our galaxy sample.

The bottom panel of Figure 4.4 is the dependence of  $\log(\text{EW}_{\text{H}\alpha})$  on the  $\log(\text{O32})$  index. We find that the equivalent widths of  $\text{H}\alpha$  also increase with O32, although not as significantly as  $[\text{OIII}]$ . Since we lack objects with  $\text{EW}_{\text{H}\alpha} < 100\text{\AA}$ , the selection bias towards strong emitters still exists in our galaxy sample. This bias might result in a shallower slope compared to that reported in Reddy et al. (2018).

Indeed, the overall trend between the equivalent width and ISM properties observed in our sample is similar to the findings in Section 4.1. We find that the equivalent widths of  $[\text{OIII}]$  are the most sensitive to the ionization parameter in the ISM, followed by  $\text{H}\alpha$  emission. On the other hand, the  $[\text{OII}]$  equivalent widths show



**Figure 4.4.** Relationship between the emission line equivalent widths and ionization-sensitive line index,  $\log(\text{O32})$ , for the HAEs at  $z \sim 2.3$  in our study. Upper:  $\log(\text{EW}_{[\text{OIII}]})$  vs.  $\log(\text{O32})$ . Those HAEs with only [OIII] detection are presented by rightward arrows. Upper Middle:  $\log(\text{EW}_{[\text{OII}]})$  vs.  $\log(\text{O32})$ . HAEs with only [OII] detection are presented by leftward arrows. Bottom:  $\log(\text{EW}_{\text{H}\alpha})$  vs.  $\log(\text{O32})$ .

almost no dependence on the ionization parameter. This indicates that the [OIII] emission line is particularly responsive to the ionization state of the ISM, while [OII] is less affected by these ISM properties. These results highlight the different behaviors of emission lines in response to the ionization conditions within galaxies.

In the upper left panels of Figure 4.1, there is an offset in  $EW_{[\text{OIII}]}$  between our sample and SDSS galaxies at a fixed mass, which may be associated with the increase in SFR with redshift at a given mass (e.g., Whitaker et al. 2014), and/or the decrease in metallicity with redshift (e.g., Sanders et al. 2020). The former is indicated above, where our sample have higher average SFR than the SDSS sample because of the increasing molecular gas fraction. To further explore the cause of this evolution, we look into the correlation between  $EW_{[\text{OIII}]}$ , SFR and O32 in more detail. Following the method in Reddy et al. (2018), we calculate how the relationship between the residuals in  $EW_{[\text{OIII}]}$  versus residuals in SFR varies in bins of residual O32, which are computed by the deviation of each galaxy’s  $EW_{[\text{OIII}]}$ ,  $\text{SFR}(\text{H}\alpha)$ , O32 from the best-fit values at the same stellar mass. The best-fit values of  $\text{SFR}(\text{H}\alpha)$  are taken from Equation 3.1, and those of  $EW_{[\text{OIII}]}$  and O32 are calculated from the best-fit linear functions of 626 HAEs that have both [OIII] and [OII] emission lines detected. The best-fit results of these linear relations are:

$$\log(EW_{[\text{OIII}]}) = -0.54 \times \log M_* + 7.41. \quad (4.1)$$

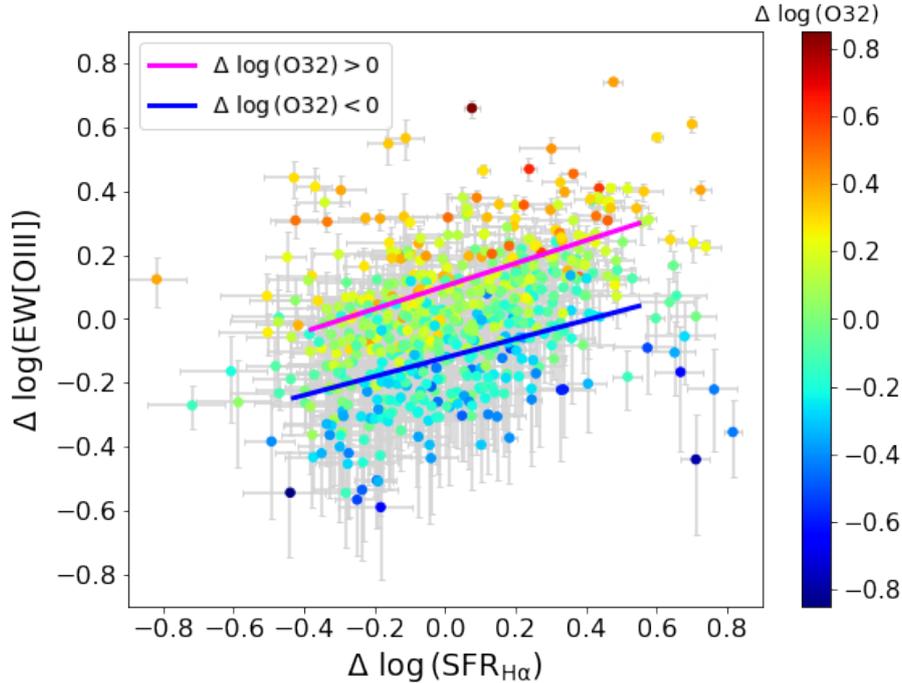
$$\log(\text{O32}) = -0.24 \times \log M_* + 2.25. \quad (4.2)$$

Figure 4.5 shows the result that the equivalent width changes with O32 at a fixed offset from the SFMS. As explained above, O32 are primarily correlated with the ionization parameter and also have a anti-correlation to the gas-phase metallicity. Our result proves that the redshift evolution of decreasing metallicity is also a factor to explain the increase in  $EW_{[\text{OIII}]}$  with redshift at a fixed stellar mass. Physically, the decreasing gas-phase metallicity enable to produce more high-temperature massive stars, which would produce much more ionized photons. These ionized photons are more likely to ionize the surrounding neutral ISM into [OIII], leading to the redshift evolution of  $EW_{[\text{OIII}]}$ , which is also observed by Khostovan et al. (2016).

### 4.3 The ionization parameters of low-mass HAEs

By applying the relationships in Table 4.1, the high  $EW_{[\text{OIII}]}$  observed in the low-mass HAEs indicates the presence of a young stellar population ( $< 100\text{Myr}$ ) and a high ionization state ( $\text{O32} \sim 3$ ) in the system. Also, such an elevated  $EW_{[\text{OIII}]}$  value ( $EW_{[\text{OIII}],\text{med}} \simeq 695\text{\AA}$ ) is typically reported in extreme O3Es studies, like Yang et al. (2017b) for “green pea” galaxies in the local universe with  $EW_{[\text{OIII}],\text{med}} \simeq 733\text{\AA}$ ; Tang et al. (2019) for O3Es at  $1.3 < z < 2.4$  with  $EW_{[\text{OIII}],\text{med}} \simeq 676\text{\AA}$ ; and Onodera et al. (2020) for O3Es at  $3.0 < z < 3.7$  with  $EW_{[\text{OIII}],\text{med}} \simeq 730\text{\AA}$ . These O3Es have revealed extreme ionization conditions not commonly seen in older and more massive galaxies, and the low-mass HAEs may also have similar unique properties.

Previous multiple emission lines analysis on galaxies at  $z \sim 2$  mainly focus on massive and bright galaxies, leading to a bias that limits the detailed study of low-mass galaxies at these epochs. The inclusion of a large number of low-mass HAEs



**Figure 4.5.** Residuals of  $EW_{[\text{OIII}]}$  vs. residuals of  $SFR(H\alpha)$  with the color gradient of dots shows the residuals of O32. The residuals are computed by the deviation of each galaxy’s  $EW_{[\text{OIII}]}$ ,  $SFR(H\alpha)$ , O32 from the mean values at the same stellar mass. This includes the 626 HAEs with both detection of [OIII] and [OII]. The magenta and blue solid lines show the best-fit linear functions with positive and negative residuals of O32, respectively. At a fixed SFR, the equivalent width clearly increases with O32.

in our galaxy sample at  $z \sim 2.3$  is unique compared to previous studies. The high  $EW_{[\text{OIII}]}$  values observed in these low-mass HAEs further raise interest, as they suggest the possibility of extreme galaxy and ISM properties in this population. Understanding the properties and evolutionary paths of these low-mass galaxies at high redshifts can provide valuable insights into galaxy formation and cosmology.

The relationship between the low-mass HAEs and LAEs is indeed an interesting topic of investigation. LAEs have been identified as potential candidates for LyC leakage, a process thought to be important for cosmic reionization (Robertson et al. 2013; Bouwens et al. 2015). LAEs exhibit specific physical properties, such as high ionization states and low dust absorption, that enable LyC photons to escape into the intergalactic medium (IGM). The association between strong Ly $\alpha$  and [OIII] emission lines observed in  $z > 7$  LAEs, as reported by Finkelstein et al. (2013), suggests a connection between the extreme [OIII] emission and LyC leakage. Additionally, Nakajima & Ouchi (2014) found a possible correlation between O32 and  $f_{esc}$  (the fraction of LyC photons escaping) from LAEs at  $z \sim 2$ , indicating the presence of “density-bounded” HII regions in LAEs that facilitate the escape of ionizing radiation into the IGM. Considering that a subset of low-mass HAEs in this study have high O32 values, a characteristic often observed in LAEs, it becomes intriguing to explore the relationship between these two populations. Investigating whether low-mass HAEs exhibit characteristics associated with LyC leakage, such as extreme [OIII] emission and low dust absorption, could shed light on their potential

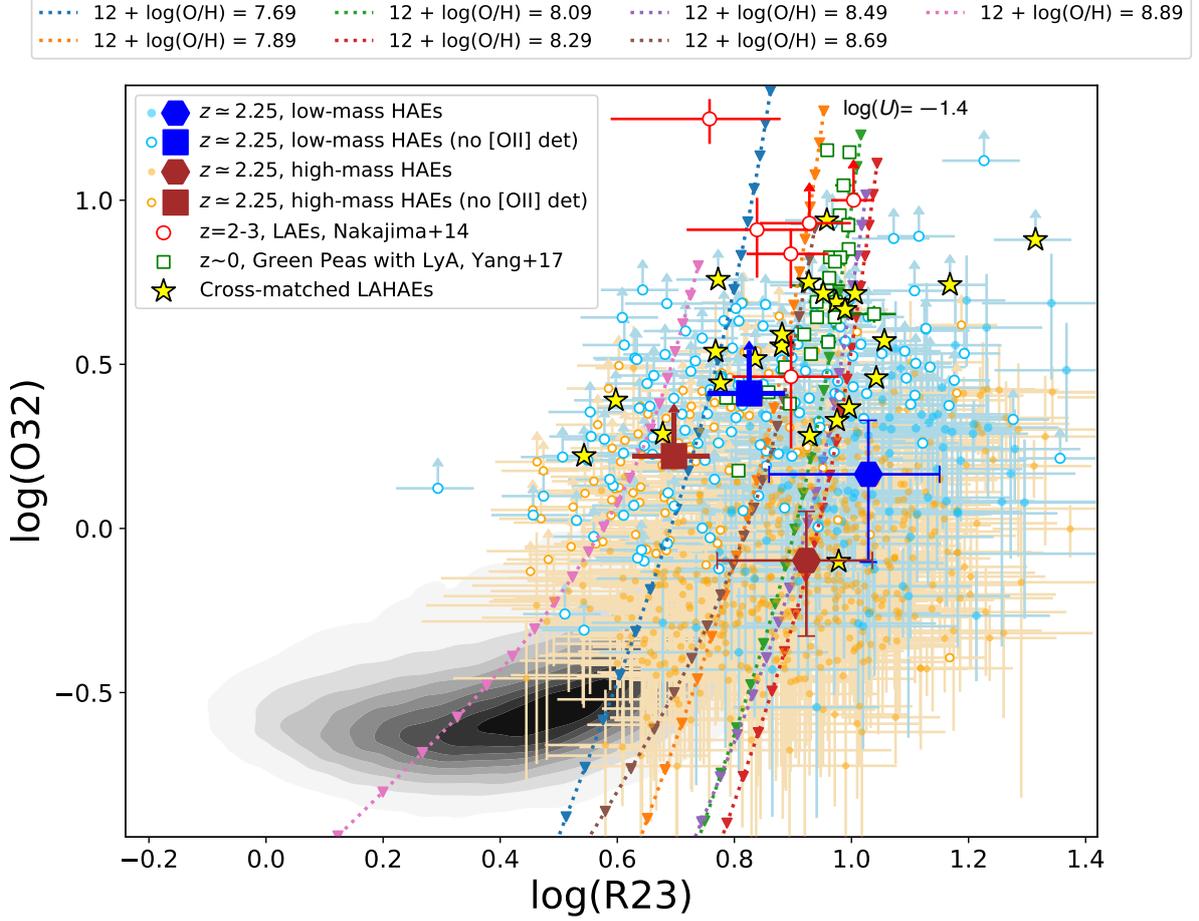
role in the cosmic reionization process. Comparing the properties and behaviors of low-mass HAEs and LAEs can provide valuable insights into the diversity and contributions of different galaxy populations.

We conduct a cross-match between all the galaxies (not only HAEs) at  $2.05 < z < 2.5$  in the ZFOURGE catalog (AGN excluded) and two LAEs catalogs: Nakajima et al. (2012) (ZF-COSMOS, UDS field) and Sobral et al. (2018) (ZF-COSMOS field). Nakajima et al. (2012) constructed a catalog of photometric-selected LAEs at  $z \sim 2.2$  ( $2.14 < z < 2.26$ ) using Subaru narrowband imaging data (NB387). Among their LAE candidates, we found 16 targets in the ZFOURGE-COSMOS field and 19 targets in the ZFOURGE-UDS field, respectively. Sobral et al. (2018) created a large sample of  $\sim 4000$  photometric-selected LAEs at redshifts ranging from  $z \sim 2$  to 6. They used deep narrow- and medium-band imaging data from the Subaru and Isaac Newton Telescopes in the COSMOS field, covering an area of  $\sim 2 \text{ deg}^2$ , and the redshift range of  $2.40 < z < 2.50$  is overlapped with our sample. We identified 10 targets in the IA427 and NB392 filters, and none of these targets overlapped with Nakajima et al. (2012).

In total, we have identified 45 cross-matched galaxies from these two LAEs catalogs. Among them, a subset of 36 cross-matched LAEs shows the detection of [OIII] emission lines. In the subsequent discussion, we will mainly focus on these 36 cross-matched LAEs. Besides, within these 36 cross-matched LAEs, 23 LAEs are also identified as HAEs in our study and 19 are classified as the low-mass ( $< 10^9 M_\odot$ ) ones. We refer to the galaxies that hold both Ly $\alpha$  and H $\alpha$  emission as cross-matched LAHAEs.

In Figure 4.6, we present the O32 versus R23-index diagram, which allows us to investigate the ionization parameter and gas-phase metallicity of galaxies (Kewley & Dopita 2002; Kobulnicky & Kewley 2004). The diagram includes the emission line ratios of HAEs at  $z \sim 2.3$  from our study with star symbols representing the cross-matched LAHAEs. Note that not every sample in our study has simultaneous observations of [OIII] and [OII]. Therefore, we only include the samples with [OIII] detection in this figure. For those with [OIII] detection but no [OII] detection, we also use the  $2\sigma$  upper-limit fluxes of [OII] as lower limits for O32, and they are depicted as open circles with arrows. To differentiate between the low-mass ( $< 10^9 M_\odot$ ) HAEs and high-mass ( $> 10^9 M_\odot$ ) HAEs, we assign them different colors. Additionally, we plot the median values of these groups as larger and darker symbols for clarity. For comparison, we include spectroscopic observations of  $z \simeq 2\text{--}3$  LAEs from Nakajima & Uchi (2014), local “green pea” galaxies with detection of strong Ly $\alpha$  line ( $EW_{\text{Ly}\alpha} > 20 \text{ \AA}$ ) from Yang et al. (2017b), as well as local SDSS star-forming and star-burst galaxies. To provide a reference framework, we overlay the O32 versus R23-index curves from the Cloudy (Ferland et al. 2013) + BPASS (Stanway & Eldridge 2018) photo-ionization models. Each curve corresponds to a different gas-phase metallicity value ranging from  $12 + \log(\text{O}/\text{H}) = 7.69$  to 8.89. The triangles along each curve represent the ionization parameter  $\log(U)$ , which increase from  $-3.6$  (bottom) to  $-1.4$  (top). By comparing the observed data with the model curves, we can gain insights into the ionization states and metallicities of the galaxies in our sample.

The median values of our sample demonstrate that low-mass HAEs tend to pos-



**Figure 4.6.** Relation between O32 and R23-index for H&A&Es at  $z \sim 2.3$ . [OIII] and [OII] are derived from the flux excesses in  $H_s/H_l$  and  $J_2/J_3$  photometry, respectively. Blue (orange) filled circles show the low-mass H&A&Es (high-mass H&A&Es) with both detection of [OIII] and [OII] emission lines, while blue (orange) open circles represent H&A&Es with only [OIII] detection but no [OII] detection. Big deeper points are the median stacks of each classification. The 23 cross-matched LAHAEs from Sobral et al. (2018) and Nakajima et al. (2012) are marked as yellow stars. Red open circles represent the spectroscopic observations of  $z \simeq 2-3$  LAEs from Nakajima & Ouchi (2014), and green squares are the “green pea” with strong Ly $\alpha$  emission from Yang et al. (2017b). Local SDSS star-forming/star-burst galaxies are represented by contours. Cloudy+BPASS model emission-line ratios are shown on this O32 vs.R23 diagram. The triangle data points on the curves increase in size as  $\log(U)$  increases at fixed nebular metallicity.  $\log(U)$  is varied in steps of 0.10 dex from  $\log(U) = -3.6$  to  $-1.4$  and have nebular metallicity range between  $0.1Z_\odot$  to  $1.5Z_\odot$ . The distribution of H&A&Es fit well to the model estimation. On the other hand, the low-mass H&A&Es seem to have higher ionization properties than high-mass H&A&Es in our study, but lower than those of LAEs.

sess higher ionization parameters compared to high-mass HAEs. Notably, LAEs from Nakajima & Ouchi (2014), Yang et al. (2017b) and those cross-matched LAHAEs in our study have nearly the highest ionization properties among all the samples. Separately, LAEs in Nakajima & Ouchi (2014) have a highest median  $O32_{\text{med}}$  of  $\sim 7.5$ ; those in Yang et al. (2017b) have  $O32_{\text{med}} \sim 5$ ; while our LAHAEs have  $O32_{\text{med}} \sim 4$ . This trend suggests that LAEs are more likely to exhibit the highest ionization states and potentially leak ionizing photons into the IGM. Although low-mass HAEs generally show lower ionization properties than LAEs, there are several individual objects within this group that display comparably high  $O32$ -index values, similar to those observed in LAEs.

Also, the local LAEs from Yang et al. (2017b) have a median  $W_{\text{H}\alpha, \text{med}} \sim 500\text{\AA}$  with a median stellar mass of  $\sim 10^9 M_{\odot}$ , while the LAHAEs in our sample have  $EW_{\text{H}\alpha, \text{med}} \sim 600\text{\AA}$  with a median stellar mass of  $\sim 10^{8.5} M_{\odot}$ . These low- $z$  and high- $z$  galaxies hold very close stellar mass,  $EW_{\text{H}\alpha}$  and the  $O32$  ratio, demonstrating the similarities in galaxy and ISM properties. This may reveal that the “green pea” galaxies from Yang et al. (2017b) are the low- $z$  analogs of the LAHAEs in our study.

We further investigate the  $\text{Ly}\alpha$  detection rate, i.e., the  $\text{Ly}\alpha$  escape fraction, of the HAEs in our study. Previous studies that examined the dual emitters of  $\text{Ly}\alpha$  and  $\text{H}\alpha$  have revealed relatively low  $\text{Ly}\alpha$  detection rates within their HAEs sample. For instance, Hayes et al. (2010b) reported a detection rate of 6 dual emitters out of 55 HAEs, Oteo et al. (2015) reported 7 out of 158, and Matthee et al. (2016) reported 17 out of 488. Through the cross-matching analysis in this study, we identify 23 cross-matched LAHAEs out of a total sample of 316 HAEs at  $2.14 < z < 2.26$  and  $2.40 < z < 2.50$ , aligning with these earlier findings. However, if we narrow our focus to the low-mass HAEs, a notably higher rate of dual emitters (19/77) emerges. LAEs are typically characterized as blue, low-mass galaxies with low gas-phase metallicity and high star formation rates, which resemble the properties of the low-mass HAEs in our study. Consequently, it is reasonable to interpret the observed higher detection rate of LAHAEs within the low-mass HAEs.

Conversely, we note that approximately half (19/36) of the cross-matched LAEs are classified as the low-mass HAEs, indicating that these two populations still have differences in their properties. One possible explanation is that LAEs and low-mass HAEs reside in different IGM environments, leading to variations in their galaxy populations. Momose et al. (2021) have discovered a distinctive characteristic of LAEs, that is their tendency to inhabit regions of higher IGM density while avoiding density peaks. Also, Shimakawa et al. (2017) have found a lack of LAEs in the galaxy cluster cores. Both findings support the depletion of  $\text{Ly}\alpha$  emission along the line of sight when penetrating the densest IGM regions. In contrast, other populations such as O3Es and HAEs, do not exhibit this particular preference of IGM environment. Another explanation lies in the selection bias towards strong emitters ( $EW_{\text{H}\alpha} > 100\text{\AA}$ ) in our study. LAEs identified in Nakajima et al. (2012) are selected down to  $EW_{\text{Ly}\alpha} \sim 20\text{\AA}$  through narrow-band observation. Those LAEs with modest  $EW_{\text{H}\alpha}$  may not be captured with the broad-band technique in this study.

Based on the comparison of galaxy properties, we propose a conceptual “Iceberg” model to explain the relationship between LAEs and the low-mass HAEs. In this

model, we draw an analogy between these two populations and the visible and submerged parts of an iceberg. LAEs are often regarded as potential LyC leaking candidates due to their extreme properties such as high ionization states and low dust absorption. However, their population is limited, similar to the visible part of an iceberg above the water surface. These LAEs represent a small fraction of the overall galaxy population. On the other hand, the low-mass HAEs identified in our study can be likened to the submerged part of the iceberg below the water surface. They demonstrate approximately three times higher number density compared to LAEs and generally possess milder ionization states on average. While, it is important to highlight that there exists a significant number of low-mass HAEs that exhibit extreme ISM properties.

# Chapter 5

## The Ionizing properties of the HAEs

In the previous chapter, we highlight that higher redshift analogs of the low-mass HAEs could be significant contributors to the cosmic reionization, which is also proposed by some past studies as the UV radiation budget is dominated by faint galaxies from the UV ( $\sim 1500\text{\AA}$ ) luminosity function at  $z > 6$  (e.g., [Atek et al. 2015](#); [Bouwens et al. 2015](#); [Finkelstein et al. 2019](#)).

In order to fully understand how the universe was reionized, we must constrain the following, (i) the ionizing photon production efficiency,  $\xi_{\text{ion}}$ , defined by the number of LyC photons, i.e, ionizing photons, produced per UV luminosity ([Robertson et al. 2013](#)) and (ii) the escape fraction of ionizing photons,  $f_{\text{esc}}$ , defined as a ratio of transmitted ionizing photons ( $\lambda < 912\text{\AA}$ ) to input ionizing photons. Rest-frame optical emission lines serve as valuable tools for determining these parameters. The primary method for estimating  $\xi_{\text{ion}}$  involves using nebular recombination lines, such as  $\text{H}\alpha$ , to infer the ionizing photon production rate.  $f_{\text{esc}}$ , though derived from LyC flux emitted, exhibits correlations with various observational features of emission lines. However, accurately measuring these parameters at  $z > 6$  presents considerable challenges. Directly measuring LyC flux is impossible due to the attenuation from the IGM along the line of sight. Meanwhile, galaxies at that epoch are too faint to construct a large and precise sample of emission line information.

At lower redshifts, the estimation of  $\xi_{\text{ion}}$  for individual galaxies has already been achieved using nebular recombination lines (e.g., [Bouwens et al. 2016](#); [Nakajima et al. 2016](#); [Shivaei et al. 2018](#); [Nanayakkara et al. 2020](#); [Emami et al. 2020](#); [Atek et al. 2022](#); [Stefanon et al. 2022](#); [Matthee et al. 2022](#)). Previous studies on  $\xi_{\text{ion}}$  have used various samples: (i) Large sample of relatively massive galaxies ( $> 10^{9.5} M_{\odot}$ ) (e.g., [Shivaei et al. 2018](#)), (ii) small sample of low-mass galaxies (e.g., [Hayashi et al. 2016](#); [Emami et al. 2020](#)), (iii) large sample of low-mass galaxies but at relatively low redshift ( $z \sim 1$ ; e.g., [Atek et al. 2022](#)). To gain a comprehensive understanding of  $\xi_{\text{ion}}$  closer to EOR, pushing forward to large sample of low-mass galaxies at high redshift would be an important and necessary task.

## 5.1 Measurement of $\xi_{\text{ion}}$

The ionizing photon production efficiency ( $\xi_{\text{ion}}$ ) is defined as the ratio of the production rate of ionizing photons,  $N(\text{H}^0)$ , in units of  $\text{s}^{-1}$  to the intrinsic UV luminosity ( $L_{\text{UV},\text{cor}}$ ) in units of  $\text{erg s}^{-1} \text{Hz}^{-1}$ :

$$\xi_{\text{ion}} = \frac{N(\text{H}^0)}{L_{\text{UV},\text{cor}}} [\text{s}^{-1}/\text{erg s}^{-1} \text{Hz}^{-1}]. \quad (5.1)$$

In this study,  $L_{\text{UV},\text{cor}}$  is derived at the rest-frame  $1500\text{\AA}$ , and  $N(\text{H}^0)$  is calculated as follows. For an ionization-bounded nebula, we could derive  $N(\text{H}^0)$  from the intrinsic  $\text{H}\alpha$  luminosity,  $L_{\text{H}\alpha,\text{cor}}$ , through the relation of [Leitherer & Heckman \(1995\)](#):

$$N(\text{H}^0) [\text{s}^{-1}] = \frac{1}{1.36} \times 10^{12} L_{\text{H}\alpha,\text{cor}} [\text{erg s}^{-1}]. \quad (5.2)$$

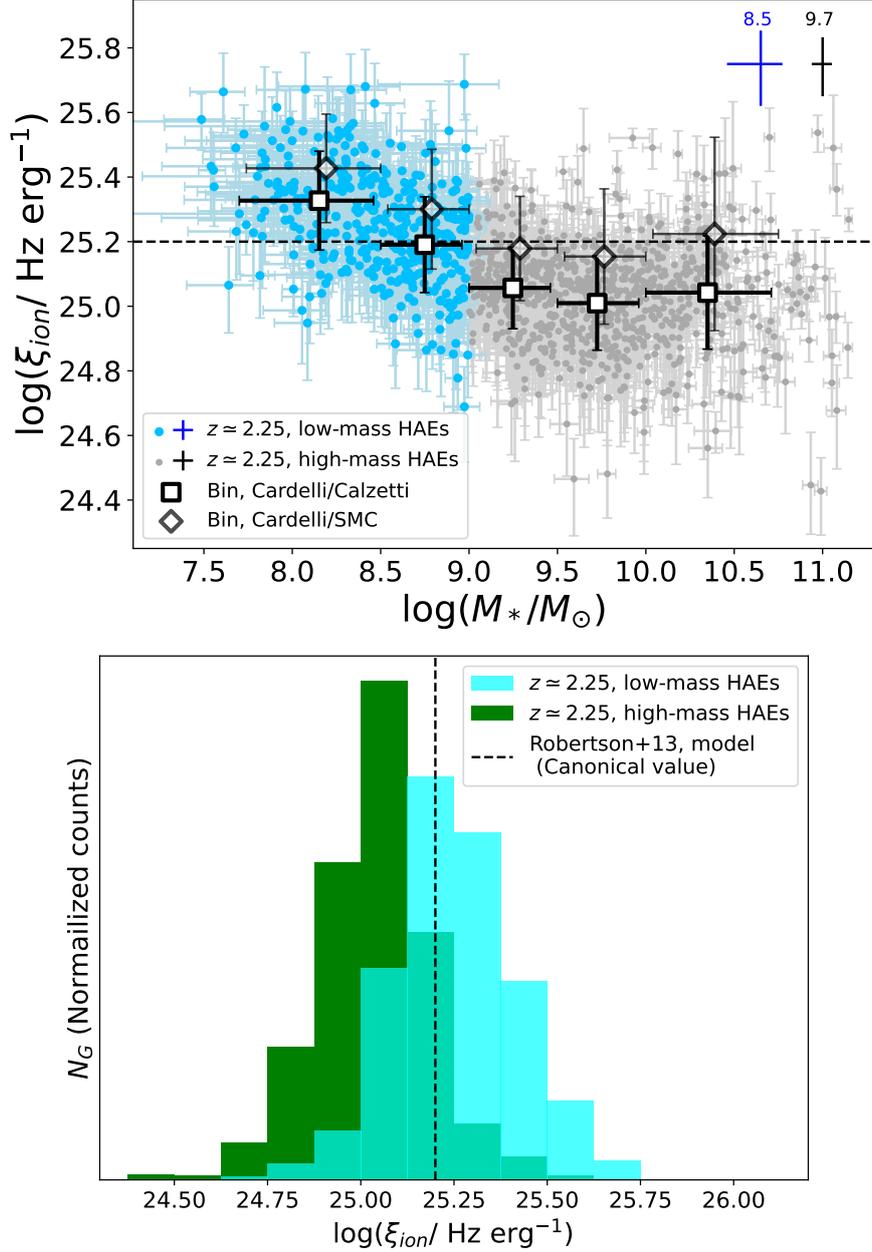
One important assumption in calculating  $\xi_{\text{ion}}$  above is the zero LyC escape fraction, i.e.,  $f_{\text{esc}} = 0$ , since we assume the rate of recombinations balances the production rate of ionizing photons. The combination of Equation 5.1 and 5.2 actually gives us  $\xi_{\text{ion},0}$ , while the intrinsic  $\xi_{\text{ion}}$  should be  $\xi_{\text{ion}} = \xi_{\text{ion},0}/(1 - f_{\text{esc}})$ .

It is estimated that if the universe is totally reionized by galaxies, a large  $f_{\text{esc}} \sim 20\%$  at  $z > 7$  is needed ([Inoue et al. 2006](#); [Ouchi et al. 2009](#)). In contrast, after the universe is fully reionized, based on direct LyC imagings at lower redshift ( $2 < z < 4$ , e.g., [Vanzella et al. 2010](#); [Grazian et al. 2016](#); [Matthee et al. 2017](#)), an upper limit of the escape fraction is  $f_{\text{esc}} < 6\%$ . Therefore, ignoring the mentioned  $f_{\text{esc}}$  correction would not significantly alter the trends we find in this analysis. We will assume  $\xi_{\text{ion}} \simeq \xi_{\text{ion},0}$  when investigating the relationship between  $\xi_{\text{ion}}$  and galaxy properties in the following sections.

Typically, either the [Calzetti et al. \(2000\)](#) curve or the SMC curve is assumed for the reddening of the stellar continuum in high-redshift galaxies. Because the SMC curve has a steeper intrinsic UV slope, these two dust attenuation curves lead to different  $\xi_{\text{ion}}$ , varying from less than 0.1 dex (e.g., [Bouwens et al. 2016](#); [Tang et al. 2019](#)) to more than 0.2 dex (e.g., [Shivaei et al. 2018](#); [Atek et al. 2022](#)). Thus, we also perform SED fitting using the SMC curve for the stellar continuum, while still adopting the Milky Way curve ([Cardelli et al. 1989](#)) for fitting the nebular emission in this chapter.

## 5.2 $\xi_{\text{ion}}$ evolution with galaxy properties

Here, we further investigate the ionizing photon production efficiency,  $\xi_{\text{ion}}$ , of the low-mass HAEs and their relationship to various galaxy properties. Since we have applied two different recipes, Cardelli/Calzetti and Cardelli/SMC, for the nebular/continuum dust correction in our SED fitting analysis, we present the results from both approaches in the following sections, with the individual data points in the figures corresponding to the Cardelli/Calzetti curve.



**Figure 5.1.** The ionizing photon production efficiency ( $\xi_{\text{ion}}$ ) of all HAEs in our study.  $\xi_{\text{ion}}$  is estimated from the  $\text{H}\alpha$  and UV luminosities followed by Equation 5.1 and 5.2.  $\xi_{\text{ion}}$  of the low-mass HAEs is higher than other HAEs in our study by  $\sim 0.2$  dex, indicating a possible mass dependence. Upper: Dependence of  $\xi_{\text{ion}}$  on the stellar mass. The galaxy sample are separated into two populations, the high-mass HAEs and low-mass HAEs as in Figure 3.1. The error bars on the upper-right corner represent the median uncertainty of low-mass HAEs (blue) and high-mass HAEs (black), added with the median stellar mass of each subsample. Open squares and diamonds are median stacks in 5 mass bins, while the error bars on them represent the scatter in each mass bin. The square-shape stacks assume the Calzetti curve with for the UV dust correction, and the diamond-shape stacks assume an SMC curve. Bottom: The  $\xi_{\text{ion}}$  distribution in the high-mass HAEs and low-mass HAEs assuming the Cardelli/Calzetti dust correction. The median values and errors for each population are  $\log(\xi_{\text{ion}}/\text{Hz erg}^{-1}) = 25.24^{+0.10}_{-0.13}$ ,  $25.05^{+0.08}_{-0.10}$ . The dashed lines in each figure indicate the canonical value of  $\log(\xi_{\text{ion}}/\text{Hz erg}^{-1}) = 25.20$  from Robertson et al. (2013).

### 5.2.1 $\xi_{\text{ion}}$ and stellar mass

We display the relationship between  $\xi_{\text{ion}}$  and the stellar mass in our sample in the left panel of Figure 5.1. We also include open circles to represent the median  $\xi_{\text{ion}}$  in different mass bins. To balance the number of samples in each bin, we combine the two most massive mass bins in Figure 3.1. The median  $\xi_{\text{ion}}$  of the low-mass HAEs is  $\log(\xi_{\text{ion}}/\text{Hz erg}^{-1}) = 25.24_{-0.13}^{+0.10}$  ( $25.35_{-0.15}^{+0.12}$ ), assuming the Cardelli/Calzetti (Cardelli/SMC) curve. This result is very close to the O3Es subsample in Tang et al. (2019) with  $EW_{[\text{OIII}]} \simeq 300 - 600 \text{ \AA}$  (25.22). On the other hand, the median  $\xi_{\text{ion}}$  of the high-mass HAEs is  $\log(\xi_{\text{ion}}/\text{Hz erg}^{-1}) = 25.05_{-0.10}^{+0.08}$  ( $25.19_{-0.13}^{+0.10}$ ), which is quite similar to that of the galaxies from the MOSDEF survey (25.06; Shivaiei et al. 2018). When using an SMC extinction curve for the continuum, the  $\xi_{\text{ion}}$  values are higher by  $\sim 0.15$  dex compared to the Calzetti curve.

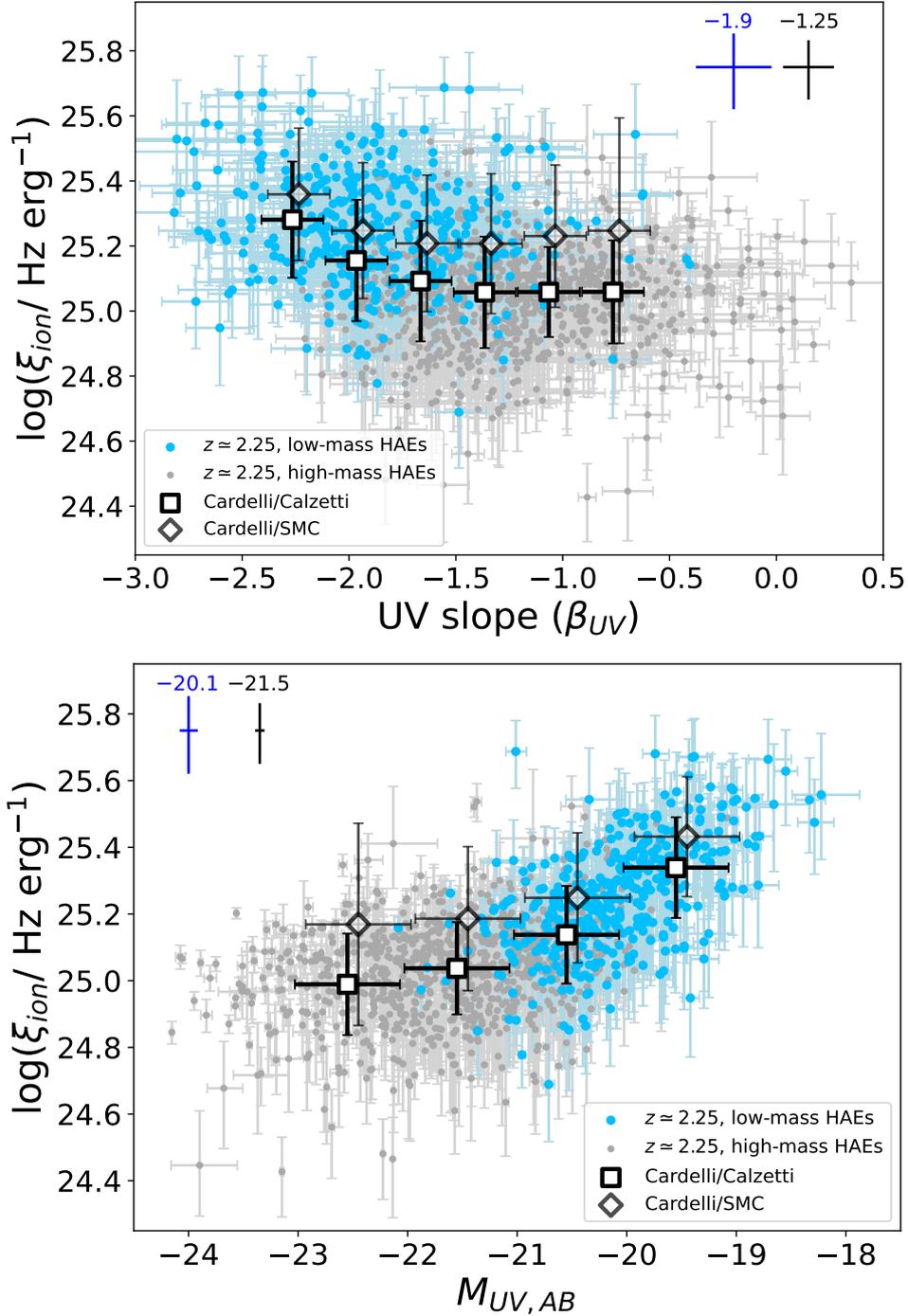
The  $\xi_{\text{ion}}$  distributions of low-mass HAEs and high-mass HAEs are exhibited in the bottom panel of Figure 5.1. The intrinsic scatter, represented by the standard deviation of the distribution, is 0.16 (0.19) dex for low-mass HAEs assuming the Cardelli/Calzetti (Cardelli/SMC) curve. For high-mass HAEs, the intrinsic scatter is 0.16 (0.26) dex.

We find that, on average, the low-mass HAEs have  $\sim 0.2$  dex higher  $\xi_{\text{ion}}$  compared to the high-mass HAEs, indicating a higher efficiency of producing ionizing photons in the low-mass galaxies. In the high-mass regime, we observe no significant evolution of  $\xi_{\text{ion}}$  with stellar mass. This trend is consistent with Shivaiei et al. (2018) at  $z \sim 2$  using MOSDEF spectroscopic data. The  $\xi_{\text{ion}}$  values from MOSDEF remain relatively constant down to  $10^{9.5} M_{\odot}$  but show an increase in the lowest mass bin of  $10^9 M_{\odot}$ , suggesting a possible mass dependence of  $\xi_{\text{ion}}$ . Our method successfully extends the mass range and provides evidence for the existence of a mass dependence of  $\xi_{\text{ion}}$  below  $10^9 M_{\odot}$ . In contrast, at similar redshift, Emami et al. (2020) has found that  $\xi_{\text{ion}}$  is generally independent of the stellar mass in their sample of 28 lensed dwarf galaxies, which span the range of  $8.0 < \log(M_*/M_{\odot}) < 10.0$ . Their sample exhibits a nearly twice as large intrinsic scatter compared to our study. It is worth noting that high lensing magnification can also introduce significant magnification differences across the galaxy sample, leading to uncertainties when measuring  $\xi_{\text{ion}}$ . Differences in sample selection and observational techniques may contribute to the discrepancy found in these studies.

### 5.2.2 $\xi_{\text{ion}}$ and UV Properties

It has been suggested that UV-faint galaxies could be significant contributors to cosmic reionization, based on several observations. One reason is the higher number density, indicated by the steep slope at the faint-end of the UV luminosity function (e.g., Atek et al. 2015; Bouwens et al. 2015). Another factor is the potentially higher LyC escape fraction for faint galaxies (Grazian et al. 2017). Investigating the UV properties of the low-mass HAEs can provide valuable insights into EoR. We report our results in Figure 5.2 and Table 5.1.

The upper panel of Figure 5.2 shows the relationship between  $\xi_{\text{ion}}$  and UV slope ( $\beta_{\text{UV}}$ ) for our sample. We observe a gradual increase in  $\xi_{\text{ion}}$  for galaxies with  $\beta_{\text{UV}} < -2.0$ , with the bluest  $\beta_{\text{UV}}$  objects showing an elevation of more than 0.2 dex. The



**Figure 5.2.**  $\xi_{ion}$  as a function of the UV spectral slope ( $\beta_{UV}$ , Upper), the absolute UV magnitude at  $1500\text{\AA}$  ( $M_{UV}$ , Bottom). Symbols are as in Figure 5.1. Open squares and diamonds are median stacks for these two galaxy properties in Table 5.1, but represent different curve for the UV dust correction, while the error bars on them are the scatter in each bin. The error bars on the corner represent the median uncertainty of low-mass HAEs (blue) and high-mass HAEs (black), added with the number value of median UV slope and UV magnitude. It is clear that galaxies with bluer  $\beta_{UV}$  and fainter UV luminosity are likely to hold larger  $\xi_{ion}$  than others, encouraging that these galaxies would play an important role in reionizing the universe.

**Table 5.1.** Median  $\xi_{\text{ion}}$  for HAEs at  $z \sim 2.3$  separated into different bins of stellar masses, UV luminosities and UV-slopes.

(Sub)sample	$N_{\text{gal}}$	$\log \xi_{\text{ion}} / [\text{Hz erg}^{-1}]$	
		Calzetti <sup>a</sup>	SMC
All HAEs <sup>b</sup>	1318	$25.09^{+0.09}_{-0.11}$	$25.24^{+0.11}_{-0.14}$
Low-mass HAEs <sup>c</sup>	401	$25.24^{+0.10}_{-0.13}$	$25.35^{+0.12}_{-0.15}$
High-mass HAEs	917	$25.05^{+0.08}_{-0.10}$	$25.19^{+0.10}_{-0.13}$
$\log(M_*/M_\odot) > 10.0$	270	$25.04^{+0.07}_{-0.08}$	$25.21^{+0.10}_{-0.12}$
$9.5 < \log(M_*/M_\odot) < 10.0$	326	$25.01^{+0.08}_{-0.10}$	$25.16^{+0.10}_{-0.13}$
$9.0 < \log(M_*/M_\odot) < 9.5$	321	$25.06^{+0.10}_{-0.12}$	$25.18^{+0.11}_{-0.14}$
$8.5 < \log(M_*/M_\odot) < 9.0$	221	$25.19^{+0.10}_{-0.13}$	$25.30^{+0.12}_{-0.16}$
$\log(M_*/M_\odot) < 8.5$	180	$25.33^{+0.11}_{-0.13}$	$25.43^{+0.12}_{-0.15}$
$\beta_{\text{UV}} < -2.1$	121	$25.28^{+0.10}_{-0.13}$	$25.36^{+0.11}_{-0.15}$
$-2.1 < \beta_{\text{UV}} < -1.8$	212	$25.16^{+0.10}_{-0.13}$	$25.25^{+0.11}_{-0.15}$
$-1.8 < \beta_{\text{UV}} < -1.5$	265	$25.09^{+0.09}_{-0.12}$	$25.21^{+0.11}_{-0.14}$
$-1.5 < \beta_{\text{UV}} < -1.2$	306	$25.06^{+0.09}_{-0.11}$	$25.21^{+0.11}_{-0.14}$
$-1.2 < \beta_{\text{UV}} < -0.9$	194	$25.06^{+0.08}_{-0.09}$	$25.23^{+0.10}_{-0.13}$
$\beta_{\text{UV}} > -0.9$	220	$25.06^{+0.07}_{-0.07}$	$25.24^{+0.11}_{-0.14}$
$M_{\text{UV}} < -22.0$	284	$24.99^{+0.06}_{-0.06}$	$25.16^{+0.08}_{-0.09}$
$-22.0 < M_{\text{UV}} < -21.0$	441	$25.04^{+0.08}_{-0.10}$	$25.19^{+0.11}_{-0.14}$
$-21.0 < M_{\text{UV}} < -20.0$	405	$25.14^{+0.10}_{-0.13}$	$25.25^{+0.11}_{-0.15}$
$M_{\text{UV}} > -20.0$	188	$25.34^{+0.11}_{-0.15}$	$25.43^{+0.12}_{-0.16}$

**Notes.** Measurements of the UV properties are only based on the observed photometry mentioned in section 3.1, which have no relationship to the SED fittings.

<sup>a</sup> The uncertainty on each number is the median uncertainty of the subsample, which is different from the scatter shown in Figure 5.1 and 5.2.

<sup>b</sup> The low-mass HAEs are those with stellar mass  $< 10^9 M_\odot$ .

UV slope is sensitive to the age of stellar population of massive stars (O- and B-type stars) in a galaxy. In ionization-bounded photoionization models (e.g., Topping et al. 2022), the bluest expected UV slope is  $\beta_{\text{UV}} \simeq -2.6$  in a dust-free case with stellar age of  $< 30$  Myr. An older stellar population would lead to a redder UV slope up to  $\beta_{\text{UV}} \sim -2.0$  and less production of ionizing photons, resulting in a lower  $\xi_{\text{ion}}$ . Otherwise,  $\xi_{\text{ion}}$  remains nearly unchanged for galaxies with redder  $\beta_{\text{UV}}$  values ( $> -1.5$ ). The effect of dust attenuation might become dominant in this region. Applying the SMC correction for the UV continuum leads to a similar result, albeit with a slightly lower elevation of  $\sim 0.15$  dex.

Our results are consistent with the literature results from Shivaei et al. (2018) at similar redshift and Bouwens et al. (2016) at  $z \simeq 4 - 5$ , both of which inferred an elevated  $\xi_{\text{ion}}$  in galaxies with the bluest  $\beta_{\text{UV}}$ . On the other hand, studies by Emami et al. (2020) and Onodera et al. (2020) have found no correlation between  $\xi_{\text{ion}}$  and  $\beta_{\text{UV}}$ . As mentioned above, Emami et al. (2020) used a sample of 28 lensed dwarf galaxies, with rest-frame equivalent widths up to  $1500 \text{ \AA}$ . Also, the results from Onodera et al. (2020) are based on  $\sim 20$  extreme O3Es with rest-frame equivalent widths up to  $2000 \text{ \AA}$ . We speculate that the discrepancy between our results and previous studies may be attributed to the selection biases, given that their samples

encompass only a small number of extreme emitters, whereas our sample is more inclusive, encompassing galaxies across a broad range of stellar masses and both high and low equivalent widths.

The low-mass HAEs in our study have a median value of  $\beta_{\text{UV}} = -1.90$ , much higher than normal HAEs by  $\sim 0.7$ . Comparing with the dust-free value of  $\beta_0 = -2.23$  from the Meurer et al. (1999) calibration, we still have 72 low-mass HAEs (more than 1/6) with the bluest  $\beta_{\text{UV}} < -2.23$ . Such blue  $\beta_{\text{UV}}$  values suggest a very young stellar population with minimal dust content in the system.

Next, we explore the relationship between  $\xi_{\text{ion}}$  and UV absolute magnitude ( $M_{\text{UV}}$ ) in the bottom panels of Figure 5.2. In our sample, we observe an increase in  $\xi_{\text{ion}}$  for the faintest galaxies compared to the brighter ones, with a difference of more than 0.3 dex. This trend suggests a dependence between  $\xi_{\text{ion}}$  and  $M_{\text{UV}}$  for the HAEs in our study. These results are similar to those observed in Ly $\alpha$  emitters (LAEs) at  $z \sim 3$  from Nakajima et al. (2018, 2020) with the faint end of UV magnitude to  $M_{\text{UV}} \simeq -19.5$  mag. On the other hand, Shivaie et al. (2018) and Emami et al. (2020) did not find a strong correlation between these two parameters in their studies. While, it should be noted that the galaxy sample in Shivaie et al. (2018) has been selected through spectroscopy, which may introduce a bias towards brighter objects  $M_{\text{UV}} \leq -21$  mag. It can be also inferred from our results that brighter galaxies exhibit a weaker dependence between  $\xi_{\text{ion}}$  and  $M_{\text{UV}}$ .

### 5.2.3 $\xi_{\text{ion}}$ and nebular emission lines

The optical nebular emission lines in galaxies provide a wealth of information on the physical parameters, including the stellar population, chemical abundance, and ionization parameter, which may also be related to  $\xi_{\text{ion}}$ .

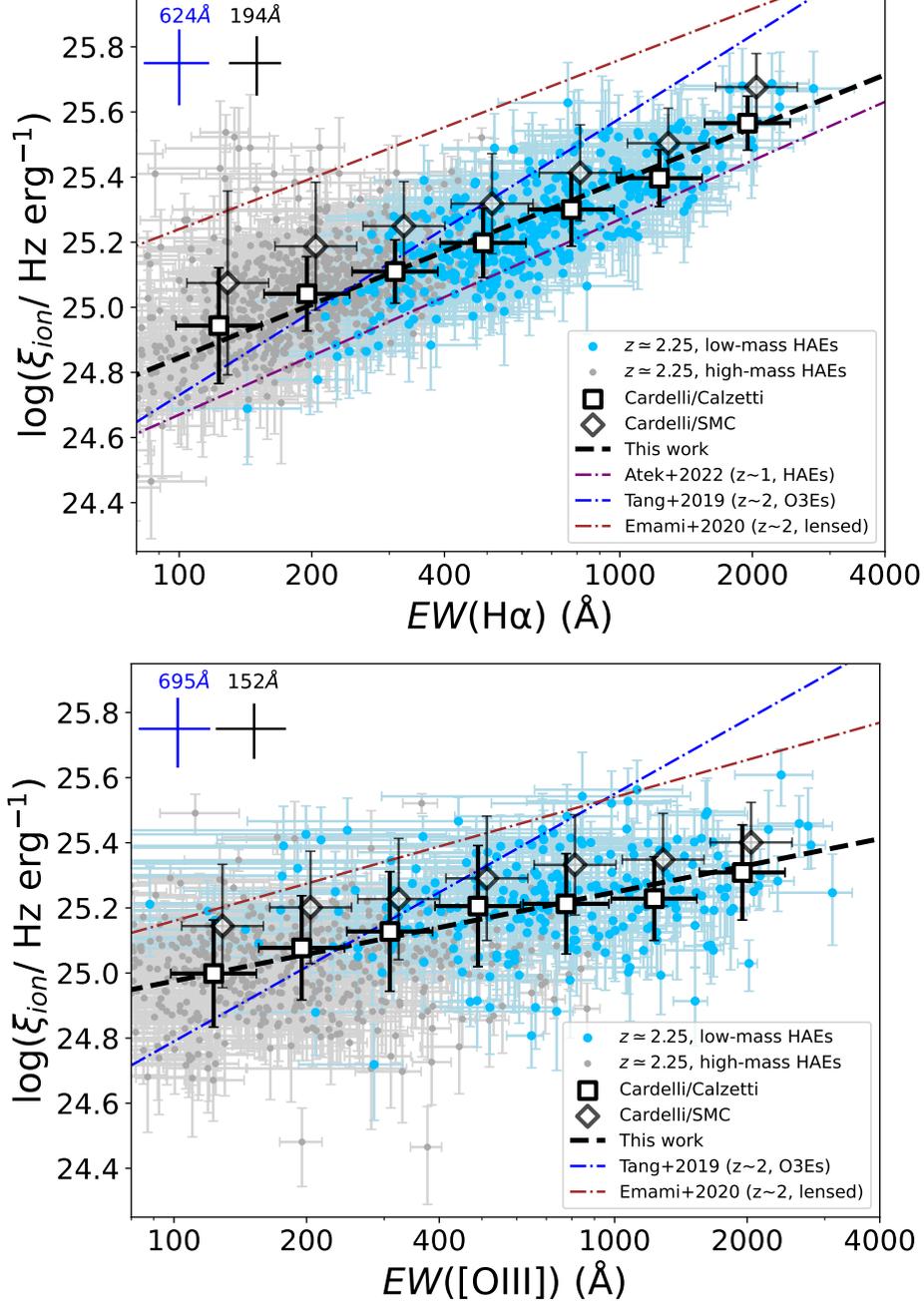
As  $\xi_{\text{ion}}$  serves as a useful indicator of the stellar populations in galaxies, with younger populations contributing more to the H $\alpha$  emission lines, we expect to find a universal relationship between  $\xi_{\text{ion}}$  and the equivalent width of H $\alpha$ . This relationship is depicted in the left panel of Figure 5.3. Assuming the Cardelli/Calzetti curve, we fit a linear relationship between these two attributes:

$$\begin{aligned} \log \xi_{\text{ion}} = & (0.54 \pm 0.03) \times \log(EW_{\text{H}\alpha}) \\ & + (23.76 \pm 0.09). \end{aligned} \quad (5.3)$$

For the Cardelli/SMC curve, we find a relationship with a slope of  $0.51 \pm 0.04$  and an intercept of  $23.96 \pm 0.11$ .

Since large [OIII] equivalent widths are typically produced by massive stellar populations with sub-solar metallicities, which meanwhile produce large amounts of ionizing photons, it is suggested that there also exists a correlation between  $\xi_{\text{ion}}$  and  $EW_{[\text{OIII}]}$ . This correlation has been observed in both local star-forming galaxies (Chevallard et al. 2018) and high-redshift emitters (Tang et al. 2019; Emami et al. 2020; Nakajima et al. 2020), indicating that systems with higher  $EW_{[\text{OIII}]}$  are more efficient in producing ionizing photons. Our results also indicate a similar relationship between these two attributes, assuming the Cardelli/Calzetti curve:

$$\begin{aligned} \log \xi_{\text{ion}} = & (0.27 \pm 0.04) \times \log(EW_{[\text{OIII}]}) \\ & + (24.43 \pm 0.11). \end{aligned} \quad (5.4)$$



**Figure 5.3.**  $\xi_{\text{ion}}$  as a function of the equivalent width of  $\text{H}\alpha$  ( $EW_{\text{H}\alpha}$ , Left), equivalent width of  $[\text{OIII}]$  ( $EW_{[\text{OIII}]}$ , Right).  $[\text{OIII}]$  are derived from the flux excesses in  $H_s/H_l$  photometry. Symbols are as in Figure 5.1. Open squares and diamonds are median stacks for various galaxy properties, but represent different curve for the UV dust correction. The error bars on the corner represent the median uncertainty of low-mass HAEs (blue) and normal HAEs (black), added with the number value of median  $EW_{\text{H}\alpha}$  and  $EW_{[\text{OIII}]}$ . For both panels, the best-fit relation between  $\xi_{\text{ion}}$  and  $EW_{\text{H}\alpha}$  ( $EW_{[\text{OIII}]}$ ) is shown with a black dashed line where we use the Calzetti curve for dust correction. Other relationship derived for HAEs at  $0.7 < z < 1.5$  (Atek et al. 2022),  $[\text{OIII}]$  emitters at  $1.3 < z < 2.4$  (Tang et al. 2019), lensed low-mass galaxies at  $z \sim 2$  (Emami et al. 2020) are shown with purple, blue, brown dashed-dotted line, respectively. Note that Emami et al. (2020) and Atek et al. (2022) used the SMC curve for dust correction, which results in a higher estimation of  $\xi_{\text{ion}}$ .

Similarly, for the Cardelli/SMC curve, we find a relationship with a slope of  $0.24 \pm 0.04$  and an intercept of  $24.62 \pm 0.10$ .

For reference, we also overlay the best-fitting relations from [Tang et al. \(2019\)](#), [Emami et al. \(2020\)](#) and [Atek et al. \(2022\)](#) in [Figure 5.3](#). Although our sample shows a similar trend to literature, we find a clear discrepancy in terms of the slopes and intercepts. One possible reason for this discrepancy could be the treatment of dust extinction. [Tang et al. \(2019\)](#) has derived the relationship for their [OIII] emitters (O3Es) using both the Calzetti law and the SMC law. Similar to our sample, applying the SMC law results in a similar slope but an elevation in the intercept by  $\sim 0.15$  dex. The difference in galaxy samples may also contribute to the discrepancy in slopes. [Emami et al. \(2020\)](#) has suggested that the slope in  $\log \xi_{\text{ion}}$  and  $\log EW_{\text{H}\alpha}$  ( $\log EW_{[\text{OIII}]}$ ) becomes shallower at lower equivalent widths. To explore this further, we derived the best-fitting result only for HAEs with  $EW_{\text{H}\alpha}$  ( $EW_{[\text{OIII}]}$ )  $> 250\text{\AA}$ . Both curves shows a steeper slope of  $0.61 \pm 0.03$  ( $0.32 \pm 0.05$ ), supporting the idea in [Emami et al. \(2020\)](#). While, even when considering larger equivalent widths and fitting the galaxies accordingly, the discrepancy with [Tang et al. \(2019\)](#) still remains. The difference in sample selection, with [Tang et al. \(2019\)](#) focusing on the most intense O3Es, may contribute to the discrepancy in the best-fitting relationship.

In any case, it is worth noting that the relationship between  $\xi_{\text{ion}}$  and the equivalent widths of nebular emission lines, such as  $\text{H}\alpha$  and [OIII], is observed globally at  $z \sim 2$ . These strong nebular emission lines can serve as proxies for measuring  $\xi_{\text{ion}}$ .

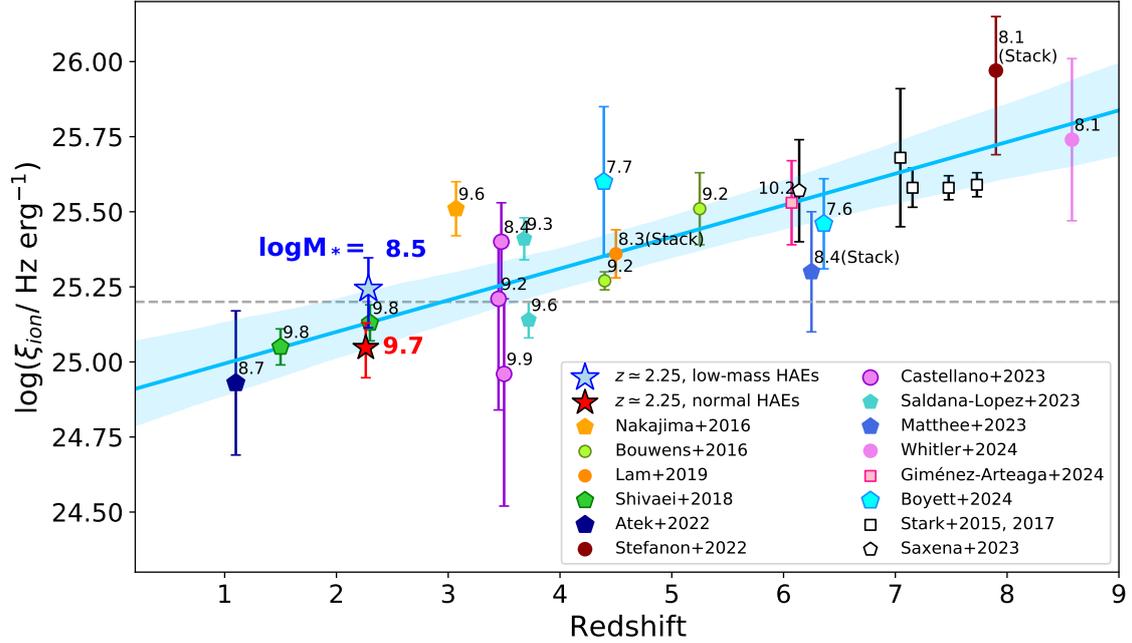
### 5.3 $\xi_{\text{ion}}$ Evolution with Redshift

Constraining  $\xi_{\text{ion}}$  during EoR is an important task for cosmic reionization models. Measuring  $\xi_{\text{ion}}$  for a large number of low-mass galaxies at  $z > 6$  remains very challenging. [Shivaei et al. \(2018\)](#) has suggested a possible evolution of  $\xi_{\text{ion}}$  with redshift, which could provide insights for extrapolating  $\xi_{\text{ion}}$  to higher redshift. We compare our results with previous studies of  $\xi_{\text{ion}}$  at various redshifts in [Figure 5.4](#). Note that the galaxy samples included here are compiled in a heterogeneous manner, encompassing continuum-selected, stacked, and emission line-selected objects, which might introduce additional scatter.

[Atek et al. \(2022\)](#) investigated  $\xi_{\text{ion}}$  of a large sample of 1167 HAEs at  $0.7 < z < 1.5$  by using 3D-HST grism and imaging data. More than half of their 3D-HST sample are low-mass galaxies below  $10^9 M_{\odot}$ , which have a similar stellar mass range to our sample. Since the low-mass 3D-HST galaxies are also the highest-EW galaxies, we consider them as analog sample at lower redshift.

[Shivaei et al. \(2018\)](#) used the spectroscopic and UV imaging data from the MOSDEF survey to measure  $\xi_{\text{ion}}$  at  $z \sim 1.5$  and  $z \sim 2.3$  of 673 galaxies. The lowest stellar mass bin of [Shivaei et al. \(2018\)](#) is  $10^{9.3} M_{\odot}$ , lacking galaxies with  $\log(M_{*}/M_{\odot}) < 9.0$ . The large number of low-mass HAEs found in our study are essential to complete the whole picture at  $z \sim 2$ .

[Nakajima et al. \(2016\)](#) used the spectroscopic measurement of  $\text{H}\beta$  emission line and SED-derived UV continuum to infer  $\xi_{\text{ion}}$  of 13 LAEs as well as 2 Lyman Break Galaxies (LBGs) at  $z \simeq 3.1 - 3.7$ . Their LAE sample consists of less dusty system



**Figure 5.4.** Evolution of estimated ionizing photon production efficiency ( $\xi_{\text{ion}}$ ) with redshift. Our measurement at  $z \sim 2.25$  (blue and red stars) are compared with literature results at a wide range of redshifts, include Stark et al. (2015, 2017); Bouwens et al. (2016); Nakajima et al. (2016); Shivaie et al. (2018); Atek et al. (2022); Stefanon et al. (2022); Castellano et al. (2023); Saldana-Lopez et al. (2023); Matthee et al. (2022); Saxena et al. (2023); Whitley et al. (2023); Giménez-Arteaga et al. (2024); Boyett et al. (2024). The spectroscopic-selected samples are marked as pentagonal symbols, the photometric-selected sample as circular ones and the individual sample as square ones. If multiple values exist in the literature, we take the result based on the Calzetti curve for UV continuum. The numbers next to the markers are the average stellar mass from each sample, and those without stellar mass are marked as open symbols. The HAEs from our sample are marked with different colors. The dashed horizontal line indicates the canonical value of  $\log(\xi_{\text{ion}}) = 25.20$  (Robertson et al. 2013). The blue solid line and the shaded area show the best-fitting linear regression results to all data with 95% confidence interval. We claim a evolution of  $\xi_{\text{ion}}$  with lookback time, with the largest sample at  $z \sim 2$  so far. This evolution can be explained by more bursty star formation on average at higher redshifts.

with color excess  $E(B - V)$  almost consistent with zero. Also, the large  $[\text{OIII}]/[\text{OII}]$  ratio indicates the higher ionization properties of their LAE sample.

Bouwens et al. (2016) measured the  $\text{H}\alpha$  emission line from the observed IRAC fluxes in  $\sim 300$  galaxies at  $z \simeq 3.8 - 5.0$  and 22 galaxies at  $z \simeq 5.1 - 5.7$ . They calculated  $\text{H}\alpha$  flux through the flux excesses in IRAC data, i.e, the  $[3.6]-[4.5]$  color. Besides, their result showed that assuming the SMC extinction law led to  $\sim 0.1$  dex higher  $\xi_{\text{ion}}$  than that with the Calzetti one.

Lam et al. (2019) performed a similar estimation of  $\xi_{\text{ion}}$  from the flux excesses in IRAC as Bouwens et al. (2016) but by using the stacked IRAC colors in a sample of galaxies. Comparing to Bouwens et al. (2016), Lam et al. (2019) extended  $\xi_{\text{ion}}$  measurements to even lower UV luminosities. The distribution of stellar mass and the large number sample in Lam et al. (2019) was comparable to the low-mass HAEs in our study, indicating the similarity of the galaxy samples.

Stark et al. (2015, 2017) discussed 4 individual, bright, lensed galaxies at  $z \sim 7$ .  $\xi_{\text{ion}}$  is inferred from stellar population and photo-ionization models. Significant detection of FUV emission lines of these galaxies, including Ly $\alpha$ , Similar to the sample from Nakajima et al. (2016)

Stefanon et al. (2022) measured the median-stacked galaxy properties of 102 LBGs at  $z \sim 8$ . H $\alpha$  flux was inferred from the flux excess in stacked IRAC 5.8  $\mu\text{m}$  band to IRAC 3.6  $\mu\text{m}$  band, and  $\xi_{\text{ion}}$  was further derived from the UV continuum computed from the stacked SED with the assumption of no dust correction. Their result constituted one of the largest  $\xi_{\text{ion}}$  estimation.

Castellano et al. (2023) calculated the  $\xi_{\text{ion}}$  of more than 1000 VANDELS galaxies at  $2.5 < z < 5$  from a multi-band SED fitting approach with BEAGLE. They found no clear evolution of  $\xi_{\text{ion}}$  with redshift within their probed range, but clear correlations with respect to stellar mass. Also Saldana-Lopez et al. (2023) directly measured  $\xi_{\text{ion}}$  of the LAEs and non-LAEs from their rest-frame UV spectra. Their LAEs have  $\sim 0.3$  dex higher  $\xi_{\text{ion}}$  respect to normal non-LAEs.

Matthee et al. (2022) obtained emission line fluxes and physical properties for a sample of 117 O3Es at  $z \simeq 5.3 - 6.9$ , using the deep *JWST/NIRCam* wide field slitless spectroscopic observations. Measurements of physical properties in their study were based on the median stack spectra of O3Es and following SED fitting.  $\xi_{\text{ion}}$  was obtained from H $\beta$  emission line and SED derived UV luminosity. Also, a subset of 58 O3Es had spectral coverage of H $\gamma$  with  $E(B - V)_{\text{neb}} = 0.14$  estimated from the observed H $\gamma$ /H $\beta$  ratio.

Whitler et al. (2023) selected a candidates of 27 galaxies from the large-scale galaxy overdensities surrounding UV luminous LAEs in the CEERS *JWST/NIRCam* imaging (Bagley et al. 2023). They modelled the SEDs of these galaxies and obtained the SED-inferred  $\xi_{\text{ion}}$  by using the BEAGLE code.

Giménez-Arteaga et al. (2024) performed resolved SED modelling on a highly-lensed galaxy at  $z = 6.072$  from *JWST/NIRCam* imaging. The SED results retrieve consistent measurement of emission lines to the IFU spectra. Based on the resolved SED fitting results, they provided 2D distribution of  $\xi_{\text{ion}}$  of their target.

Saxena et al. (2023) used *JWST/NIRCam* spectroscopic data from the JADES survey (Eisenstein et al. 2023a) and directly measured  $\xi_{\text{ion}}$  of 17 faint LAEs. Similarly, Boyett et al. (2024) used the same dataset and measured  $\xi_{\text{ion}}$  of 28 [OIII] or H $\alpha$  EELGs. The stellar mass of each EELG is fitted by BEAGLE.

We perform a fit to all the listed data points, and the results indicate a clear evolution of  $\xi_{\text{ion}}$  with redshift, which is consistent with previous studies in the literature (e.g., Shivaiei et al. 2018; Atek et al. 2022). The best-fitting relation between  $\xi_{\text{ion}}$  and redshift yields the following result,

$$\log \xi_{\text{ion}} = (0.10 \pm 0.02) \times z + (24.92 \pm 0.10). \quad (5.5)$$

In Figure 5.1, we present the relationship between  $\xi_{\text{ion}}$  and the stellar mass, revealing a mass-dependent evolution of  $\xi_{\text{ion}}$ . Thus, we also fit the aforementioned data points based on their average stellar masses, which are indicated by the numbers next to the markers. After excluding stacked and individual samples, we perform separate fits for galaxies with average stellar masses below  $10^9 M_{\odot}$  and above  $10^9 M_{\odot}$ , respectively. The results are as follows:

$$\begin{aligned}
\log \xi_{\text{ion}} &= (0.09 \pm 0.03) \times z + (25.01 \pm 0.15) & \log(M_*/M_\odot) < 9; \\
\log \xi_{\text{ion}} &= (0.10 \pm 0.03) \times z + (24.85 \pm 0.16) & \log(M_*/M_\odot) \geq 9.
\end{aligned}
\tag{5.6}$$

We find that low-mass galaxies exhibit a slightly higher normalization compared to high-mass galaxies, while both populations follow a similar slope in the relationship between  $\xi_{\text{ion}}$  and redshift, consistent with the results in Section 5.2.1.

It is important to note that the intrinsic differences among the sample selection in different studies should be taken into consideration. The current limitations in data availability make it challenging to identify a large number of individual galaxies at higher redshifts. Stacking methods, while useful, can introduce systematic differences among the samples.

Constructing large samples of individual low-mass galaxies is a challenging task, and currently, it has been undertaken in Atek et al. (2022) and our study. To obtain a clearer understanding of  $\xi_{\text{ion}}$  at higher redshifts, it is crucial to construct large samples of individual low-mass galaxies.

## 5.4 $\xi_{\text{ion}}$ in SPS models and Implications for Reionization

In section 5.2.1, we found an enhancement of  $\xi_{\text{ion}}$  at low-masses. We compared our findings with several studies that also examined  $\xi_{\text{ion}}$  and  $M_*$  relationships below  $10^9 M_\odot$ , but at different redshifts. Atek et al. (2022) has analyzed 3D-HST galaxies at  $z \sim 1$  and also observed an enhancement of  $\xi_{\text{ion}}$  at lower mass, around 0.5 dex, which is larger than our sample’s  $\sim 0.2$  dex enhancement. Conversely, Lam et al. (2019) has stacked IRAC images of galaxies at  $z \simeq 4 - 5$  and found an almost independent relationship between  $\xi_{\text{ion}}$  and  $M_*$ . By combining our result at  $z \sim 2.3$  in Figure 5.5, we speculate a possible ”downsizing” relationship between  $\xi_{\text{ion}}$  and  $M_*$  over cosmic time, suggesting that the correlation between  $\xi_{\text{ion}}$  and  $M_*$  weakens from lower redshift to higher redshift.

At the low-mass end, the difference in  $\xi_{\text{ion}}$  among various redshifts is smaller than 0.5 dex. This discrepancy could potentially be mitigated or even canceled out if we extrapolate the  $M_* - \xi_{\text{ion}}$  relationship from our study and Atek et al. (2022) to lower masses, around  $\log(M_*/M_\odot) \simeq 7.5$ , as was in Lam et al. (2019). This trend suggests the possibility of an upper limit for  $\xi_{\text{ion}}$  theoretically, and the low-mass galaxies are gradually approaching the upper limit  $\xi_{\text{ion}}$  value.

To explore the theoretical limits of  $\xi_{\text{ion}}$ , we investigate several stellar population synthesis (SPS) models. The models we consider include the “Binary Population and Spectral Synthesis” (BPASS v2.2.1) models (Stanway & Eldridge 2018) and GALAXEV (BC03 v2016) models (Bruzual & Charlot 2003) with different metallicities. The initial mass function is using Chabrier IMF (Chabrier 2003) but with varying upper mass cutoffs. In Table 5.2 and Figure 5.6, we present the results from these models under the assumption of a constant SFH. If we assume a stellar age of 30 Myr in the constant SFH scenario, the estimation of  $\xi_{\text{ion}}$  from the BPASS model with an upper mass cutoff of  $100 M_\odot$  and sub-solar metallicity is consistent with

the observed  $\log \xi_{\text{ion}} \sim 25.5$  at  $z \simeq 4 - 5$ . It should be noted that the actual stellar ages may vary among the samples, and even younger stellar populations with an age of 10 Myr would exhibit  $\sim 0.15$  dex higher  $\xi_{\text{ion}}$  compared to those at 30 Myr. This estimation is in closer agreement with recent observations from *JWST* at  $z > 8$  (Whitler et al. 2023). Also, the realistic SFHs of galaxies are more complex than a simple assumption of a constant SFH. If a galaxy’s star formation rate is increasing, it is expected to have a higher  $\xi_{\text{ion}}$  than that of a galaxy with a constant SFH at the same stellar age. Given that low-mass galaxies are more likely to have an increasing star formation rate, their upper limits of  $\xi_{\text{ion}}$  may be even higher.

At the high-mass end, we find that the  $\xi_{\text{ion}}$  of massive galaxies at  $z \sim 1$  is almost one magnitude lower than that at  $z \simeq 4 - 5$ . We propose that these galaxies may be at different stages of their star formation. Those galaxies at  $z \sim 1$  may have decreasing star formation rates, resulting in the lower observed  $\xi_{\text{ion}}$ . In Figure 5.6, we also exhibit the evolution of  $\xi_{\text{ion}}$  for a simple stellar population (SSP) from the BPASS model. If a galaxy undergoes quenching,  $\xi_{\text{ion}}$  would rapidly decrease to  $\log \xi_{\text{ion}} \sim 24.0$  within  $10^8$  yr. On the other hand, massive galaxies at higher redshift in Figure 5.5 may still be undergoing continuous star formation. The discrepancy in their  $\xi_{\text{ion}}$  is likely due to differences in their stellar populations. Referring to Figure 5.6,  $\xi_{\text{ion}}$  would gradually decrease but eventually converge after  $10^9$  yr. For comparison, we include the observed results for massive galaxies at various redshifts in Figure 5.6. It is suggested that the BPASS models would be more representative for galaxies at  $z \simeq 4 - 5$ , while the BC03 models for our sample at  $z \sim 2.3$ .

Frameworks on cosmic reionization (e.g., Robertson et al. 2013, 2015) highlighted the importance of constraining two key parameters: the escape fraction  $f_{\text{esc}}$  and the ionizing photon production efficiency  $\xi_{\text{ion}}$ , at  $z > 6$ . Bouwens et al. (2015) has established a well-matched evolution of the cosmic ionizing emissivity and the galaxy UV-luminosity density, suggesting that galaxies are the major sources responsible for reionization. They also constrained a lower limit of  $\log(\xi_{\text{ion}} f_{\text{esc}}) = 24.50 \pm 0.1$ , if only galaxies contributed to the reionization.

In our study, we assume that the low-mass HAEs at  $z \sim 2.3$  can serve as the analog population of the galaxies that reionized the universe at  $z > 6$ . The median  $\xi_{\text{ion}}$  of these low-mass HAEs is comparable to the canonical value of  $\log \xi_{\text{ion}} = 25.2$  with an escape fraction  $f_{\text{esc}} = 0.2$  (Robertson et al. 2013; Bouwens et al. 2015). Note that our measurement of  $\xi_{\text{ion}}$  assumes no escape fraction. Therefore, we can infer that the required escape fraction to ionize the universe at  $z > 6$  is likely no larger than 0.2.

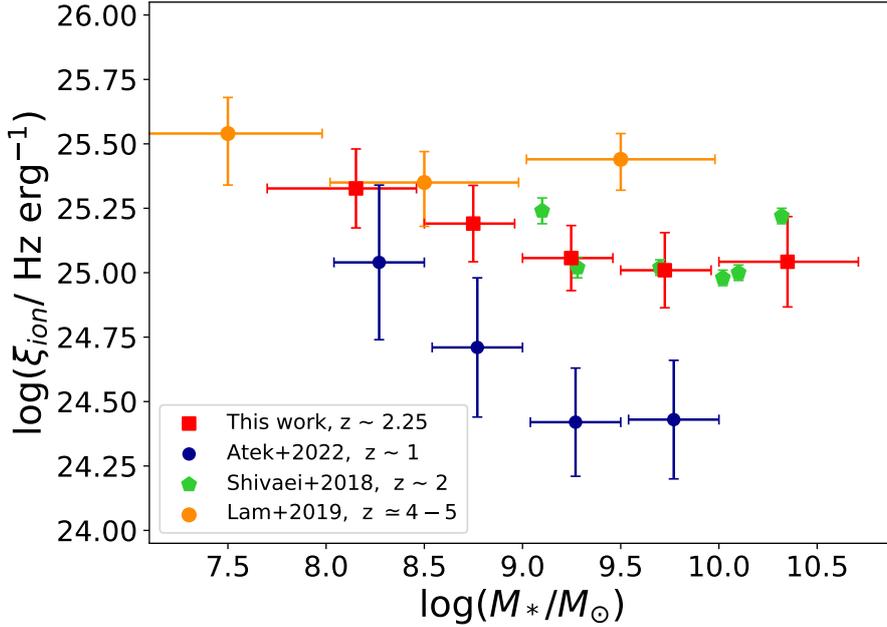
Hydrodynamical simulations show that successive starburst activities are prevalent in low-mass galaxies (e.g., Domínguez et al. 2015; Sparre et al. 2017; Emami et al. 2019). While, variations in the underlying physical conditions of ISM over cosmic time, such as changing metallicities, can lead to different levels of starbursts, resulting in varying equivalent widths of emission lines and different observed  $\xi_{\text{ion}}$  at different redshifts. From the SPS models, we inferred a possible upper limit of  $\xi_{\text{ion}}$  at  $\log \xi_{\text{ion}} \simeq 25.5$ . If we accept this value and assume  $\log(\xi_{\text{ion}} f_{\text{esc}}) = 24.60$ ,  $f_{\text{esc}}$  should be 0.13 at maximum for galaxies during reionization to reionize the universe. Observations of LyC signal and the corresponding  $f_{\text{esc}}$  at  $z \simeq 2 - 4$  have been conducted by several studies in the past decade. Some surveys have reported very few

individual detection of LyC emission, but inferred escape fraction  $f_{\text{esc}} \sim 0.05 - 0.1$  based on stacking imaging (e.g., [Marchi et al. 2017](#); [Naidu et al. 2018](#); [Steidel et al. 2018](#); [Pahl et al. 2021](#)). Still, there are also several individual galaxies that have been observed with very high escape fraction (e.g., [Marques-Chaves et al. 2022](#)).

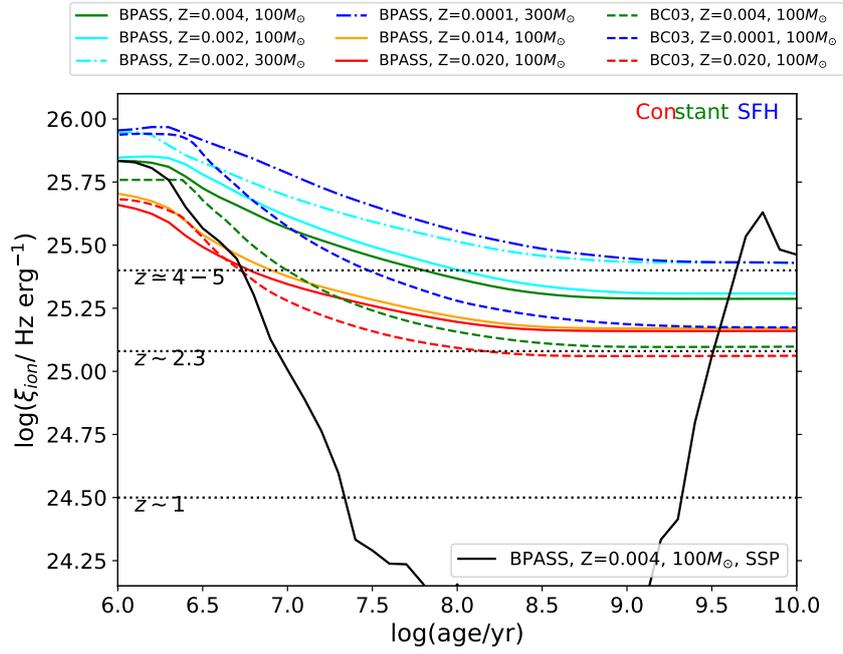
**Table 5.2.** Various SPS models and corresponding  $\xi_{\text{ion}}$

Model (upper mass cutoff)	Metallicity ( $Z_*$ )	SFH & stellar age	$\log(\xi_{\text{ion,model}} / [\text{Hz erg}^{-1}])^{\text{a}}$
BPASS v2.2.1 ( $300 M_{\odot}$ )	0.0001	Constant, $t = 10^8$ yr	25.56
BPASS v2.2.1 ( $300 M_{\odot}$ )	0.002	Constant, $t = 10^8$ yr	25.51
BPASS v2.2.1 ( $100 M_{\odot}$ )	0.002	Constant, $t = 10^8$ yr	25.40
BPASS v2.2.1 ( $100 M_{\odot}$ )	0.004	Constant, $t = 10^8$ yr	25.37
BPASS v2.2.1 ( $100 M_{\odot}$ )	0.014 ( $Z_{\odot}$ )	Constant, $t = 10^8$ yr	25.21
BPASS v2.2.1 ( $100 M_{\odot}$ )	0.004	Constant, $t = 10^{7.5}$ yr	25.46
BC03 v2016 ( $100 M_{\odot}$ )	0.004	Constant, $t = 10^8$ yr	25.16

**Notes.** <sup>a</sup>  $\xi_{\text{ion,model}}$  is derived from the ionizing photon production rate and the luminosity in the FUV band from the spectral synthesis outputs of the models.



**Figure 5.5.** The relation between  $\xi_{\text{ion}}$  and stellar mass at different redshifts. Red squares are stacks in 5 mass bins, same as the upper panel in Figure 5.1. A possible downsizing of the relation between  $\xi_{\text{ion}}$  and  $M_{\odot}$  with the increase of redshift is shown here.



**Figure 5.6.** The evolution of  $\xi_{ion}$  with stellar age in various SPS models with different upper mass cutoff and metallicities. Models with lower metallicities and higher upper mass cutoff are having higher  $\xi_{ion}$  more than a factor of 2. Here, if we suppose a constant SFH,  $\xi_{ion}$  would gradually drop but finally converge. For reference, we also include one typical BPASS SSP model as black solid line and observed results for massive galaxies at  $z \sim 1$ ,  $z \sim 2.3$  and  $z \simeq 4 - 5$  as black dotted line.



# Chapter 6

## Resolved studies of the HAEs from JWST observations

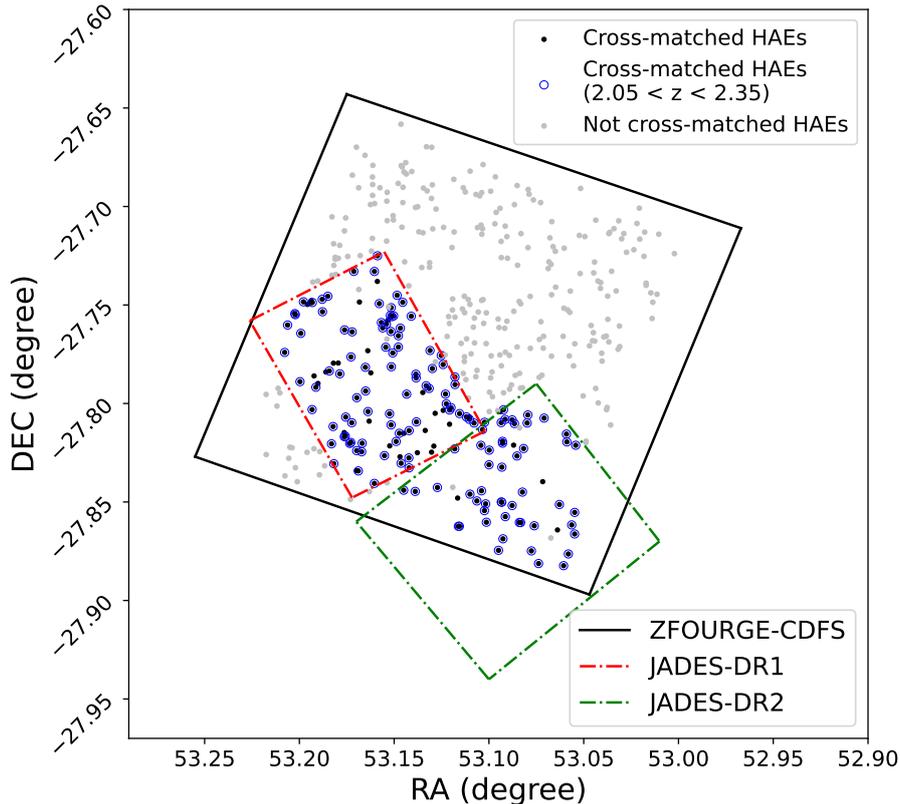
Resolved studies of galaxy properties are important to understand the spatially dependent processes that govern galaxy formation and evolution. Using the unprecedentedly deep and high-resolution data from JWST, we can now resolve the rest-frame optical structures of emission line galaxies in remarkable detail.

### 6.1 JWST observations in ZFOURGE field

In this study, the full sample of HAEs are from the ZFOURGE survey, included in three well-known extragalactic fields: the GOODS-S, COSMOS, and UDS fields, which overlap with the CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011). The GOODS-S field was observed by the Cycle 1 the JWST Advanced Deep Extragalactic Survey (JADES; Eisenstein et al. 2023a; Rieke et al. 2023; Eisenstein et al. 2023b), which includes infrared imaging with nine broad-bandwidth filters from JWST/NIRCam. With exposure times exceeding 10 hours for each filter, JADES provides the deepest ever near-infrared view of this field, reaching a  $5\sigma$  depth of approximately 30 AB magnitude in each filter. JADES First and Second Data Release (DR1 & DR2) provide mosaics images over a total area of 67.7 arcmin<sup>2</sup>, including the F090W, F115W, F150W, F200W, F277W, F335M, F356W, F410M, and F444W filters. The footprint of JADES and ZFOURGE-CDFS fields is presented in Figure 6.1. Additionally, JADES released a 23-band space-based photometric catalog of 94,000 distinct objects, all PSF-matched to the F444W images. In this study, we focus on the spatially-resolved structure of the HAE sample. Therefore, we utilize the F090W, F115W, F150W, F200W, F277W imaging, which provides better angular resolution. The empirical PSF FWHM of the F277W filter is 0."092.

#### 6.1.1 Astrometric correction

The JADES images were astrometrically corrected by registering bright images to the Guide Star Catalogue coordinates from Gaia DR2 (Gaia Collaboration et al. 2018), known as GSC 2.4. The astrometry of ZFOURGE GOODS-South images agrees almost perfectly with that of the CANDELS HST images (Grogin et al. 2011;



**Figure 6.1.** Sky locations of the ZFOURGE survey and the JADES survey in the GOOD-South field. The ZFOURGE footprint covers an area of around  $13' \times 13'$ , while JADES has a total coverage of  $67.7 \text{ arcmin}^2$ . We make a cross-match between the HAE catalog, which contains HAEs at  $2.05 < z < 2.5$ , and the JADES catalog (Rieke et al. 2023). The 170 cross-matched HAEs are initially marked as black dots, with the remaining HAEs shown as grey dots. Within the JADES footprint, seven HAEs are lost in detector gaps and artefacts around bright stars. Finally, the 135 HAEs are selected at  $2.05 < z < 2.35$ , where the [OIII] emission line falls in the F150W filter, are highlighted with blue circles.

Koekemoer et al. 2011), which were registered to the Guide Star Catalog II (GSC 2.3; Lasker et al. 2008). The different versions of astrometric reference stars lead to astrometric errors reported to be  $0.''2 - 0.''7$ , according to the Mikulski Archive for Space Telescopes (MAST).

To correct for the differences in astrometric alignment and to update the CANDELS and ZFOURGE catalogs to the newest astrometry, we cross-match the objects in the CANDELS catalog (Guo et al. 2013) and the JADES catalog with a maximum angular separation of  $1''$ . This results in 17,394 cross-matched objects. We measure a median deviation in right ascension ( $\text{RA}_{\text{JADES}} - \text{RA}_{\text{CANDELS}}$ ) of  $\Delta \text{RA}_{\text{med}} = 0.''107$ ; and in declination ( $\text{DEC}_{\text{JADES}} - \text{DEC}_{\text{CANDELS}}$ ) of  $\Delta \text{DEC}_{\text{med}} = -0.''244$ . A similar result is obtained from cross-matching the ZFOURGE catalog with the JADES catalog, with  $\Delta \text{RA}_{\text{med}} = 0.''122$  and  $\Delta \text{DEC}_{\text{med}} = -0.''247$  from 12,375 matched objects.

Based on these results, we correct all the right ascensions by  $0.''12$  (4 pixels in JADES image) and all declinations by  $-0.''24$  (8 pixels in JADES image) for the original coordinates of the HAE sample in the GOODS-South field.

## 6.1.2 The subsample of HAEs in JADES

We cross-match between the HAE and JADES catalogs (Rieke et al. 2023), requiring an angular separation of less than  $0.''5$  and the availability of photometric data in all the F090W, F115W, F150W, F200W, and F277W filters. We find that 170 HAEs meet these criteria, with a median separation of  $0.''06$ . These HAEs are marked as black dots in Figure 6.1. Given that the JADES images provide the deepest near-infrared view of this field, each HAE is expected to have a corresponding detection in the JADES catalog. However, within the JADES DR1 & DR2 footprint (see also in Figure 6.1), we find that three of these HAEs (ZF-14229, 19965, 23087) fall between the detector gaps of the F200W image, and four others (ZF-7627, 10297, 18528, 21984) overlap with the hexagon-shaped artefacts around bright stars in the JADES images, which were masked when processing photometry.

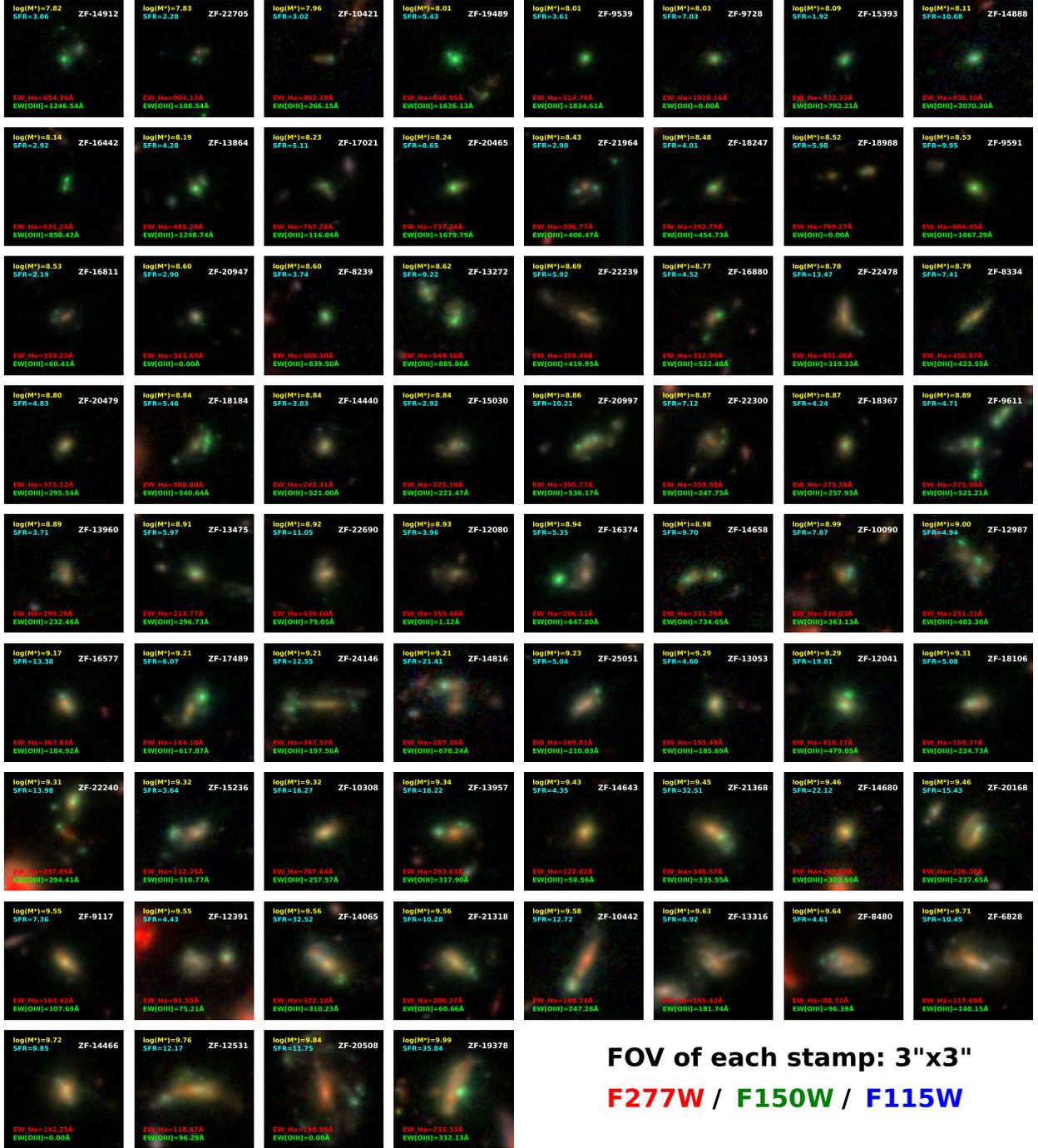
For our HAE sample, the [OIII] emission line at  $2.05 < z < 2.35$  is redshifted into the F150W filter of JWST/NIRCam, while the  $H\alpha$  emission line at  $2.05 < z < 2.40$  is redshifted into the F200W filter. To spatially resolve these two emission lines simultaneously, we therefore require that HAEs fall within the redshift range of  $2.05 < z < 2.35$ . This results in a final sample of 135 HAEs, which are the primary targets of this work. Their sky distributions are added with blue circles around the corresponding black dots in Figure 6.1.

The median redshift of this subsample is  $z_{med} = 2.23$ , along with a distribution peaks at  $z \sim 2.3$ , which may be due to clustered regions in the field. The median stellar continuum dust reddening is  $E(B - V)_{cont} = 0.15$  mag, with only 15 of 135 galaxies exhibiting high dust attenuation of  $E(B - V)_{cont} > 0.3$  mag.

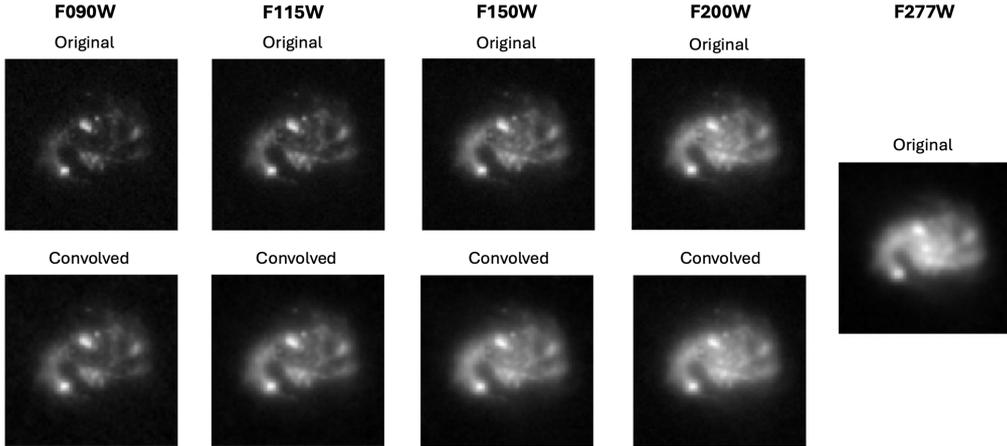
Figure 6.2 displays cutout color images from a selection of the HAEs, created by combining the F115W, F150W, and F277W images. Point-spread functions (PSF) of the shorter wavelength images are matched to that of the F277W image with  $0.''03$  per pixel (See Section 6.2 for details). These three filters directly track the rest-frame blue, green, and far-red colors of the galaxies, roughly corresponding to  $b$ ,  $g$  and  $z$  filters in the rest frame. Notably, the composite color images reveal a prominent number of resolved compact green components, representing strong [OIII]+ $H\beta$  emission line regions with high EW more than several hundreds of angstroms (Cardamone et al. 2009), and indicating extreme ISM properties. In the next section, we demonstrate our unique methodology for extracting and quantitatively analyzing these [OIII]+ $H\beta$  regions. Note that the remaining HAEs that do not contain such a substructure in [OIII]+ $H\beta$  regions are displayed in Appendix C.

## 6.2 Resolved continuum and emission line map

We produce  $3'' \times 3''$  ( $101 \times 101$  NIRCAM pixel) cutout images of each cross-matched HAE from the F090W, F115W, F150W, F200W, and F277W images. The central pixel of each galaxy is determined as the source position of the JADES catalog from the SExtractor (Bertin & Arnouts 1996; Barbary 2016). In Ji et al. (2024), a direct comparison between the JADES empirical PSF and the WebbPSF-predicted model PSF was conducted. Their study found excellent agreement between these two PSFs, with a typical difference of less than 1%. Thus, we match the PSFs of



**Figure 6.2.** JWST/NIRCam cutout images of 68 selected HAEs, each with a size of  $3'' \times 3''$ . The RGB images are built by combining the F115W (B), F150W (G), and F277W (R) PSF-matched images to the F277W filter. For each filter, cutouts with dimensions of  $101 \times 101$  pixels ( $0.03''$  pixel scale) are made. The ZFOURGE ID (Stratman et al. 2016) of each cross-matched HAE is labeled at the upper-right corner of each cutout. The SED-derived integrated galaxy properties: total stellar mass, SFR, and the integrated EW, are also labeled here. These cutout images are ordered by stellar mass, from the lowest to the highest. Notably, these HAEs exhibit several compact [OIII]+H $\beta$  regions (green substructures) embedded within the galaxies.



**Figure 6.3.** Cutout images of a cross-matched HAE imaged in the F090W, F115W, F150W, F200W, and F277W filters. The top row displays the observed images at their original resolution, i.e., direct cutout stamps from the JADES images, while the bottom row shows the images convolved with a kernel that matches the PSF of the F277W filter.

all images to that of the F277W image using the WebbPSF Python package (Perrin et al. 2012, 2014). WebbPSF produces simulated PSFs for the selected JWST instrument and filter. We set the target model PSF as that of the F277W image, and then derive a convolution kernel for the other bands using the scikit-image Python package (Van der Walt et al. 2014) through the Wiener filtering method. In Figure 6.3, we present a sample galaxy (ZF-12763) from the cross-matched HAEs to visualize the PSF-matching process. The top row displays original images from NIRCam, and the bottom images convolved.

Based on the best-fit SEDs of HAEs at  $z \sim 2$ , we discover that strong optical emission lines, such as [OIII] and  $H\alpha$ , can cause noticeable flux excesses in broadband photometric data, specifically in the F150W and F200W images in our sample. On the other hand, other broadband photometric data can roughly trace the stellar continuum because there is no inclusion of strong emission lines. Therefore, we assume that the F115W and F277W photometric data represent the stellar continuum and that a linear relationship of the stellar continuum between them can be applied.

At an arbitrary wavelength ( $\lambda$ ) between F115W and F277W, the stellar continuum can be expressed as follows:

$$f_{\lambda,cont} = f_{F115W} + (f_{F277W} - f_{F115W}) \times \frac{\lambda (\mu m) - 1.15 \mu m}{2.77 \mu m - 1.15 \mu m}, \quad (6.1)$$

where  $f_{F115W}$  and  $f_{F277W}$  are the photometric fluxes in the F115W and F277W filters, respectively. In most cases, the observed F150W and F200W photometric fluxes are higher than the stellar continuum at the corresponding wavelengths from Equation 6.1 due to strong optical emission lines like [OIII] and  $H\alpha$ . This assumption is also applicable in the resolved view. Using the cutout images from F115W and F277W, it is possible to construct a pixel-by-pixel stellar continuum map for the F150W filter using Equation 6.1, as shown in panel (1) of Figure 6.4.

Since the strong optical emission lines can lead to significant flux excesses, it is reasonable to assume that:

$$F_{\text{EL}} S(\lambda_{\text{EL}}) = (f_{\lambda, \text{obs}} - f_{\lambda, \text{cont}}) \int_{\lambda_1}^{\lambda_2} S(\lambda) d\lambda, \quad (6.2)$$

where  $F_{\text{EL}}$  represents the emission line flux fall in the filter,  $f_{\lambda, \text{obs}}$  is the observed photometric data, and  $f_{\lambda, \text{cont}}$  is the stellar continuum derived from Equation 6.1.  $S(\lambda)$  is the throughput function of the filter, which varies with wavelength, and  $S(\lambda_{\text{EL}})$  is the exact throughput where the emission line drop in. The uncertainties in the emission-line flux originate from the observed flux errors, which can be estimated from the noise level of each pixel ( $\sigma_{\text{pixel}}$ ).

In this study, we also apply Equation 6.2 to derive the resolved continuum-subtracted emission line map. For example, after constructing the stellar continuum map at  $1.5 \mu\text{m}$ , we subtracted it from the observed F150W image. This resulting residual image represents the emission line map of [OIII] (and  $\text{H}\beta$ ), which falls in the F150W filter at  $z \sim 2$ . As shown in the panel (2) of Figure 6.4, we present the [OIII]+ $\text{H}\beta$  emission line map of a sample galaxy. Note that, in Section 6.4, we apply statistical corrections to account for contamination from the weaker  $\text{H}\beta$  emission line in each emission line region.

Since the prominent green color in the JWST RGB images results from high  $\text{EW}_{[\text{OIII}]+\text{H}\beta}$ , we apply an EW-limited selection to identify these regions; using the resolved stellar continuum map and emission line map, we can further extract the resolved equivalent width (EW) map for each HAE. The rest-frame EW map is constructed as follows:

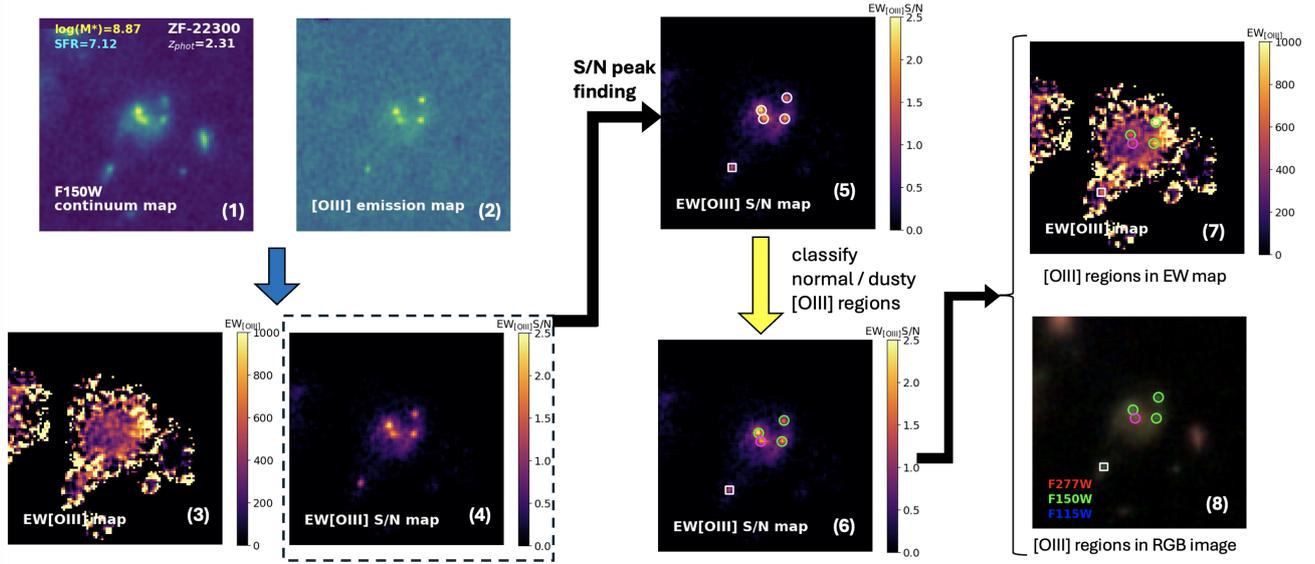
$$\text{EW} = \frac{F_{\text{EL}}}{f_{\lambda, \text{cont}}} \times \frac{\Delta\lambda_{\text{filter}}}{1+z}, \quad (6.3)$$

where  $z$  is the redshift of the HAE, and  $\Delta\lambda_{\text{filter}}$  is the bandwidth of the corresponding filter, which are  $\Delta\lambda_{\text{F150W}} \sim 3180\text{\AA}$  and  $\Delta\lambda_{\text{F200W}} \sim 4630\text{\AA}$ . Panel (3) of Figure 6.4 shows the rest-frame EW map of a sample galaxy, masked by a segmentation map from the JADES data release (Rieke et al. 2023). However, even after masking with the segmentation map, the EW map is still dominated by noise patterns, especially in the outer regions where the stellar continuum is quite low.

To identify and select regions with sufficiently high underlying stellar continuum, we refer to the EW S/N map. In order to construct the EW S/N map, we first obtain the standard deviation of the background pixels. In this study, the pixel-by-pixel error map is referred from the ‘‘ERR’’ uncertainty extensions of the JADES images (Rieke et al. 2023). Next, we calculate the uncertainties in EW on the pixel-by-pixel basis using the following formula:

$$\sigma_{\text{EW}} = \sqrt{\left(\frac{1}{f_{\lambda, \text{cont}}} \times \sigma_{\text{pixel}}\right)^2 + \left(\frac{f_{\text{EL}}}{f_{\lambda, \text{cont}}^2} \times \sigma_{\text{pixel}}\right)^2} \times \frac{\Delta\lambda_{\text{filter}}}{1+z}, \quad (6.4)$$

where  $\sigma_{\text{pixel}}$  is noise of a pixel, originally obtained from the ERR uncertainty extensions of JADES F150W images. Panel (4) of Figure 6.4 displays the  $\text{EW}_{[\text{OIII}]+\text{H}\beta}$



**Figure 6.4.** The flowchart of our methodology to identify “Green Seeds” in a sample galaxy. The stellar continuum map at  $1.5\ \mu m$  (1) is derived by assuming a linear relationship between the stellar continuum in the F115W and F277W filters. The residual image obtained by subtracting this continuum map from the F150W image is the [OIII] emission line map (2) in the unit of  $n_{\text{Jy}}/\text{pixel}$ . From the continuum and emission line images, we generate the EW map (3) and the EW S/N map (4). The EW S/N map reveals several bubble-like structures with high EW. We apply a peak finding algorithm to the EW S/N map to extract these resolved [OIII]+H $\beta$  emission line regions, requiring  $S/N > 2.5$ , which are indicated by white circles in panel (5). White squares represent regions with  $S/N < 2.5$ . We then classify the green (normal) and red (dusty and/or older) [OIII]+H $\beta$  regions through an SED-based color diagram, labeling the former ones as Green Seeds due to their appearance in the RGB images. Green Seeds are highlighted as green circles in panel (6), while Red Seeds are shown as magenta circles. Panel (7) & (8) display these Green Seeds on the EW map and the RGB image, respectively.

S/N map of a sample galaxy. This S/N map efficiently excludes the noise patterns observed in the  $EW_{[\text{OIII}]+\text{H}\beta}$  map, typically shown as unconnected noise pixels. Note that the  $\text{H}\alpha$  emission line map, as well as the  $EW_{\text{H}\alpha}$  and S/N maps, are also derived using the combination of F115W, F200W, and F277W filters, where the  $\text{H}\alpha$  emission lines certainly fall in the F200W filter.

When comparing panel (2) and (4) of Figure 6.4, we observe that the  $[\text{OIII}]+\text{H}\beta$  regions are located at the same positions as those of the  $EW_{[\text{OIII}]+\text{H}\beta}$  S/N peaks. To extract these emission line regions, we apply the `findpeaks` Python Package on the  $EW_{[\text{OIII}]+\text{H}\beta}$  S/N map. After identifying and confirming the pixel coordinates of the  $EW_{[\text{OIII}]+\text{H}\beta}$  S/N peaks, we place a circular aperture with  $0.''15$  (5 pixels) diameter on each peak, corresponding to  $\sim 1.2$  kpc at the median redshift  $z_{\text{med}} = 2.23$ , and to nearly twice the FWHM of the F277W image. We require an average  $EW_{[\text{OIII}]+\text{H}\beta}$  S/N within the circular aperture to be greater than 2.5, as shown by the white circles in panel (5) of Figure 6.4. Those peaks identified with an average S/N  $< 2.5$  are marked with white squares.

This process is repeated iteratively for all 135 cross-matched HAEs to determine the  $EW_{[\text{OIII}]+\text{H}\beta}$  S/N peaks with circular apertures. By applying the same circular aperture on the other scientific frames of the same galaxy, we successfully obtain the total luminosity, emission line flux, and  $EW_{[\text{OIII}]+\text{H}\beta}$  for each emission line region. Among all the S/N peaks, 187 circular apertures have S/N  $> 2.5$  and  $EW_{[\text{OIII}]+\text{H}\beta} > 200\text{\AA}$ , which are considered as the  $[\text{OIII}]+\text{H}\beta$  regions. In this study, we focus on these 187  $[\text{OIII}]+\text{H}\beta$  regions and analyze their properties quantitatively.

### 6.3 “Green Seeds” and “Red Seeds”

In the next step, we visually inspect the  $[\text{OIII}]+\text{H}\beta$  regions on their highly sensitive NIRC*am* RGB (F115W+F150W+F277W) images. Upon inspection, we find a continuous range of the properties of these regions. On one end, there is a population showing greenish color in the RGB images, which are thought to be the most extreme starbursts having the largest  $EW_{[\text{OIII}]+\text{H}\beta}$  values (greater than  $500\text{\AA}$ ) and the bluest underlying stellar continuum from the largest proportion of very young stars. On the other end, there is a population of redder color, which are expected to have dusty spectra with prominent  $EW_{[\text{OIII}]+\text{H}\beta}$  or older populations having Balmer break and no emission lines.

We quantify this sequence in a color-color diagram (F090W – F150W vs. F150W – F277W) and compare these observed colors with simple models of stars and gas in Figure 6.5. First, we generate a series of SED models using BC03 (Bruzual & Charlot 2003) with a Chabrier (2003) IMF. These models assume a instantaneous burst star formation history (SFH), which happened at 3, 5, 10, 20, 50, 100, 200 Myrs before. The sequences of pure stellar light are shown in solid blue line. Emission-line templates generated by CLOUDY (Ferland et al. 1998, 2013) are then added to these SED models to show their effects on broad-band photometry. The stellar metallicity and gas-phase metallicity are fixed at  $Z = 0.008$  (0.6 of Solar metallicity), typical for galaxies at Cosmic Noon (e.g., Erb et al. 2006; Steidel et al. 2014; Sanders et al. 2015, 2021). The free parameters in these SED models are stellar age, dust attenuation and ionization parameters. Older stellar ages and higher dust attenuation lead to redder

spectra, while higher  $\log U$  increases emission line fluxes in the models, resulting in more pronounced flux excesses in certain filters (e.g., F150W and F200W in this study). We then convolve the JWST filter transmission curves to these SED models to obtain the corresponding color indices.

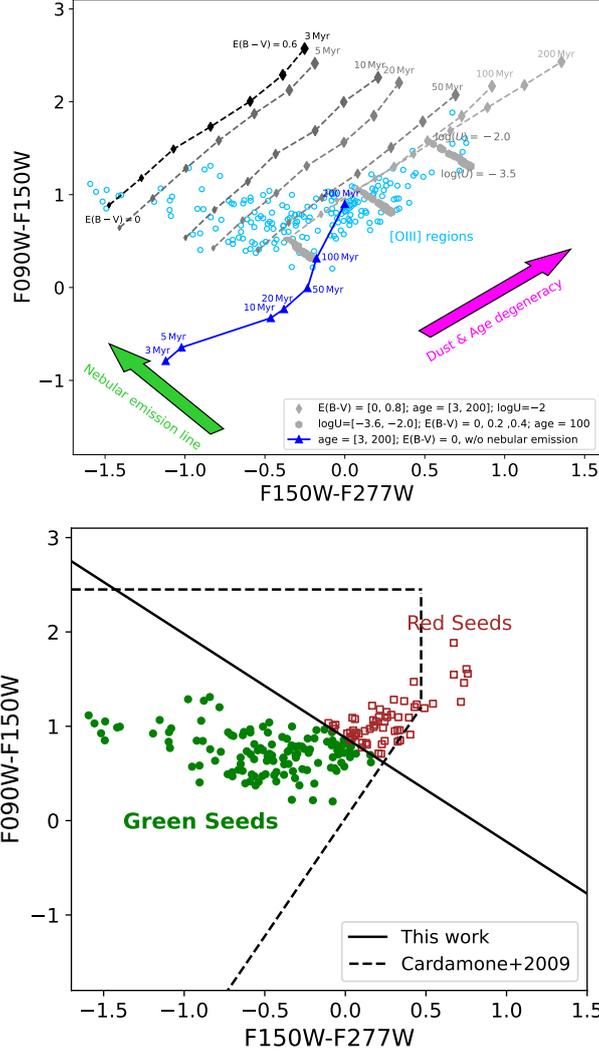
We overlay the color indices from the SED models as grids on the upper panel of Figure 6.5. The diamond grids with dashed lines exhibit how the color indices change with dust attenuation and stellar age at a fixed  $\log U = -2$ . The sizes of the diamond grids increase with dust attenuation level from  $E(B - V) = 0$  to 0.8. The grayscale of the diamonds grids represents the stellar ages, with the darkest shading indicating the youngest age of 3 Myr and the lightest the oldest age of 500 Myr. As suggested by blue solid line and black dashed line, both increasing the age of a stellar population and increasing the amount of dust can redden the spectra in F090W – F150W and F150W – F277W, known as the age-dust degeneracy occurring at the rest-frame optical light. Also, spectra with younger stellar populations exhibit more obvious flux excesses in F150W from [OIII]+H $\beta$  emission line at fixed  $\log U$  because of their lower stellar continuum.

The age-dust degeneracy raise an issue on whether the identified redder [OIII]+H $\beta$  regions are dusty young starburst or containing older stellar population. Actually, the stellar continuum could become discontinuous by Balmer break, and the amount of the break depends on the age of the galaxy. This may raise uncertainties on the estimation of [OIII]+H $\beta$  emission line fluxes. Thus, we develop a new color selection criterion by fitting the SED model grids. For young galaxies with little dust content,  $E(B - V)$  values are typically small, ranging from 0 to 0.15 magnitudes (e.g., Nakajima et al. 2016; Tang et al. 2019). Thus, we set  $E(B - V) = 0.15$  and stellar age equals to 100 Myr as the threshold to classify two populations. We apply the linear fits to the SED model grids based on ionization parameters from  $\log U = -2$  to  $-3.6$ . The resulting color limit is:

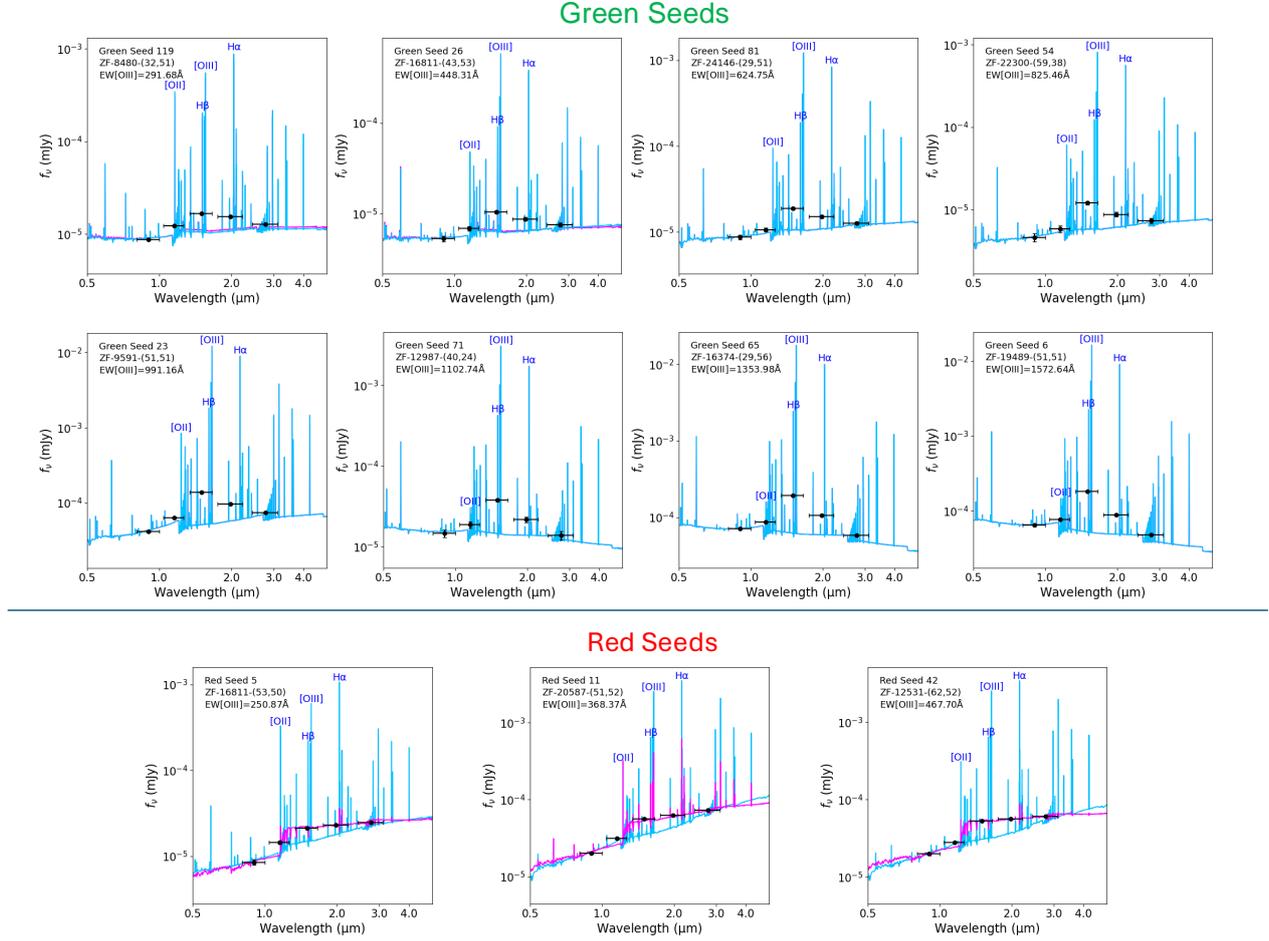
$$\text{F090W} - \text{F150W} = -1.1 \times (\text{F150W} - \text{F277W}) + 0.9 \quad (6.5)$$

We display this color selection in the bottom panel of Figure 6.5, separating all the [OIII]+H $\beta$  regions into 128 normal ones and 59 older and/or dusty ones, shown as green filled circles and brown open squares, respectively. Those 128 [OIII]+H $\beta$  regions are named as ‘‘Green Seeds’’, in reference to the famous Green Pea galaxies at  $z \sim 0.2$  (Cardamone et al. 2009). The Green Pea galaxies were discovered in the citizen science project Galaxy Zoo, exhibiting peculiar bright green colors in Sloan Digital Sky Survey (SDSS; York et al. 2000; Kauffmann et al. 2003) imaging due to very strong [OIII]+H $\beta$  emission lines with EWs up to  $\sim 1000$  Å. Green Seeds in this study have a similar appearance on the JWST RGB images. On the other hand, other emission line regions are named as ‘‘Red Seeds’’.

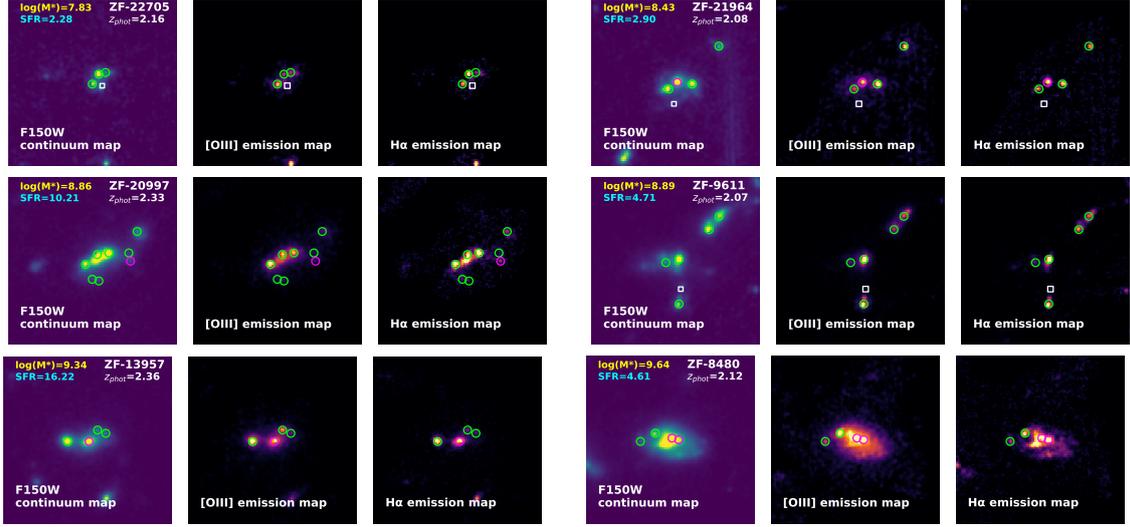
Cardamone et al. (2009) defined an  $r - z$  vs.  $u - r$  color selection criteria by comparing Green Pea galaxies to a large sample of local star-forming galaxies (SFGs) and quasars. Since the three SDSS filters ( $u$ ,  $r$ ,  $z$ ) have similar rest-frame wavelengths to the JWST filters (F090W, F150W, F277W), we also incorporate their color selection criteria in the bottom panel of Figure 6.5. The color selection from Cardamone et al. (2009) does not clearly distinguish Green Seeds from Red Seeds, as identified by our color selection (Equation 6.5). In Figure 6.6, we present a series



**Figure 6.5.** Upper: The F150W – F277W versus F090W – F150W color-color diagram for the 187 [OIII]+H $\beta$  regions. A series of SED model grids are overlaid with varying symbols. The blue triangles represent the stellar continuum of a burst SFH from 3 to 200 Myr without emission lines and dust attenuation. Emission-line templates from CLOUDY (Ferland et al. 1998, 2013) and dusty attenuation are then added to these pure SED models. The diamonds represent SED models with varying stellar ages and dust attenuation at a fixed ionization parameter of  $\log U = -2$ . The sizes of the diamond grids increase with dust attenuation from  $E(B - V) = 0$  to 0.8, as labeled. The grayscale of the diamond grids indicates the stellar ages of the SED models, with black representing 3 Myr and silver representing 200 Myr. The hexagons show SED models with varying ionization parameters and dust attenuation at a fixed stellar age of 100 Myr. From the bottom-left to upper-right, the grids have dust attenuation levels of  $E(B - V) = 0, 0.2, 0.4$ . The sizes of the hexagon grids increase with decreasing ionization parameter from  $\log U = -2$  to  $-3.6$ , evolving from upper-left to bottom-right, as labeled. Bottom: The color selection derived from the linear fitting of these SED model grids at a fixed  $E(B - V) = 0.15$  and 100 Myr with varying  $\log U$  is shown as the black solid line. On the bottom-left side of this color selection, we classify 128 [OIII]+H $\beta$  regions with young age and little dust, and designate them as “Green Seeds”, referencing the local Green Pea galaxies (Cardamone et al. 2009). These Green Seeds are marked as green filled circles. Red [OIII]+H $\beta$  regions are located on the other side, shown as brown open squares, named as “Red Seeds”. These regions are highly dust-attenuated and/or contain older stellar population. For comparison, the  $r - z$  vs.  $u - r$  color selection from (Cardamone et al. 2009) is also shown here, after being redshifted to  $z=2.2$  and corrected from the SDSS magnitude to AB magnitude.



**Figure 6.6.** Eight example SEDs of Green Seeds and three example SEDs of Red Seeds are given in the small panels. The blue and magenta spectra represent the best-fit SEDs, with the blue spectra constrained by the largest stellar age of 100 Myr. Green Seeds favor young stellar populations with low dust attenuation. In contrast, due to the limited photometric data, Red Seeds could either choose younger spectra with large dust content or older spectra with less dust, i.e., the age-dust degeneracy. These examples validate the feasibility of the color selection in Figure 6.5 for identifying Green Seeds with young stellar populations and low dust attenuation. Also, the spectra demonstrate that assuming a linear relationship between the stellar continuum in the F115W and F277W filters does not introduce large uncertainties.



**Figure 6.7.** The stellar continuum map at  $1.5 \mu m$ , the  $[OIII]+H\beta$  emission line map, and the  $H\alpha$  emission line map for six HAEs that contain Green Seeds. These galaxies are ordered by their stellar mass. Green Seeds are marked as green circles, while Red Seeds are shown as magenta circles. Also, white squares represent regions with  $S/N < 2.5$ , similar to Figure 6.4.

of SED samples for both Green Seeds and Red Seeds using the SED fitting code CIGALE (Noll et al. 2009; Boquien et al. 2019). The blue and magenta spectra in each panel represent both the best-fit SED, but for the blue ones, maximum stellar age of the SED fitting is limited to 100 Myr. It is clearly illustrated that Green Seeds tend to favor younger spectra with less dust content in the both fitting, and their stellar continuum can be well estimated by a linear interpolation between the F115W and F277W filters. On the other hand, Red Seeds suffer from the age-dust degeneracy, where it is unclear whether younger stellar population and strong emission lines with high dust attenuation or older population with low dust attenuation represent the true SED, given only five JWST photometric data points.

In Figure 6.7, we display the stellar continuum map at  $1.5 \mu m$ , the  $[OIII]+H\beta$  emission line map, and the  $H\alpha$  emission line map for six samples among the total 68 HAEs that contain Green Seeds. They are ordered by stellar mass, from lower to higher. The 68 HAEs have stellar mass from  $10^{7.8} M_{\odot}$  to  $10^{10} M_{\odot}$ . The continuum and emission line maps of the remaining HAEs with Green Seeds are displayed in Appendix C. In this paper, we mainly focus on the properties, origins, and evolution of Green Seeds, in conjunction with other populations, including Red Seeds and HII regions.

## 6.4 The properties of Green Seeds

### 6.4.1 Stellar mass

To derive the stellar masses of the 128 individual Green Seeds, we also use the SED fitting code CIGALE with emission-line templates. The parameter settings are

the same as those introduced in Section 2.3, and are also used in Figure 6.6 to derive the best-fit SED. While the best-fit model itself does not provide direct uncertainty estimates, we further adopt a Bayesian-like fitting approach, assigning weights to all models based on their  $\chi^2$  values. This method allows for more accurate estimates of physical properties. In this study, the uncertainties in the stellar masses of the seeds are directly obtained from CIGALE.

Given that we only utilize five JWST filters from JADES to achieve better angular resolution in this study, the limited number of photometric fluxes may lead to large uncertainties in stellar mass estimation. Thus, in addition to applying the standard SED fitting techniques, we also estimate the stellar masses of Green Seeds from the mass-to-light ratio method (Bell & de Jong 2001). McGaugh & Schombert (2014) introduced an approach to estimate stellar mass using an updated color-mass-to-light ratio (CMLR), constructed by population synthesis models with the following relation:

$$\log \Upsilon_*^k = a_k + b_k (B - V), \quad (6.6)$$

where  $k$  refers to the selected filters and the footnote  $*$  equals to the color of  $B - V$ . In this study, we prefer using the rest-frame  $I$  band (F277W), denoted as  $\Upsilon_*^I$ , to minimize the effects of underlying emission lines. The  $B - V$  color can be obtained from the stellar continuum in F115W and F150W, that the stellar continuum of F150W are obtained from Equation 6.1. The stellar mass of each Green Seeds is then derived from the rest-frame  $I$ -band luminosity by  $M_* = \Upsilon_*^I L_I$ . The uncertainties in the stellar masses are inferred from the flux errors in F277W within the aperture of the Seeds, combined with the uncertainties in the  $B - V$  color, derived from the flux errors in F115W and F150W. In Figure 6.8, we compare the stellar mass distributions of Green Seeds inferred from SED and CMLR, displaying the results as histograms. We demonstrate that these two methods provide closely aligned estimations of stellar masses.

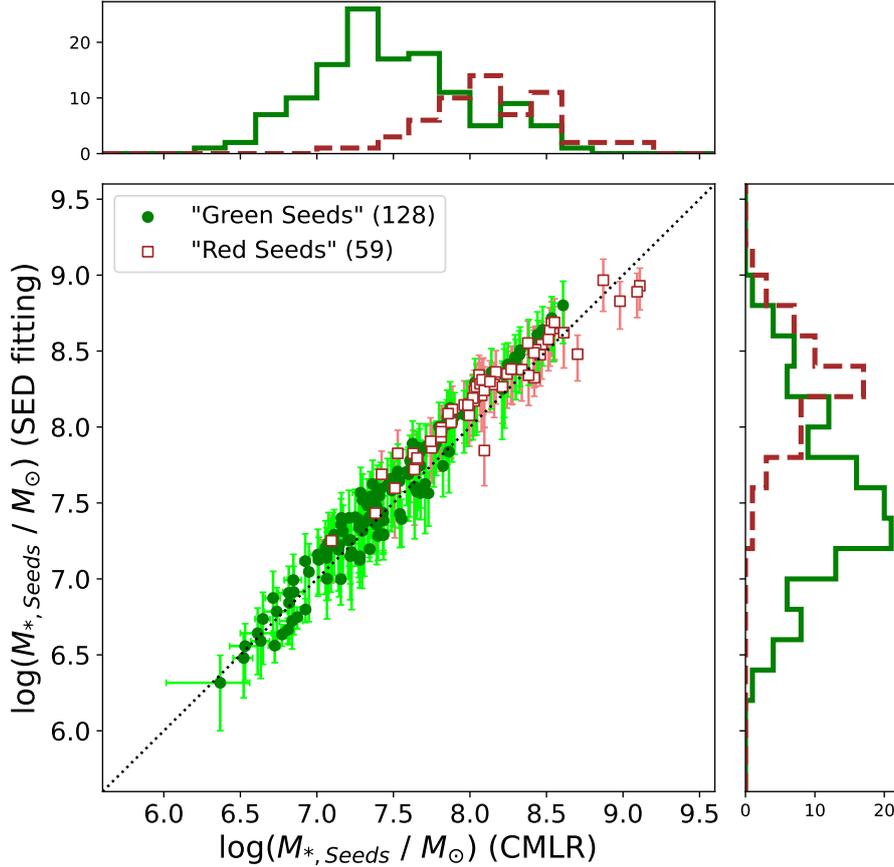
Most Green Seeds have stellar masses in the range of  $10^{6.5} M_\odot$  to  $10^{8.5} M_\odot$ , with a median stellar mass of  $10^{7.4} M_\odot$ . When compared to the total stellar mass of their host galaxies, the stellar mass of Green Seeds contributes to  $\sim 3\%$  of the total mass (with a median of 1.5 dex lower). We also show the stellar mass distribution of Red Seeds in the same panel, which exhibits a  $\sim 1$  dex higher median stellar mass than that of Green Seeds.

## 6.4.2 Star formation rate

To determine the SFRs in Green Seeds, we measure the  $H\alpha$  luminosity within the same aperture by using the  $H\alpha$  emission line map, as shown in Figure 6.7. By analyzing the flux excess in the F200W filter, we calculate the total  $H\alpha$  luminosity and convert it to SFR by the calibration from Kennicutt & Evans (2012), with a correction applied for the Chabrier (2003) IMF:

$$\log \text{SFR}(H\alpha) = \log L_{H\alpha} - 41.34. \quad (6.7)$$

The correction for dust attenuation is based on the median  $E(B - V)_{neb} = 0.1$  obtained from the SED fitting results of Green Seeds. In the F200W filter, the main



**Figure 6.8.** Comparison between the stellar masses of the resolved emission line regions estimated from SED fitting and the color-mass-light ratio (CMLR). The 128 Green Seeds are marked as green filled circles, while the 59 Red Seeds are shown as brown open squares. The overall estimation of stellar mass from SED fitting agrees well with the CMLR measurements. The stellar masses of Green Seeds and Red Seeds are also distributed as histograms on both axes in green and brown, respectively.

contaminants include [NII] and [SII] emission lines. We have applied a correction of contamination, assuming the contamination ratio of 0.1 which are median values of the best-fit SED templates.

Following the methodology described in Section 6.2, we measure the  $H\alpha$  luminosity in the same aperture as used for Green Seeds, derived from the  $H\alpha$  emission line map in the F200W filter. A majority (112 out of 128) of Green Seeds show  $H\alpha$  emission line detection, contributing to clear flux excesses of  $S/N > 2.5$  in F200W. For those Green Seeds without strong flux excesses, we take the  $2.5\sigma$  upper limit fluxes for the  $H\alpha$  measurements.

In the upper panel of Figure 6.9, we show the  $SFR(H\alpha)$  as a function of the stellar mass of Green Seeds. Generally, SFGs exhibit a correlation between stellar mass ( $M_*$ ) and SFR, known as the SFMS, which holds true up to at least  $z \sim 3$  (e.g., Whitaker et al. 2014; Speagle et al. 2014). Green Seeds also appear to follow the SFMS with similar slopes to that from Whitaker et al. (2014). However, the normalization of this relation is much higher for Green Seeds compared to their host galaxies, indicating a higher specific SFR (sSFR) for Green Seeds. Since higher

molecular gas fractions are a key driver of the SFMS evolution (higher normalization) across redshift (Carilli & Walter 2013), we infer that Green Seeds may contain a higher molecular gas fraction than their host galaxies.

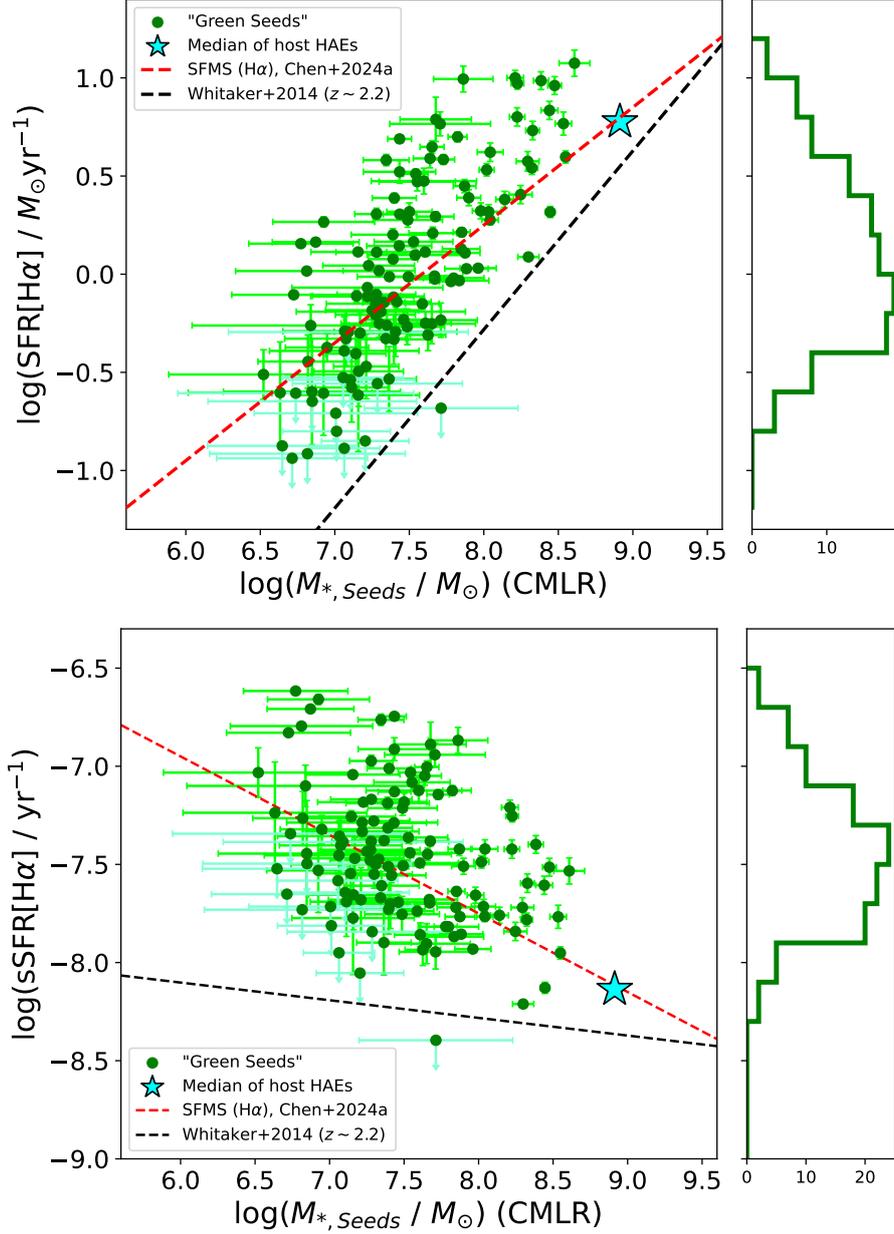
We further illustrate the SFMS by displaying the correlation between sSFR and the stellar mass in the bottom panel of Figure 6.9. The inverse of sSFR ( $1/\text{sSFR}$ ) directly indicates the time taken to double the stellar mass in the system. The median sSFR for Green Seeds reaches  $-7.6 \text{ yr}^{-1}$ , which is  $\sim 0.6$  dex higher than that of their host galaxies, suggesting  $\sim 4\times$  more rapid stellar mass assembly in Green Seeds. This supports the idea that starburst-like activities are occurring in Green Seeds.

### 6.4.3 Equivalent width

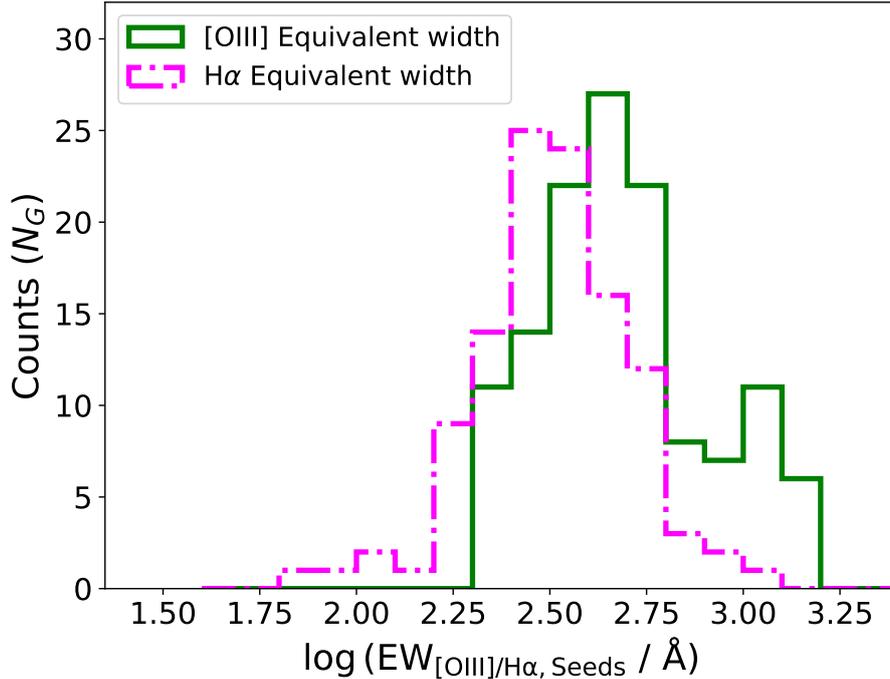
Following Equation 6.3, we construct the rest-frame  $\text{EW}_{[\text{OIII}]+\text{H}\beta}$  map of each HAE. By applying the circular apertures derived from peak finding, we obtain the average  $\text{EW}_{[\text{OIII}]+\text{H}\beta}$  of each  $[\text{OIII}]+\text{H}\beta$  emission line region. To ensure accurate measurements of  $[\text{OIII}]$  line fluxes, we assume that the total flux excesses in the F150W filter are contaminated by  $\text{H}\beta$ , following the approach in Section 2.4. Assuming Case-B recombination with  $T_e = 10,000\text{K}$  and  $n_e = 100 \text{ cm}^{-3}$ , we derive the intrinsic  $\text{H}\beta$  fluxes from the intrinsic  $\text{H}\alpha$  fluxes using  $F_{\text{H}\beta,\text{int}} = F_{\text{H}\alpha,\text{int}}/2.86$ , where the  $\text{H}\alpha$  flux is obtained as described earlier.

Figure 6.10 displays the rest-frame  $\text{EW}_{[\text{OIII}]}$  distribution for Green Seeds. Green Seeds have quite high equivalent width reaching to a median value of  $\text{EW}_{[\text{OIII}]}, \text{med} = 452\text{\AA}$ . Since  $\text{EW}_{[\text{OIII}]}$  serves as an indicator of the ionization state of the ISM (e.g., Reddy et al. 2018), this implies a very intense ionizing radiation field in Green Seeds. Similar results are also found for  $\text{EW}_{\text{H}\alpha}$ , as shown as magenta dot-dashed line in Figure 6.10. Green Seeds have a median equivalent width of  $\text{EW}_{\text{H}\alpha}, \text{med} = 320\text{\AA}$ . More directly, the intensity of  $\text{H}\alpha$  emission line traces the recent star formation activities within the past 10 Myr. The higher  $\text{H}\alpha$  EWs reflect the younger stellar populations in Green Seeds and explain the higher sSFRs observed in the bottom panel of Figure 6.9.

We find that 17 Green Seeds exhibit extremely large  $\text{EW}_{[\text{OIII}]} > 1000\text{\AA}$ , a phenomenon previously reported only in a resolved case study at  $z \sim 2$  (Zanella et al. 2015). On the other hand, such large  $[\text{OIII}]$  EWs are more commonly observed in integrated studies of EELGs at similar redshift (e.g., Nakajima & Ouchi 2014; Tang et al. 2019). These studies also suggested a close connection between extreme  $[\text{OIII}]$  emission and Lyman continuum (LyC) leakage. As suggested in the previous section that low-mass HAEs are potential lower-redshift analogs of the galaxies that reionized the universe during cosmic reionization at  $z > 6$ . Our high-resolution observations from JWST may be capturing possible resolved LyC leakage regions in galaxies at  $z \sim 2$ , which has so far only been reported in the local universe (Izotov et al. 2021). These Green Seeds with  $\text{EW}_{[\text{OIII}]} > 1000\text{\AA}$  are likely dominated by extremely hot and massive stars, leading to a more intense radiation field, and creating so-called “density-bounded” HII regions (Nakajima & Ouchi 2014) from which ionizing photon have a higher probability of leaking into the IGM.



**Figure 6.9.** Based on the F200W photometry, we require H $\alpha$  emission lines show flux excesses with  $S/N > 2.5$  in every Green Seeds, and nearly 90% of them meet this requirement. Upper: The SFR(H $\alpha$ ) as a function of stellar mass, i.e., the “SFMS” in Green Seeds distributed as green filled circles. Downward arrows indicate the  $2.5\sigma$  upper limits for the H $\alpha$ -undetected regions. The median  $M_*$  and integrated SFR of the host galaxies that contain Green Seeds is indicated by the cyan star. The red dashed line is the extrapolated best-fit SFMS from the parent sample of HAEs, while the black dashed one is from Whitaker et al. (2014) for  $z \sim 2.2$  galaxies. The mass completeness of both literatures are  $\sim 10^9 M_{\odot}$ . Bottom: The sSFR(H $\alpha$ ) as a function of stellar mass in Green Seeds, with the same outlines as in the upper panel. The cyan star shows the median value of the integrated sSFR of the host galaxies. Green Seeds have a much higher sSFR by  $\sim 0.6$  dex compared to their host galaxies. Also, they are generally above the extrapolated SFMS(H $\alpha$ ) from Whitaker et al. (2014).



**Figure 6.10.** The distribution of [OIII] and H $\alpha$  EWs for Green Seeds, shown as histograms in green and magenta, respectively. This analysis of H $\alpha$  EWs only includes those Green Seeds with H $\alpha$  flux excesses of  $S/N > 2.5$ , which contains a sample of 111 regions. The median equivalent width of  $EW_{[\text{OIII}]}$ ,  $\text{med} = 452\text{\AA}$  and  $EW_{\text{H}\alpha}$ ,  $\text{med} = 320\text{\AA}$ , respectively.

## 6.5 Comparison between “Green Seeds” and other resolved structures

### 6.5.1 Relevance between Green Seeds and HII regions

In the previous section, we estimated the SFRs in Green Seeds through H $\alpha$  luminosity. We figure out that nearly 90% of Green Seeds exhibit strong flux excesses with  $S/N > 2.5$  in the F200W filter, primarily driven by H $\alpha$  emission lines. Interestingly, 16 Green Seeds do not show strong flux excesses in F200W, but still have the excesses with at least  $S/N > 1$  for the H $\alpha$  emission lines. The intrinsically weaker H $\alpha$  emission lines (compared to [OIII]) at  $z \sim 2$  (Sanders et al. 2018; Reddy et al. 2018) could be the main reason for the lower  $S/N$ . Nevertheless, it is reasonable to assert the prevalence of both [OIII] and H $\alpha$  emission lines in the detected emission line regions.

Although not shown in Figure 6.4, we also apply the same peak finding algorithm to the  $EW_{\text{H}\alpha}$   $S/N$  maps to extract the H $\alpha$  emission line regions (HII regions) at  $z \sim 2.2$ . We identify 164 regions with  $S/N > 2.5$  and  $EW_{\text{H}\alpha} > 100\text{\AA}$ . Among these HII regions, we find that 6 regions do not have flux excesses  $S/N > 2.5$  in the F150W filter, reflecting weak [OIII] emission lines in them. Note that five of these 6 regions are located in massive galaxies that do not contain Green Seeds. This finding further suggests the ubiquity of both strong [OIII] and H $\alpha$  emission lines in the emission line regions at Cosmic Noon.

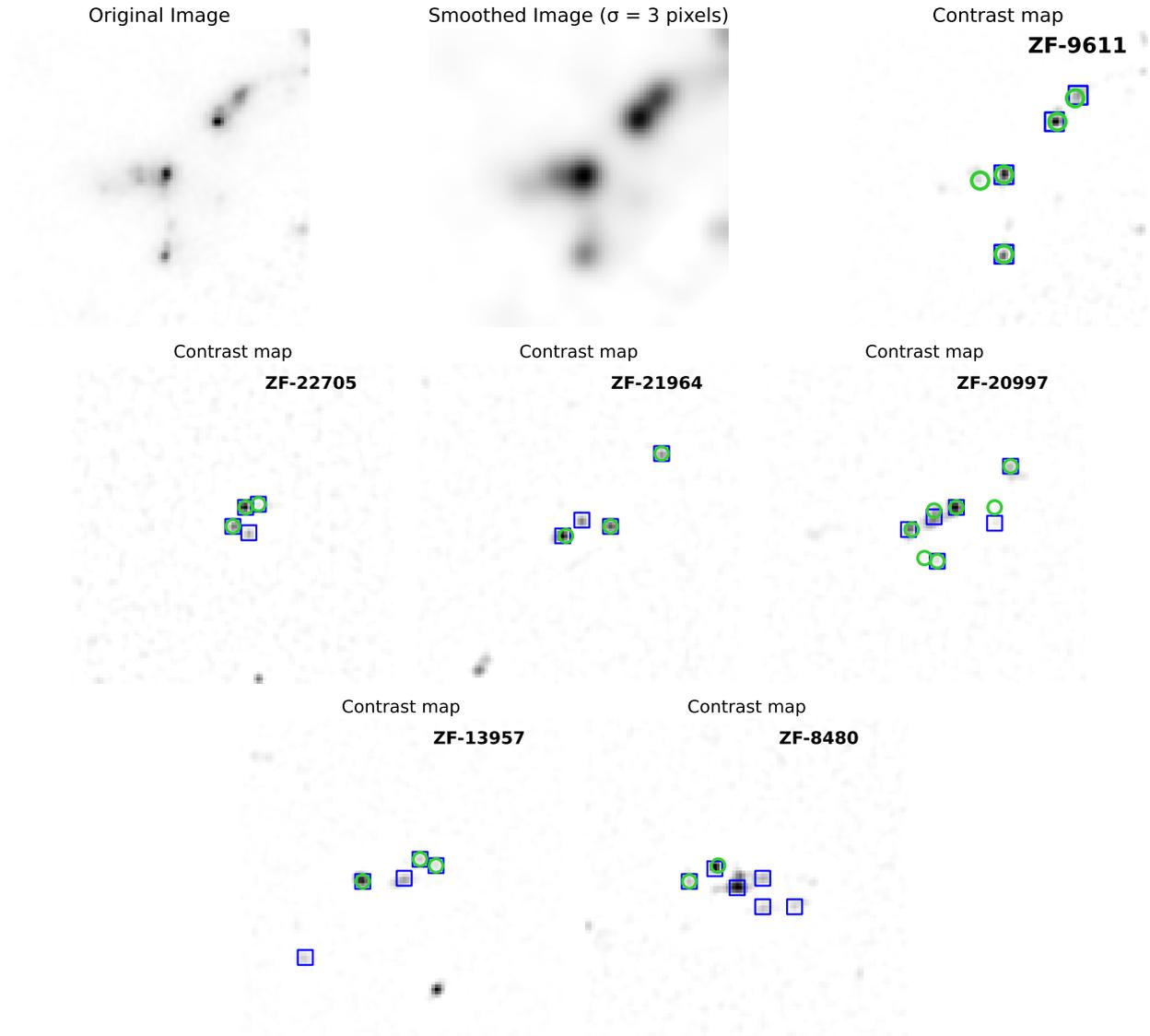
## 6.5.2 Comparison with UV star-forming clumps

The properties of these kpc-scale Green Seeds suggest that they are star-burst regions with hard ionizing radiation fields, recalling the star-forming clumps. Previous observations with the Hubble Space Telescope (HST) have revealed that many galaxies at  $z \simeq 1 - 3$  host discrete rest-frame ultraviolet (UV) star-forming clumps of similar scale (e.g., [Conselice et al. 2004](#); [Elmegreen & Elmegreen 2005](#); [Elmegreen et al. 2007](#); [Förster Schreiber et al. 2011b](#); [Wuyts et al. 2012](#); [Guo et al. 2012, 2015, 2018](#); [Tadaki et al. 2014](#); [Livermore et al. 2015](#); [Garland et al. 2015](#); [Shibuya et al. 2016](#); [Calabrò et al. 2019](#); [Claeyssens et al. 2023](#)). These UV-bright clumps have been shown to have enhanced sSFRs, higher than those of their surrounding areas by a factor of several. However, these previous studies mostly focused on clumps in massive galaxies with stellar mass larger than  $10^{10} M_{\odot}$  and within larger apertures than used in this study, such as  $0.''30$  in [Guo et al. \(2012\)](#), due to the limitations in resolution and depth of HST observations. UV star-forming clumps in low-mass hosting galaxies have rarely been studied, with most investigations relying on the gravitational lensing effect (e.g., [Vanzella et al. 2022](#)). With the unprecedented resolution and depth provided by JWST/NIRCam, it is worth exploring UV star-forming clumps in HAEs, especially those in low-mass galaxies that have only been resolved with the aid of gravitational lensing ([Vanzella et al. 2023](#)).

We utilize the convolved F090W images and apply an automated clump-detection algorithm, following methods widely used for detecting clumps (e.g., [Conselice 2003](#); [Guo et al. 2015](#); [Kalita et al. 2024](#)). We first smooth the F090W images by a Gaussian filter with size of  $\sigma = 3$  pixels. Then, we subtract the smoothed image from the original image to create a contrast map. Next, we again run the `findpeaks` Python Package on the contrast map, similar to how we extracted the emission line regions in Section 3. This method successfully selects regions with peaked stellar continuum. In Figure 6.11, we illustrate the clump-detection method using the same six HAEs as in Figure 6.7. Previous works used Gaussian filters ranging from 3 – 6 pixels (e.g., [Kalita et al. 2024](#)). In our case, we find that a 3-pixel Gaussian filter works best for detecting clumps through visual inspection, and most UV star-forming clumps with typical sizes of  $\sim 1$  kpc can be identified in the contrast image. The remaining contrast maps with UV star-forming clumps are displayed in Appendix C.

Based on these contrast maps, we compare Green Seeds with the UV star-forming clumps selected from JWST observations. We find that many Green Seeds are also identified as UV stellar continuum clumps. This result is not surprising because UV clumps observed with HST also differ from their surrounding areas in terms of younger stellar age, higher sSFR and lower dusty attenuation ([Guo et al. 2012](#)), similar to the appearance of Green Seeds. Moreover, compared to the UV clumps observed with HST at  $z \sim 2$ , which are mostly more massive than  $10^8 M_{\odot}$ , the clumps observed with JWST have a wider mass range, covering a large number of low stellar mass components down to less than  $10^7 M_{\odot}$ .

From the contrast map in Figure 6.11 and Appendix C, we find more than half (70 out of 128) of the Green Seeds are co-located with the UV star-forming clumps. [Faisst et al. \(2024\)](#) combined JWST and ALMA data, finding that galaxies hosting UV star-forming clumps are likely to have larger gas fraction and higher



**Figure 6.11.** Illustration of the clump-detection process for HAEs. First, the original F090W image is smoothed using a Gaussian filter. The smoothed image is then subtracted from the original image to create a contrast image. The UV stellar continuum clumps are identified from the contrast image, which are marked as blue squares. For reference, Green Seeds are shown as green circles. The results shown here correspond to the six HAEs presented in Figure 6.7, with additional samples listed in Appendix C.

star formation efficiency, suggesting a connection between UV star-forming clumps and gas inflows replenishment. Such gas inflow could also be responsible for the formation of Green Seeds observed in this study in the rest-frame optical wavelength range. Additionally, [Kalita et al. \(2024\)](#) analyzed F356W data of galaxies at  $z = 1-2$  and found that clumps are not limited to UV light but are also prominently detected in near-IR wavelengths. These findings suggest the presence of clump structures across a broad wavelength range, indicating their role in both star formation and stellar mass distribution.

On the other hand, more than one-third of the Green Seeds are “off-peak” to the UV star-forming clumps, with separations ranging from one pixel ( $\sim 200$  pc) to several pixels in the cutout stamps. Physically, the “off-peak” appearance of emission line regions and UV stellar continuum could partly be explained by the different timescales of SFR indicators: UV light traces star formation over the past  $\sim 100$  Myr, while emission lines such as  $H\alpha$  trace a shorter timescale of  $\sim 10$  Myr (e.g., [Kennicutt & Evans 2012](#); [Theios et al. 2019](#)). In our case, compared to UV clumps, a higher fraction ( $\sim 75\%$ ) of Green Seeds are co-located with  $H\alpha$  emission line peaks. This finding suggests that Green Seeds are more likely to track the star formation on shorter timescales. Another possible explanation could involve differences in dust attenuation. UV clumps typically experience stronger dust extinction compared to emission line regions in the galactic disk. This variation in dust extinction across the galactic disk may also contribute to the “off-peak” alignment between UV clumps and emission line regions. The third possible explanation for the “off-peak” appearance could be outflows driven by star formation feedback (e.g., [Rich et al. 2010](#); [Heckman & Thompson 2017](#); [Cresci et al. 2017](#); [Förster Schreiber et al. 2019](#)). Large amounts of ionized gas in galaxies could be pushed outward by the energy imparted from star formation activity. Observations of local SF-driven winds imply more efficient outflows in lower-mass galaxies (e.g., [Heckman et al. 2015](#)). While the complex nature of these winds make them harder to be identified at high redshift. The outflow velocity maps from future high resolution IFU observations will be necessary to further constrain this viewpoint quantitatively. Note that there is no indication of AGN activity in the HAEs that contain Green Seeds (see Appendix C)

### 6.5.3 Comparison with Local HII regions

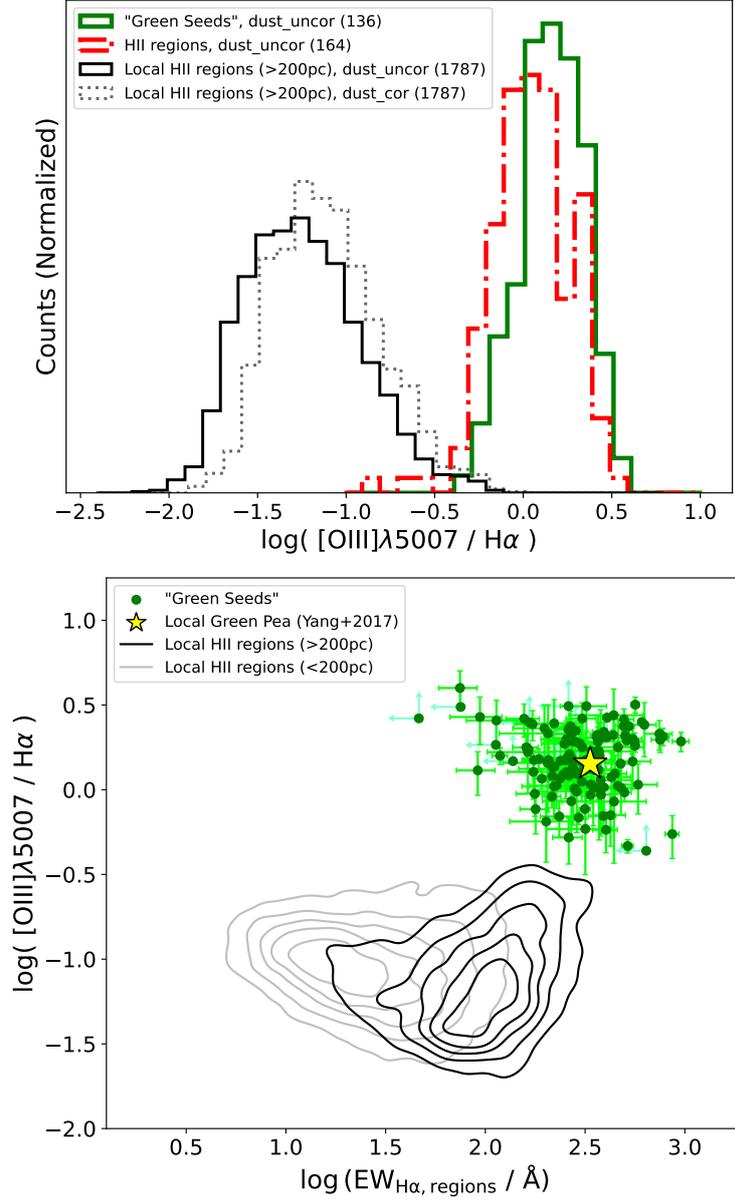
Local HII regions, to some extent, share some similarities with Green Seeds, such as higher sSFR and lower chemical abundance compared to their host galaxies. However, given the significant differences on the ISM properties between the local universe and Cosmic Noon (e.g., [Erb et al. 2006](#); [Liu et al. 2008](#); [Shapley et al. 2015](#); [Sanders et al. 2018](#); [Reddy et al. 2018](#)), it is intriguing to compare local HII regions with Green Seeds at higher redshift. This comparison may reveal how galaxy and gas properties have evolved from early epochs to the present day. For this comparison, we refer to the recent PHANGS-MUSE survey ([Emsellem et al. 2022](#)), which conducted IFU observations on 19 star-forming disc galaxies in the local universe, studying the physical properties of HII regions inside these galaxies with a physical resolution down to 100 pc. These local HII regions are constructed from

the  $H\alpha$  emission line map, with a detection threshold to  $3\sigma$  above the background. The PHANGS-MUSE HII region catalogs (Santoro et al. 2022; Groves et al. 2023) provide flux measurements and kinematic information for multiple optical emission lines in each HII region, along with the region area and physical properties.

In the PHANGS-MUSE survey, the median physical resolution is 7 – 20 times higher than that of JWST observations at  $z \sim 2$  using the F277W filter (0.7 kpc) in this study. As a result, the catalog contains local HII region with areas ranging from 40 pc to 800 pc, with an average size of  $\sim 110$  pc. This is smaller than the 1.2 kpc circular aperture ( $0.''15$ ) used to identify Green Seeds in this study. To obtain a sample comparable to Green Seeds, we select local HII regions with sizes larger than 200 pc (assuming a circular shape inferred from the `region_area` in the catalog), resulting in a sample of 1,787 local HII regions.

Based on the estimation of  $[OIII]$  and  $H\alpha$  emission line intensities in Section 6.4, we first compare the  $[OIII]\lambda 5007/H\alpha$  ratios of Green Seeds and local HII regions in the upper panel of Figure 6.12 (using the line ratio of  $[OIII]\lambda 5007 : [OIII]\lambda 4959 = 2.97 : 1$  for conversion). Here, the PHANGS data are derived from the  $H\alpha$  emission map, while Green Seeds are extracted from  $[OIII]$  emission map. The selection bias may introduce a higher  $[OIII]\lambda 5007/H\alpha$  ratio in our sample. Thus, we also utilize the HII regions at  $z \sim 2$  from Section 5.1, which were first derived from  $H\alpha$  emission map, for additional comparison. The PHANGS-MUSE survey derived dust attenuation levels from the Balmer decrement, assuming an intrinsic Balmer ratio of  $H\beta = H\alpha/2.86$ , giving a higher  $[OIII]\lambda 5007/H\alpha$  shown as the dashed histogram. On the other hand, considering the complexity of dust attenuation in HII regions at  $z \sim 2$  in this study, we do not correct for dust attenuation, instead comparing the observed  $[OIII]\lambda 5007/H\alpha$  ratios for our samples in this discussion. We find that Green Seeds have a median ratio of  $\log([OIII]\lambda 5007/H\alpha) = 0.16$ . Despite being selected from different emission line maps from JWST, Green Seeds and HII regions from the same host galaxies only show a slight discrepancy in the  $[OIII]\lambda 5007/H\alpha$  ratio, with a larger median value by  $\sim 0.1$  dex in Green Seeds. In contrast, Green Seeds exhibit a much higher  $[OIII]\lambda 5007/H\alpha$  ratio than local HII regions by  $\sim 1.5$  dex ( $\sim 30$  times). If we assume a fixed ratio between  $H\alpha$  and  $H\beta$  ( $H\beta = H\alpha/2.86$ ), our sample is constant with the well-documented obvious enhancement in  $[OIII]\lambda 5007/H\beta$  ratio at  $z \sim 2$  in previous studies (e.g., Liu et al. 2008; Yabe et al. 2014; Steidel et al. 2014; Shapley et al. 2015). The traditional BPT diagram (Baldwin et al. 1981) gives a theoretical maximum of  $\log([OIII]\lambda 5007/H\beta) \simeq 0.9$  for SFGs (Kewley et al. 2001; Steidel et al. 2014; Nakajima & Maiolino 2022), and we find that most of our Green Seeds are close to, but within, this upper limit. This “maximum starburst” limit requires a high ionization parameter ( $\log U > 2$ ) and hard ionizing spectra ( $T_{\text{eff}} \sim 50,000$  K) in the photoionization models from Steidel et al. (2014). We also find that four Green Seeds exceed the “maximum starburst” limit with their  $[OIII]\lambda 5007/H\beta$  ratios. One possible explanation is the contamination by AGNs. Alternatively, Gutkin et al. (2016) and Feltre et al. (2016) used state-of-the-art stellar population and photoionization models with a wider range of parameters to model HII regions at higher redshift without AGNs, suggesting that  $\log([OIII]\lambda 5007/H\beta)$  as high as  $\sim 1$  are possible in their models.

The PHANGS-MUSE HII region catalogs also provide  $EW_{H\alpha}$  for local HII re-



**Figure 6.12.** Upper: The  $[\text{OIII}]\lambda 5007/\text{H}\alpha$  ratios for Green Seeds, HII regions at  $z \sim 2$  from Section 5.1, and local HII regions ( $> 200$  pc) from the PHANGS-MUSE survey (Groves et al. 2023) are normalized distributed in a histogram. Green Seeds are represented by the green solid line, while HII regions at  $z \sim 2$  are shown as a red dotted-dashed line. For Green Seeds with weak  $\text{H}\alpha$  emission, we use the  $2.5\sigma$  upper limit fluxes (see Section 4.2). We do not correct for dust attenuation in Green Seeds or HII regions at  $z \sim 2$ , presenting the observed  $[\text{OIII}]\lambda 5007/\text{H}\alpha$  ratios. For comparison, both dust-corrected and uncorrected line ratios of the local HII regions are exhibited as grey and black lines, respectively. While Green Seeds and HII regions at  $z \sim 2$  exhibit similar  $[\text{OIII}]\lambda 5007/\text{H}\alpha$  ratio, there is a clear discrepancy when compared to local HII regions. Bottom: Relationship between  $\text{EW}_{\text{H}\alpha}$  and  $[\text{OIII}]\lambda 5007/\text{H}\alpha$  ratios for Green Seeds distributed as green circles. Local HII regions with sizes  $> 200$  pc ( $< 200$  pc) are represented by black (grey) contours. Despite having comparable  $\text{EW}_{\text{H}\alpha}$ , Green Seeds display the highest  $[\text{OIII}]\lambda 5007/\text{H}\alpha$  ratios. Additionally, the  $\text{EW}_{\text{H}\alpha}$  and  $[\text{OIII}]\lambda 5007/\text{H}\alpha$  ratios in Green Seeds closely correspond to those of local Green Pea galaxies from Yang et al. (2017b), with the median value of their 43 samples indicated by a yellow star.

gions. The  $\text{EW}_{\text{H}\alpha}$  shows the most solid correlation with stellar age and sSFR among the optical emission lines (Reddy et al. 2018). We further examine the relationship between the  $\text{EW}_{\text{H}\alpha}$  and  $[\text{OIII}]\lambda 5007/\text{H}\alpha$  ratio in the bottom panel of Figure 6.12. We find that local HII regions larger than 200 pc have  $\text{EW}_{\text{H}\alpha}$  comparable to Green Seeds, within 0.5 dex, but their  $[\text{OIII}]\lambda 5007/\text{H}\alpha$  ratios are almost 1.5 dex lower. This suggests that while both local HII regions and Green Seeds are primarily composed of young stellar population ( $\sim 10$  Myr) with comparable sSFR, they differ significantly in their ionization properties. The presence of top-heavy IMFs in Green Seeds may explain this phenomenon, as an increased number of high-energy photons could result in more doubly ionized oxygen ( $\text{O}^{++}$ ). For comparison, we also include the spectroscopic measurement of 43 local unresolved Green Pea galaxies from Yang et al. (2017b). The distribution of  $\text{EW}_{\text{H}\alpha}$  and  $[\text{OIII}]\lambda 5007/\text{H}\alpha$  for local Green Pea galaxies is nearly identical to that of our Green Seeds, as shown by the yellow star representing their median measurement in the panel. This result reveals that Green Seeds share very similar emission line properties with the local Green Pea galaxies. This finding is not surprising, as these local Green Pea galaxies have lower metallicity and higher sSFRs (Amorín et al. 2010), making their properties more comparable to our higher-redshift observations than to other local observations.

The bottom panel of Figure 6.12 also shows that the  $\text{EW}_{\text{H}\alpha}$  in the local HII regions may depend on their physical size, with larger HII regions ( $> 200$  pc) having  $\text{EW}_{\text{H}\alpha}$  nearly 1 dex higher than smaller ones. Simulations by Tamburello et al. (2015) suggest that larger star-forming regions are more likely to form in galaxies with high gas fractions than in those with low gas fractions. Therefore, higher gas fractions could explain the higher sSFR and corresponding higher  $\text{EW}_{\text{H}\alpha}$  observed in larger HII regions.

Overall, the significant differences in the properties of local HII regions and Green Seeds, coupled with the rarity of Green Seeds in the local universe, suggest a fundamental shift in galaxy and gas properties from early epochs to the present day.

## 6.6 The origin and fate of Green Seeds

### 6.6.1 Formation of Green Seeds

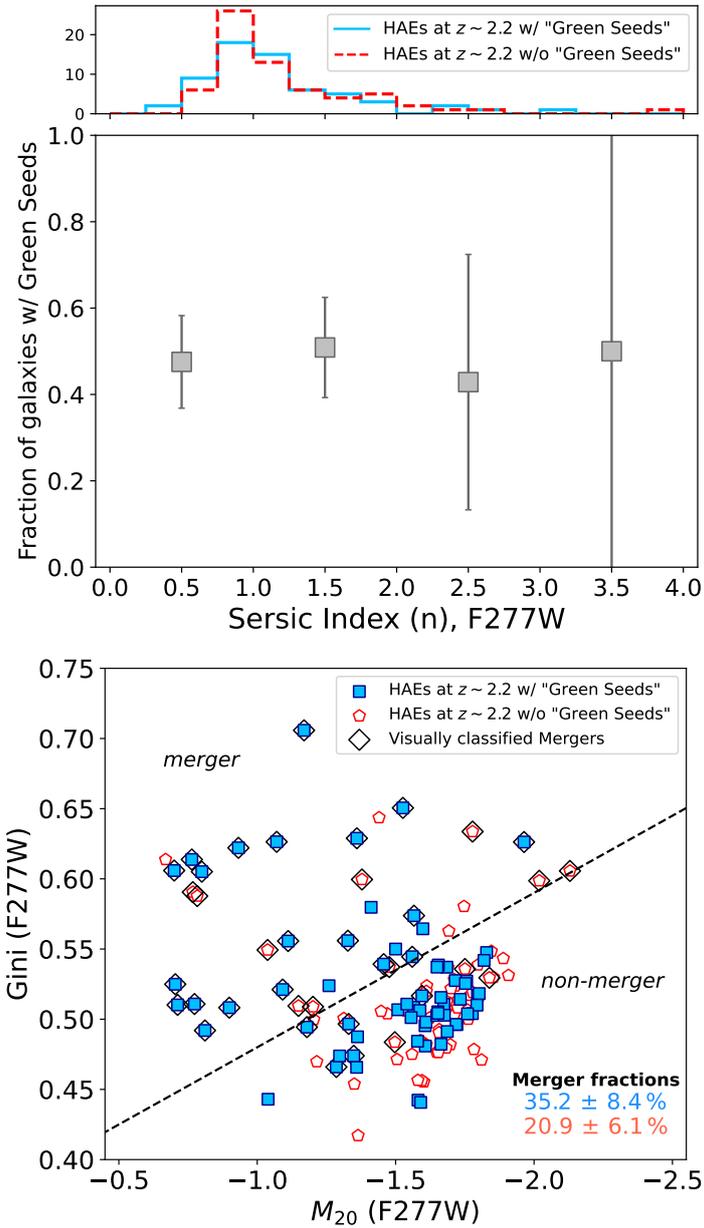
The commonly assumed framework, based on many observational and theoretical results, suggests that star-forming clumps have two possible formation mechanisms: (1) violent disk instability (VDI), regarded as “in situ” origins (Dekel et al. 2009; Mandelker et al. 2014, 2017; Dekel et al. 2022); and (2) galaxy mergers, also regarded as “ex situ” origins (Di Matteo et al. 2008; Renaud et al. 2015; Moreno et al. 2019; Sparre et al. 2022). Considering the similarities between Green Seeds and star-forming clumps discussed in Section 6.5.2, we propose that Green Seeds may share similar origins. However, previous studies of clumpy structures were mostly focusing on massive galaxies with stellar mass larger than  $10^{10} M_{\odot}$  (e.g., Elmegreen & Elmegreen 2005; Guo et al. 2012; Shibuya et al. 2016). The differing properties of their host galaxies may lead to discrepancies in the origin of Green Seeds and star-forming clumps.

In the VDI scenario, star-forming clumps are predicted to form in regions of

thick, gas-rich galaxy disks, where the high surface density of gas and young stars drives the Toomre  $Q$  parameter (Toomre 1964) below unity, leading to gravitational disc instability. Guo et al. (2015) and Shibuya et al. (2016) found that the fraction of clumpy galaxies among SFGs is consistent with the cosmological evolution of VDI from Cacciato et al. (2012), suggesting that VDI is a major mechanism for forming star-forming clumps. VDI is characterized by high gas velocity dispersion, which can only be definitively confirmed through kinematic studies using techniques such as IFU spectroscopy. In the absence of such observations, an alternative approach is to assess the structural properties of host galaxies. Since VDI requires host galaxies to have disk-like underlying components, the Sérsic index of clumpy host galaxies serves as a key diagnostic for distinguishing clump formation mechanisms. For example, galaxies in HST observations from Shibuya et al. (2015) exhibited a higher fraction of clumpy galaxies in systems with a low Sérsic index of  $n \sim 1$ , supporting the VDI scenario (Shibuya et al. 2016).

To investigate the structural parameters of the host galaxies in this study, we conduct the GALFIT profile fitting (Peng et al. 2002, 2010) on the two-dimensional surface brightness profile from the F277W cutout images of all 135 cross-matched HAEs explained in Section 6.1.2. We fit them by single Sérsic profiles centered on the galactic nuclei convolved by the F277W PSF profile from WebbPSF and optimizes the fits for  $\chi^2$  minimization. The noise images are obtained from the inverse square root of the JADES weight maps (Rieke et al. 2023), following Ono et al. (2023). The initial magnitude and effective radius used for the GALFIT profile fitting are taken from JADES catalog, but all parameters are allowed to vary during the profile fitting process. Based on the initial outputs, we exclude 22 unreliable or failed fits, mostly due to nearby counterparts in the  $3'' \times 3''$  cutout images. These outliers are re-fitted with GALFIT using a double Sérsic components mode to separate the central galaxy from its counterpart. This process provides an additional good sample, but 8 HAEs are still excluded because of their complicated components (e.g., ID:9611 in Figure 6.2). Overall, the mean Sérsic index ( $n$ ) of the HAEs with Green Seeds is slightly lower than that of the HAEs without Green Seeds by  $\sim 0.1$ , indicating a minor difference in the Sérsic profiles of these two populations. In the upper panel of Figure 6.13, we also present the relationship between the fraction of galaxies with Green Seeds and the Sérsic index. Unlike the clear correlation in Shibuya et al. (2016), where the clumpy galaxy fraction is higher in galaxies with a lower Sérsic index of  $n \sim 1$ , our sample does not show a clear correlation between these two parameters. However, as displayed in the upper histogram of the same panel, a large fraction of HAEs have a Sérsic index  $n \simeq 1$ , indicating that disk-like light profiles are generally present in the HAE sample, which could still be consistent with the VDI scenario. While considering our much smaller sample size compared to the  $\sim 190,000$  galaxies in Shibuya et al. (2016), especially the lack of galaxies with a Sérsic index  $n > 2$ , we cannot draw a definitive conclusion about VDI being a major mechanism for triggering the formation of Green Seeds. A larger sample is needed to further investigate this scenario.

During the GALFIT profile fitting, several cases of irregular morphology in HAEs with Green Seeds suggested the presence of galaxy mergers, as we find at least one or more bright counterparts in the  $3'' \times 3''$  cutout images. We determine whether



**Figure 6.13.** Upper: Dependence of the fraction of HAEs with Green Seeds on the Sérsic index ( $n$ ) measured in the F277W cutout images of all cross-matched HAEs. The error bars are given by Poisson statistics from the galaxy number counts. In the VDI scenario, a higher fraction of galaxies with Green Seeds would be expected in those with a lower Sérsic index (Shibuya et al. 2016). However, we do not observe such a clear dependence in our sample, making it hard to determine whether VDI is a major mechanism for triggering Green Seeds. Bottom: Diagram of the Gini coefficient ( $G$ ) versus the second-order moment of the brightest 20% of the galaxy's flux ( $M_{20}$ ), measured in the F277W cutout images. HAEs with Green Seeds are shown as blue squares, and those without Green Seeds are represented by red open pentagons. We also conduct a visual classification of mergers directly from the cutout images, with classified mergers indicated by black squares. The dashed line represents the threshold for ongoing mergers or non-merger, as defined by Lotz et al. (2004). We observe a potentially higher fraction of ongoing mergers among HAEs with Green Seeds.

a galaxy is an ongoing galaxy mergers by Gini coefficient ( $G$ ) and a second-order moment of a galaxy ( $M_{20}$ ), which are one of the major non-parametric methods for quantifying galaxy morphology (Lotz et al. 2004). The Gini coefficient is a statistic based on the Lorenz curve of fluxes per pixel in a galaxy and it represents the relative distribution of pixels covering the galaxy, while the other parameter  $M_{20}$  is a normalized second-order moment of pixels which is measured to be the brightest 20% flux in a galaxy. Lotz et al. (2004) defined a threshold for distinguishing mergers from non-mergers, represented by the dashed line in the bottom panel of Figure 6.13. We use the statmorph Python package (Rodriguez-Gomez et al. 2019), an affiliated package of Astropy (Astropy Collaboration et al. 2022), to calculate non-parametric morphological diagnostics on the cutout images of HAEs. Both HAEs with and without Green Seeds are analyzed, and the results are displayed in the same panel with respective symbols. Independently, we also carry out a visual classification of mergers directly from every cutout images, with visually classified mergers represented as black squares. We figure out that the non-parametric method and visual classification provide quite similar results, with fewer than 20% of visually classified mergers being outliers in the  $G - M_{20}$  diagram. For the following discussion, we use the merger fractions calculated from the non-parametric method.

We identify that the merger fraction of HAEs with Green Seeds is  $34.8 \pm 8.2\%$ , marginally higher than that of HAEs without Green Seeds of  $20.9 \pm 6.1\%$ . This finding raises the possibility of “ex situ” origins for Green Seeds. In this scenario, galaxy mergers could drive external violent processes, triggering turbulent modes in the ISM that lead to rapid fragmentation of gas and the formation of clumps (Bournaud et al. 2011; Renaud et al. 2014). Observations also support a correlation between star-forming clumps and mergers; for example, Calabrò et al. (2019) used the same  $G - M_{20}$  diagrams from HST images and found a factor of three higher clumpiness in mergers compared to the rest of the population from nearly 200 massive galaxies. Mergers and galaxy–galaxy interactions could also drive extreme emission lines, which are directly correlated with the physical background of Green Seeds. In an unresolved view, Gupta et al. (2023) found out that EELGs are mostly surrounded by massive companion galaxies and are more likely to have recently experienced strong interactions. The merger-induced bursty star formation histories produce extreme emission lines on a significantly short timescale. Notably, due to the limited sample size, the  $1\sigma$  lower limit of the merger fraction of HAEs with Green Seeds is still lower than the  $1\sigma$  upper limit of the merger fraction of HAEs without Green Seeds, indicating that we cannot make a definitive conclusion about a higher merger fraction in HAEs with Green Seeds. Furthermore, our HAEs exhibit a wide diversity on their morphology, from the smooth disk-like structure to the irregular and merger-like structure. Since these merger systems constitute less than half of the full HAE sample, it is difficult to conclude that galaxy mergers are the primary mechanism for triggering the formation of Green Seeds.

Recently, Dekel et al. (2022) simulated a wide range of clumps, from in situ clumps to ex situ clumps, and identified distinguishing features that ex situ clumps tend to populate the outer disc, while in situ clumps are expected at all radii. These simulated results generally correspond to Green Seeds through visual inspection. Although we cannot quantitatively determine which mechanism is more important

for forming Green Seeds, we suggest that both mechanisms are active in our sample.

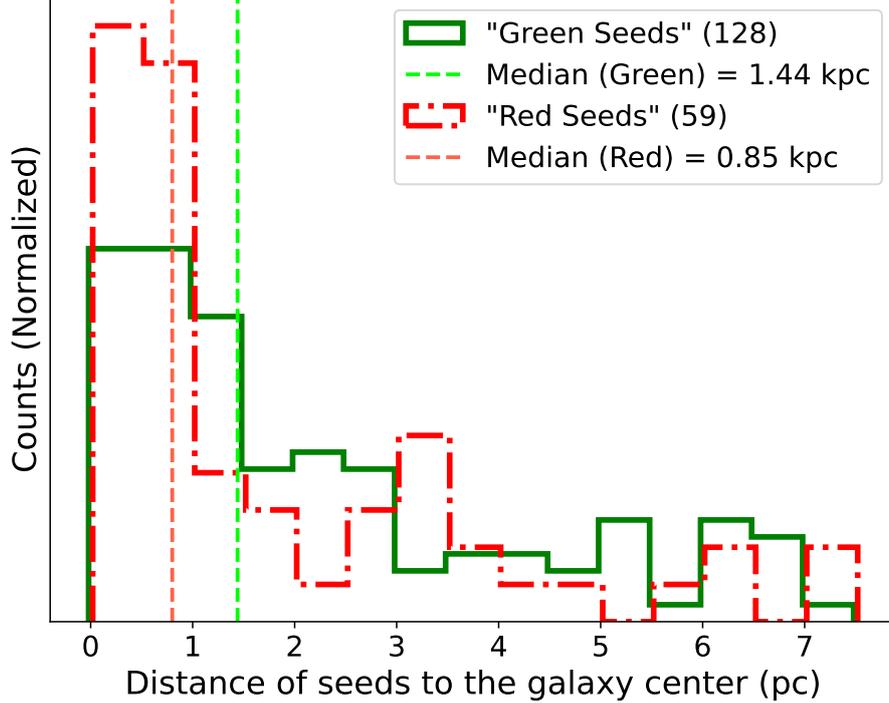
Additionally, in Figure 3.1, it is noted that a significant fraction of low-mass HAEs lie above the SFMS in the COSMOS and UDS fields, while this sample is lacking in the GOOD-South field. This enhanced SFR in host galaxies may be related to environment factors, such as the presence of galaxy clusters (Spitler et al. 2012; Shimakawa et al. 2018a; Momose et al. 2021). It remains unclear whether Green Seeds are more frequently found in clustered environments, where galaxy mergers are more common. Therefore, a larger sample of HAEs with deep JWST images in the COSMOS and UDS fields is needed to address this question.

## 6.6.2 Fate of Green Seeds

While the origins of Green Seeds may be linked to VDI and/or galaxy mergers, their ultimate fate remains a topic of debate despite various theoretical studies and numerical simulations have been done. Early simulations by Elmegreen et al. (2008) showed that clumpy structures interact strongly with disk stars, gradually lose their angular momentum, migrate toward the gravitational centers of their host galaxies, and eventually coalesce into a young bulge within 500 Myr. In Section 6.4, we have compared various properties of Green Seeds and Red Seeds. Previous studies based on HST images mainly focused on rest-frame UV-bright structures, lacking observations of these redder systems. Red Seeds are located closer to the galaxy center than Green Seeds, as shown in Figure 6.14 (a median distance of  $0.85 \pm 0.36$  kpc vs.  $1.44 \pm 0.79$  kpc). Moreover, we compare the stellar age of Green Seeds and Red Seeds from SED fitting estimation, and reveal a clear discrepancy: the median stellar age of Red Seeds ( $251 \pm 35$  Myr) is much older than that of Green Seeds ( $74 \pm 50$  Myr). These findings probably suggest that these “red” and massive clumps are likely migrating toward the gravitational center, supporting the so-called “migration” scenario.

Along with this migration scenario, Shapiro et al. (2010) proposed that these massive gas clumps may be the formation site of globular clusters (GCs; Brodie & Strader 2006). As the clumps migrate towards the centre, a small fraction of the mass is being stripped off during this process and remaining in the galaxy disc to form the metal-rich GCs. These metal-rich GCs are preferentially close to the galaxy center, typically located in the galactic thick disc (Forbes et al. 2001).

Notably, the above simulation results supporting the migration scenario are based on giant clumps with stellar mass of  $10^{8-9} M_{\odot}$ , which applies to most Red Seeds but only a few Green Seeds. In Figure 6.8, we show that Green Seeds have stellar masses ranging from  $10^{6.5} M_{\odot}$  to  $10^{8.5} M_{\odot}$ . As a result, Green Seeds with stellar mass lower than a few  $10^7 M_{\odot}$  may not be suitable for the migration scenario. Also, because the accretion of gas from the surrounding disk and gaseous outflows from stellar feedback are happening simultaneously (Bournaud et al. 2014), it is difficult for low-mass Green Seeds to rapidly accumulate their stellar masses. As suggested by Dekel et al. (2022), less massive Green Seeds are typically short-lived, in most cases  $< 200$  Myr. In this simulation, the cold gas reservoirs are expelled through stellar feedback, supernova feedback, and tidal torques due to their shallower potential wells. The short-lived nature means that low-mass Green Seeds may be disrupted



**Figure 6.14.** The distribution of distances from Green and Red Seeds to their respective galaxy centers is shown in histograms, with green and red representing each population. The galaxy center is defined as the source position in the JADES catalog, determined using the SExtractor (Bertin & Arnouts 1996; Barbary 2016). For each population, a dashed line indicates the median distance from the galaxy center. Our analysis reveals that Red Seeds are more concentrated toward the galaxy center compared to Green Seeds.

during or before migration toward galactic centers, eventually becoming part of the disk of their host galaxies.

From the continuum map and emission map in Figure 6.7 and Appendix C, we also observe several faint, low-mass Green Seeds located away from galaxy disks, which do not seem to fit the two scenarios mentioned above, likely only applicable within galaxy disks. We consider them to be an outlier populations that may follow a different evolution path. One possible assumption is that these isolated Green Seeds may be the progenitors of metal-poor GCs observed in the local universe. The metal-poor GCs are thought to be formed from young massive clusters in the halos of their host galaxies (Forbes et al. 2008; Kruijssen 2014; Forbes et al. 2018). Cosmological zoom-in simulations by Mandelker et al. (2018) presented a model in which cold filamentary accretion forms GCs. It is shown that cold streams can fragment into dense clusters not associated with the disk structure, whereas star formation occurs. The largest cluster in the simulation has a radius of  $\sim 1$  kpc and stellar mass of  $4 \times 10^6 M_{\odot}$ , corresponding to the very low-mass Green Seeds in Figure 6.8. The typical stellar masses of local GCs are in the range  $\sim 10^4 M_{\odot} - 10^6 M_{\odot}$  (Brodie & Strader 2006), while these surviving GCs are likely have undergone mass loss since their formation, requiring them to have been roughly 20 times more massive than their present-day masses in the assumed models (Kruijssen 2014). Therefore, the stellar masses of low-mass Green Seeds, which are  $\sim 10^7 M_{\odot}$ , could be reasonable

if we consider them to be proto-GCs that have experienced extensive mass loss to evolve into the local metal-poor GCs.

Another possible scenario is that these isolated low-mass Green Seeds could become typical satellite galaxies, and then evolve into ultracompact dwarf galaxies (UCDs; [Drinkwater et al. 2000](#); [Phillipps et al. 2001](#)) in the nearby Universe. The stellar masses of UCDs are ranging within  $M_* = 10^6 - 10^8 M_\odot$ , quite close to low-mass Green Seeds. While UCDs typically hold sizes less than 100 pc, recent detections of extended stellar envelopes around a number of luminous UCDs suggest that UCDs are the remnants of nucleated dwarf galaxies that have survived tidal stripping ([Liu et al. 2015, 2020](#); [Wang et al. 2023](#)). The stellar envelopes and expected tidal radius for these objects are extended to several 100 pc, which is close to typical sizes of isolated low-mass Green Seeds.

Overall, we list several possible scenarios for the evolution path of Green Seeds mainly based on their average stellar mass and size. These results are driven from the multi-band photometric data from JWST/NIRCam. However, there is a lack of critical kinematic information on Green Seeds to further support these scenarios. To address these limitations and make further progress, we anticipate that the JWST/NIRSpec IFU spectroscopy will play a crucial role.



# Chapter 7

## Conclusions

### 7.1 Summary points

In this study, we have carried out a systematic search for HAEs at  $z \sim 2.3$  in three ZFOURGE fields. The selection process for identifying HAE candidates involved examining the flux excess detected in the ZFOURGE- $K_s$  filters, which is indicative of strong  $H\alpha$  emission lines compared to the underlying stellar continuum estimated through SED fitting. To ensure the reliability of our sample, we applied a conservative selection criterion of  $3\sigma$ , resulting in the identification of 1318 HAEs. Following the same strategy, we also extract the [OIII] and [OII] emission lines of these HAEs from the ZFOURGE medium  $J/H$ -band filters.

1. We have identified more than 1300 HAEs in three ZFOURGE fields. Considering the limiting volume of the ZFOURGE survey ( $\Delta V = 6.8 \times 10^5 \text{ Mpc}^3$ ), our method has proven to be highly efficient in identifying emitters. Additionally, the derived emission line fluxes, including  $H\alpha$ , [OIII] and [OII], exhibit a high level of consistency with measurements obtained through spectroscopy (and grism) from the MOSDEF (and 3D-HST) Emission-Line Catalogs. Specifically, more than 80% of the detected fluxes show consistent values within a factor of 2. This demonstrates the reliability and accuracy of our method in determining the emission line properties.
2. The  $\text{SFR} - M_*$  relation, i.e., SFMS, derived from the  $H\alpha$  luminosity, exhibits a slope of  $0.56 \pm 0.03$  above the stellar mass completeness. When comparing our results with those from the literature, we find that the shallower slope is primarily influenced by sample selection biases. Meanwhile, we identify a subset of 401 low-mass HAEs ( $< 10^9 M_\odot$ ) that deviate from the SFMS( $H\alpha$ ) by  $\sim 0.3$  dex. We regard these low-mass HAEs as a large population of **EELGs** at Cosmic Noon, characterized by their strong  $H\alpha$  and [OIII] emission lines and high sSFRs.

We investigate the correlations between the equivalent widths of these emission lines and various galaxy and ISM properties, including stellar mass, stellar age, SFR, sSFR, and ionization state. The main findings of our analysis can be summarized as follows. These findings shed light on the relationships between emission line

equivalent widths and galaxy/ISM properties, providing insights into the nature of intense emitters.

3. The stellar mass, stellar age, sSFR, exhibit significant correlations with the [OIII] and H $\alpha$  equivalent widths. These properties serve as useful indicators for identifying intense emitters. Our method successfully extends these correlations to the lower mass domain, around  $\sim 10^8 M_\odot$ , supporting the notion that high EW<sub>[OIII]</sub> and EW<sub>H $\alpha$</sub>  are prevalent in low-mass galaxies at high redshift.
4. The [OII] equivalent widths display the weakest correlations with the aforementioned attributes. This suggests that neutral oxygen is more likely to be doubly ionized at higher redshifts. This observation aligns with the presence of a considerable number of low-mass galaxies exhibiting weak [OII] emission at high redshift.
5. Our sample reveals that the ionization-sensitive line index, O32, increases with the [OIII] equivalent widths, indicating extreme ISM properties for the most intense [OIII] emitters. In contrast, the H $\alpha$  equivalent widths show a much weaker correlation with ionization states, and the [OII] equivalent widths are largely independent of O32. This implies that optical emission lines have varying sensitivities to ISM properties.

We have compared the galaxy properties of the low-mass HAEs in our study with those of Ly $\alpha$  emitters (LAEs), which are known to have high ionization parameters and are considered as potential LyC leakages. While the low-mass HAEs exhibit milder ionization states on average, a considerable number of them still possess extreme ISM properties. We propose an “Iceberg” model to connect LAEs and low-mass HAEs, highlighting the importance of low-mass HAEs during cosmic reionization. Thus, we further investigate the ionizing photon production efficiency,  $\xi_{\text{ion}}$ , of our sample.

6. The ionizing photon production efficiency,  $\xi_{\text{ion}}$ , of the low-mass HAEs is found to be  $\log(\xi_{\text{ion}}/\text{erg}^{-1}Hz) = 25.24^{+0.10}_{-0.13}$  ( $25.35^{+0.12}_{-0.15}$ ), assuming the Calzetti (SMC) curve for UV dust correction. This result is higher by  $\sim 0.2$  dex compared to other HAEs in our study, suggesting a possible mass dependence of  $\xi_{\text{ion}}$ .
7. We observe a correlation between  $\xi_{\text{ion}}$  and both the UV slope ( $\beta_{\text{UV}}$ ) and UV absolute magnitude ( $M_{\text{UV}}$ ). Galaxies with bluest UV slopes and faintest UV luminosities exhibit an enhanced  $\xi_{\text{ion}}$  by nearly a factor of two compared to the median  $\xi_{\text{ion}}$  of our sample. Our results also confirm that  $\xi_{\text{ion}}$  is related to the equivalent widths of H $\alpha$  and [OIII] (EW<sub>[OIII]</sub> and EW<sub>H $\alpha$</sub> ). This indicates that the equivalent width of these strong optical lines can serve as a proxy for estimating  $\xi_{\text{ion}}$ .
8. By combining a comprehensive analysis of literature results, we have strengthened the evidence for the evolution of  $\xi_{\text{ion}}$  with redshift, extending our study to a significant number of low-mass galaxies at  $z \sim 2$ . Moreover, our findings suggest a potential “downsizing” relationship between  $\xi_{\text{ion}}$  and stellar mass as

we trace back in cosmic time. We utilize stellar population synthesis (SPS) models to highlight that  $\xi_{\text{ion}}$  observed in low-mass galaxies may approach the upper limit predicted by these models. This finding suggests that low-mass galaxies could be reaching the maximum efficiency of ionizing photon production according to current SPS models.

With the unprecedentedly deep and high-resolution data from JWST, we are able to resolve the rest-frame optical emission lines of these low-mass HAEs in remarkable detail. We investigate the highly resolved and deep NIRCcam imaging from the JWST Advanced Deep Extragalactic Survey (JADES; [Rieke et al. 2023](#)) on a cross-matched sample of 135 HAEs. From the convolved RGB (F115W+F150W+F277W) images, we identify a large number of kiloparsec-scale resolved [OIII]+H $\beta$  emission line regions, exhibiting prominent green colors (flux excesses in F150W) due to their high equivalent widths (EWs). We apply an EW-limited algorithm to extract these regions and designate them as “Green Seeds”.

9. Utilizing the resolved stellar continuum map and emission line map in the F150W filter, we successfully identify 187 [OIII]+H $\beta$  emission line regions with  $EW_{[\text{OIII}]+\text{H}\beta} > 200\text{\AA}$ . Based on a series of SED model grids, we develop a color selection to separate them into two populations: 128 Green Seeds from 68 HAEs, characterized by their distinctly green colors, and 59 Red Seeds. The substantial number of Green Seeds enable us to statistically study their physical properties.
10. We estimate the stellar mass ( $M_*$ ) of each Green Seed, using both color-mass-to-light ratio and SED fitting methods, which provide closely aligned estimations. The stellar masses of Green Seeds are mostly ranging from  $10^{6.5} M_\odot$  to  $10^{8.5} M_\odot$ , with a median stellar mass of  $10^{7.4} M_\odot$ . In contrast, Red Seeds are more massive than Green Seeds by  $\sim 1$  dex. Most Green Seeds also exhibit significant flux excesses in the F200W filter, driven by H $\alpha$  emission lines. We derive the SFRs of Green Seeds from their intrinsic H $\alpha$  luminosities. The  $M_*$ -SFR relation of Green Seeds has a higher normalization than that of host galaxies by  $\sim 0.6$  dex, indicating  $\sim 4\times$  more rapid stellar mass assembly in Green Seeds. Green Seeds have a median equivalent width of  $EW_{[\text{OIII}], \text{med}} = 431\text{\AA}$ , due to the intense ionizing radiation fields in Green Seeds. We identify 17 Green Seeds with extremely large rest-frame  $EW_{[\text{OIII}]} > 1000\text{\AA}$ , potentially indicating resolved LyC leakage regions in galaxies at  $z \sim 2$ .
11. We compare Green Seeds with UV star-forming clumps (from F090W) and HII regions (from F200W) in the same host galaxies. As all these spatially-resolved structures are associated with star-forming activities, we observe the co-location among them in most cases. While Green Seeds are sometimes “off-peak” relative to UV star-forming clumps, which could be explained by differences in the timescales of star formation or ionized gas outflows driven by star formation feedback. A comparison between Green Seeds and large HII regions in the local universe from the PHANGS-MUSE survey ([Groves et al. 2023](#)) reveals a much higher [OIII] $\lambda$ 5007/H $\alpha$  ratios by  $\sim 1.5$  dex in Green Seeds, but comparable  $EW_{\text{H}\alpha}$  between the two populations. This suggests that

Green Seeds may have lower metallicity and top-heavy IMFs, indicating the discrepancy in the gas and galaxy properties between the local universe and Cosmic Noon.

12. Our analysis on the morphology of host galaxies supports two different formation mechanisms proposed by theoretical works, that Green Seeds may form through violent gravitational instability (in situ) and/or galaxy mergers (ex situ). Still, we cannot definitively determine which scenario plays the dominant role in the formation of Green Seeds. Comparing with Green Seeds and Red Seeds on galaxy disks, we observe radial and stellar age variations between them. This may suggest the migration scenario, in which Green Seeds gradually migrate toward the centers of their host galaxies, evolve into Red Seeds, and eventually coalesce to build the central bulge of their host galaxies. While those low-mass ( $< 10^8 M_{\odot}$ ) Green Seeds are more likely to dissipate within the galactic disks. Additionally, we connect Green Seeds to the formation of local galactic substructures. These isolated Green Seeds could be considered progenitors of GCs or UCDs in the local universe. Our observational results of Green Seeds align well with theoretical models on stellar mass and size.

## 7.2 Future Aspects

Ongoing observations with the JWST are allowing us to explore the early universe in unprecedented detail. The data are transforming our approach to study galaxies, enhancing our understanding of their internal structures and mass assembly processes, especially in the critical low-mass regime and at rest-frame optical or even longer wavelengths. The systematic study of emission line regions conducted in this work has the potential to significantly deepen our understanding of galaxy structure and evolution at Cosmic Noon. Looking ahead, the Extremely Large Telescopes (ELTs), such as the Giant Magellan Telescope (GMT), will become the premier ground-based optical-infrared facilities from the 2030s onward. These ELTs are designed to achieve wider FOV and even higher resolutions, facilitated by advanced adaptive optics (AO) systems, surpassing the capabilities of JWST. Pilot studies using JWST data can lay a strong foundation for more extensive research with these future instruments.

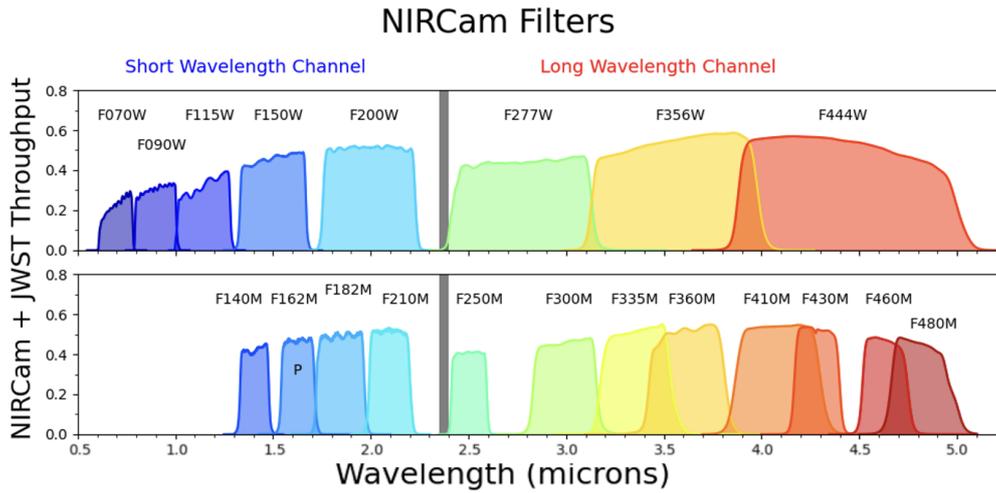
1. In this study, the selection process for identifying HAE candidates and the analysis of Green Seeds demonstrated the feasibility of isolating strong optical emission line (regions) from flux excesses in BB at  $z \sim 2$ . Ideally, by combining BB and MB filters, we can achieve more precise emission line sampling and better constrained stellar continuum. In most case for emitters, the corresponding emission lines will boost one of the MB filters, while others trace the stellar continuum. Future observations that include an extensive set of filters, from BB to NB, coupled with a large FOV, will enable the construction of a larger and more diverse emitters sample.
2. As shown in Figure 2.3, we exhibit several best-fit SED models for HAEs. Among the panel, the rightmost two black circles are IRAC 1/2 channel and

the existing IRAC data are not sufficient to accurately constrain the stellar continuum at longer wavelengths. This is primarily due to the shallower depth of the IRAC observations and the uncertainties caused by source confusion, which is a result of the relatively large PSF size. These limitations in the IRAC data may introduce inaccuracies and reduce the reliability of the SED fitting results. To overcome these limitations, JWST observation with the F277W, F444W, and F770W filters would significantly improve the accuracy of SED fitting and enable more precise measurements of the rest-frame NIR stellar continuum.

3. The higher sensitivity of JWST data enables the detection of [OIII] and [OII] emission lines more effectively. In Section 4.1, we mentioned the sSFR- $EW$  relationship is not exactly redshift-invariant for [OIII] and [OII]. Our sample predominantly consists of galaxies with  $EW_{[\text{OIII}]}$   $> 30\text{\AA}$ , while the SDSS sample reaches the detection limit at several angstroms. Due to the  $EW$  incompleteness of our sample, it is difficult to determine whether low  $EW$  galaxies would follow the same slope as our high  $EW$  sample or align with the SDSS sample region. In order to further investigate this trend, follow-up measurements with high sensitivity are indispensable.
4. To reliably evaluate the redshift evolution of  $\xi_{\text{ion}}$ , it is essential to construct large samples of individual HAEs across various redshifts. However, this remains a challenging task and is currently being undertaken in this study and [Atek et al. \(2022\)](#). The vast amount of data from JWST will be instrumental in filling this gap. In the JADES field, there are overlapping regions with a medium-band survey: JEMS, which provides MB observations at  $2\ \mu\text{m}$  (two filters) and  $4\ \mu\text{m}$  (three filters). This setup enables accurate sampling of H $\alpha$  emission lines at  $z \sim 2$  and  $z \sim 5$ . Beyond JEMS, future JWST MB observations will allow for the construction of a large number of HAEs spanning  $1 < z < 6.6$ . In Figure 7.1, we further present the available HAEs at various redshifts from JWST MB observations.
5. What are the major formation mechanisms of isolated emission line regions within galaxies? They may form due to galaxy mergers, violent disk instabilities, or cold filamentary accretion from the IGM. The environment of the host galaxies could also play a role; galaxies in overdense regions might experience more frequent mergers. Future wide-field surveys from JWST or ELTs offer a promising pathway to address these questions. By sampling a wide range of environments, from field regions to galaxy overdensities, these surveys can provide statistically significant data to disentangle the relative contributions of different formation mechanisms.
6. A potential relationship may exist between isolated emission line regions and the process of cosmic reionization. Previous studies have suggested a close connection between extreme [OIII] emission and LyC leakage, where ionizing photons escape from galaxies to reionize the surrounding intergalactic medium. The advent of ELTs, with their optical wavelength coverage and powerful sensitivity, offers a unique opportunity to directly detect LyC leakage from

Green Seeds at  $z < 6$ . Such observations will provide direct evidence of this phenomenon and further our understanding of the role these regions play in galaxy evolution and cosmic reionization.

- Low-mass galaxies at "Cosmic Noon" serve as the building blocks of more massive galaxies at later times and are therefore a key population for understanding galaxy formation and evolution. Through our selection criterion, we have identified 401 low-mass HAEs scattering above the SFMS in the upper panel of Figure 2.1. Although we conduct simulations of mock HAEs, accounting for photometric errors and fluctuations of stellar continuum in Figure 3.3, we cannot definitely confirm whether this observation reflects the intrinsic scatter of the SFMS or the existence of a low-mass sequence in the  $H\alpha$  indicator. Discovery of even lower-mass dwarf galaxies is particularly exciting. JWST Slitless Spectroscopy has proven capable of capturing emission-line galaxies, providing their rest-frame optical emission line measurements for an unbiased sample of galaxies down to stellar masses of  $\sim 10^7 M_\odot$  at  $z \sim 2$  (Malkan et al. 2024, accepted by ApJ).



$H\alpha$  emitters at various redshifts in NIRCam MB filters:

F140M: $1.03 < z < 1.25$	F162M: $1.35 < z < 1.61$	F182M: $1.63 < z < 1.99$
F210M: $2.04 < z < 2.35$	F250M: $2.68 < z < 2.95$	F300M: $3.32 < z < 3.81$
F335M: $3.86 < z < 4.39$	F360M: $4.24 < z < 4.80$	F410M: $4.89 < z < 5.55$
F430M: $5.35 < z < 5.70$	F460M: $5.88 < z < 6.23$	F480M: $6.11 < z < 6.56$

**Figure 7.1.** The throughput of the NIRCam broad-band (BB) and medium-band (MB) filters from 1 to  $5 \mu\text{m}$ , which is available from <https://jwst-docs.stsci.edu/jwst-near-infrared-camera/nircam-instrumentation/nircam-filters>. Besides, the redshift bins corresponding to  $H\alpha$  emission lines shifting into these filters are listed below.

# Appendix A

## SWIMS Imaging in this study

### A.1 SWIMS imaging pipeline: SWSRED

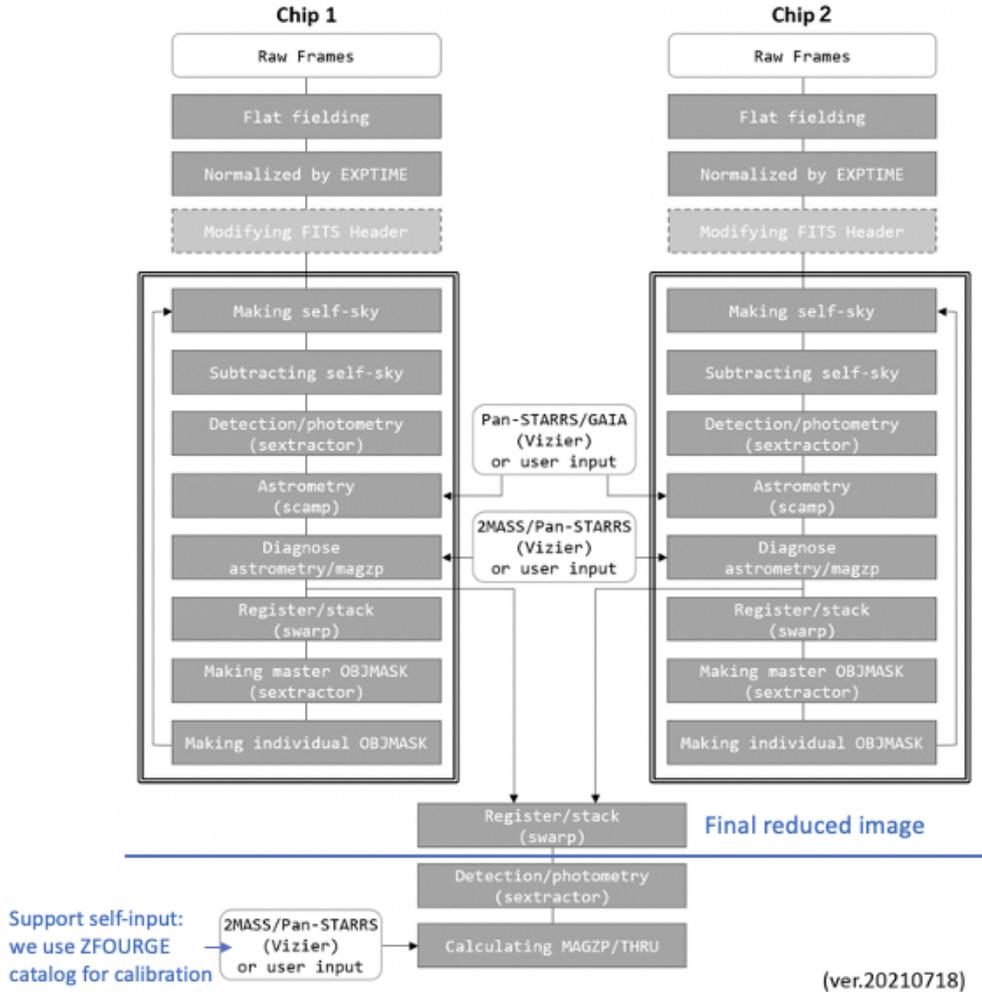
As introduced in Section 2.1.2, the SWIMS data are reduced using a custom Python pipeline, called “SWSRED”, written by Dr. Masahiro Konishi. The latest flowchart for the reduction pipeline is displayed in Figure A.1 on the [SWIMS webpage](#). While SWSRED currently lacks formal documentation, most options and parameter settings can be configured in the “swsred/reduce\_all.py” file. In this section, we follow the flowchart to outline the key steps involved in one iteration of the image reduction process.

**Flat fielding:** Dome flat frames are used for flat field. In each semester of SWIMS observations, dome flat frames of both chips in each filter would be created by IRAF, available from the same website as the flow chart. The pipeline would automatically read the semester from the raw images and choose the appropriate dome flat frame.

**Making self-sky:** The sky background is computed by averaging a user-defined number of images taken before and after a certain exposure except the frame on the same dithering position. The number of frames used to create a self sky frame is defined by `n_sky` in `reduce_all.py`. Also the combine type and other detailed setting can be adjusted in `make_sky.py`.

**Detection/Photometry:** The detection threshold for sources used for astrometry is defined by `detect_thresh_wcs` in `reduce_all.py`. The default setting is 3.0. While in some cases, for example, the number of bright stars is very rare in the field, or the seeing is not good enough, this value have to be set to a lower value. Otherwise, if the number of detected objects is very small, errors would happen during the astrometry.

**Astrometry:** The brightest (faintest) source selected for astrometry is defined by `mag_min_wcs` (`mag_max_wcs`) in `reduce_all.py`. The default setting of `mag_max_wcs` is 18.5. While, same as above, for the sky-field with very little bright stars, we may relax this setting to larger value (recommended value: 20). The reference catalog for astrometry can be selected from Pan-STARRS1 ([Chambers et al. 2016](#)), GAIA-DR2 ([Gaia Collaboration et al. 2018](#)) and 2MASS ([Skrutskie et al. 2006](#)). Notably, the pipeline encounters most of the errors during this process and the reduction pipeline would stop here.



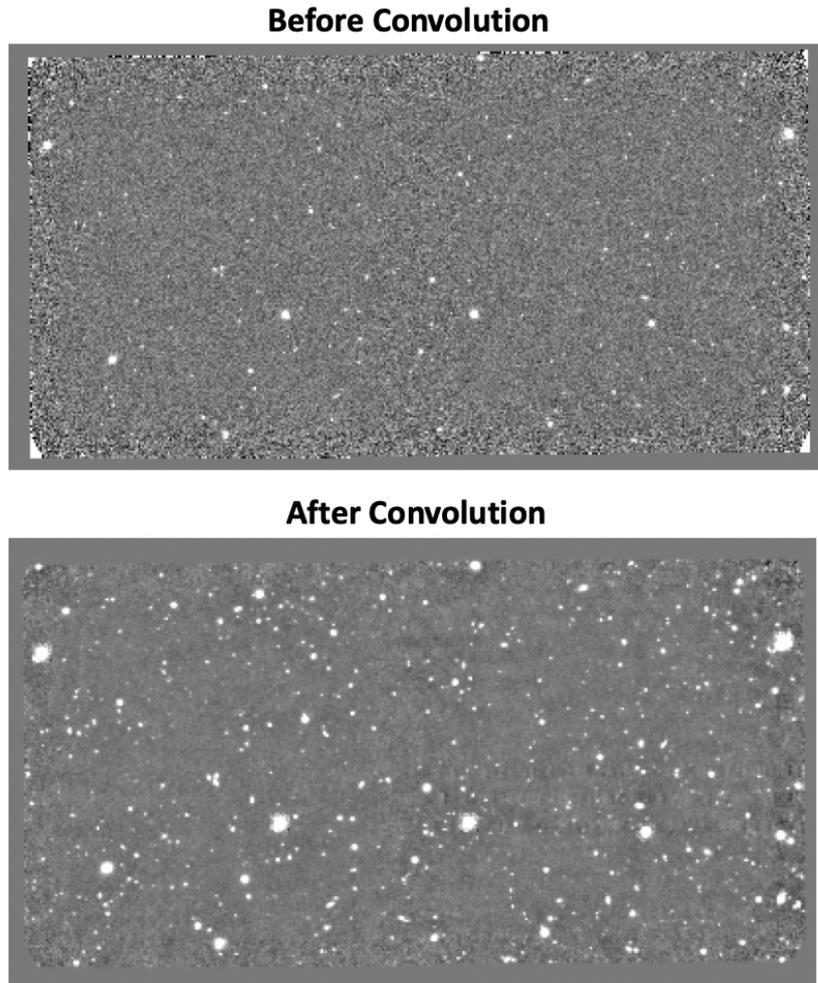
**Figure A.1.** The flow chart of SWSRED from [http://www.ioa.s.u-tokyo.ac.jp/TAO/swims/?Data\\_Reduction/Imaging\\_Data\\_Reduction](http://www.ioa.s.u-tokyo.ac.jp/TAO/swims/?Data_Reduction/Imaging_Data_Reduction).

**Register/stack:** After the correction of astrometry, all the frames in the set are registered and SWARP is ran to stack these frames together into one stacked frame. The default `COMBINE_TYPE` is `CLIPPED`. This option is not introduced in the User’s guide of SWARP (Bertin 2010). In short, this `COMBINE_TYPE` exclude pixel values that off from the mean by a value greater then a factor times of the standard deviation (Gruen et al. 2014). While, we can still choose other options which are listed in the User’s guide of SWARP.

**Making master/individual OBJMASK:** In the last part of the iteration, object masks are created for each frame in order to obtain a better self-sky in the next iteration.

## A.2 PSF matching on SWIMS reduced image

To merge our SWIMS sources into the ZFOURGE catalog, we adopt the identical PSF-matching method used in ZFOURGE. After reprojecting the  $K_1/K_2$  image to



**Figure A.2.** The SWIMS  $K_2$  image before and after convolution following the process introduced in Appendix A.2.

match the ZFOURGE pixel scale, we first select unsaturated stars with high S/N in the reprojected images and extract them as postage stamps of  $10''.65 \times 10''.65$ . Following the procedure described by [Straatman et al. \(2016\)](#), we carefully select a homogeneous sample of stars to derive the final median PSF.

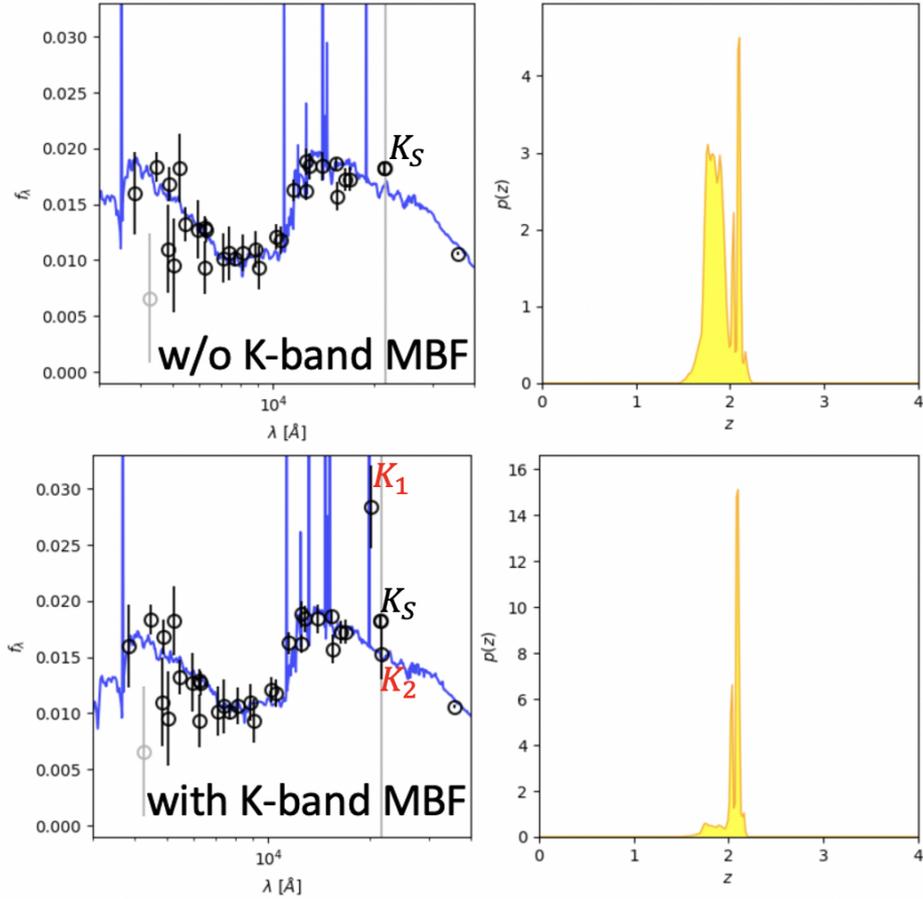
All the individual PSF were convolved into a target PSF same as the Moffat profile ([Moffat 1969](#)) with  $\text{FWHM} = 0''.9$  and  $\beta = 2.5$ . A convolution kernel was generated for each PSF during this process. Finally, the original images were convolved with their respective kernels to match the target Moffat PSF. The final convolved  $K_2$  image is shown in [Figure A.2](#).

### A.3 Better constrained Photometric redshifts

[Figure A.3](#) is an example that SWIMS  $K_1/K_2$  fluxes is included in the ZFOUREG catalog. We rerun the EAZY code and update  $z_{\text{phot}}$  (`z_peak`) with the new outputs. As this object does not show strong color excess in medium  $J/H$  band, it shows a bimodal distribution in  $p(z)$  when EAZY was run without the SWIMS  $K_1/K_2$  data.

After SWIMS data being included, the  $K_1$  filter shows strong color excess, likely to be boosted by  $H\alpha$  emission line, while the  $K_2$  filter may indicate the level of stellar continuum. With the additional SWIMS data, EAZY no longer gives a bimodal distribution but a very constrained distribution of  $p(z)$ .

We again obtain the  $\sigma_z = |z_{\text{phot}} - z_{\text{spec}}|/(1 + z_{\text{spec}})$  as ZFOURGE was done. After including our SWIMS MB data,  $\sigma_z$  drop from 0.03 to 0.02 in the ZFOURGE-COSMOS field. Statistically, the overall correspondence is ever better after adding SWIMS  $K_1/K_2$  data, as indicated by the smaller scatter in the difference between photometric and spectroscopic redshifts.



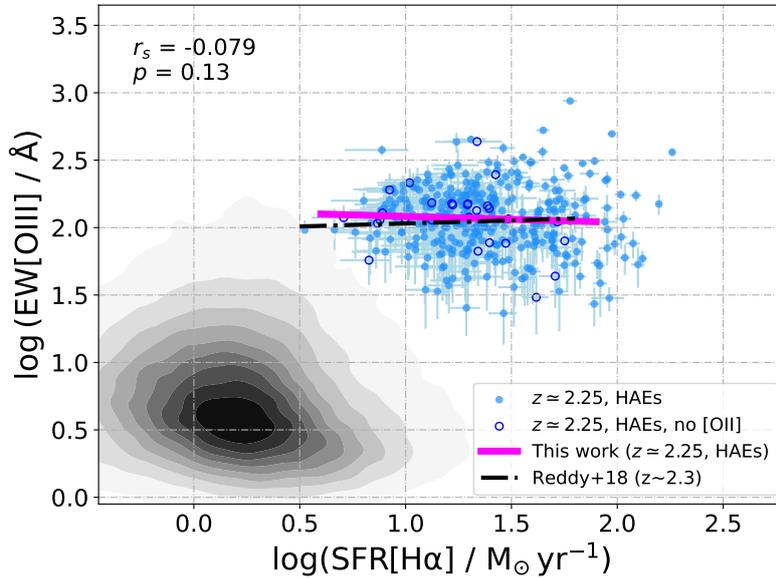
**Figure A.3.** An example galaxy fitted by EAZY templates to obtain  $z_{\text{phot}}$ . Open circles represent flux of the galaxy in every filter. In the upper panel we show the EAZY result without the SWIMS  $K_1/K_2$  filters. Blue line represents the best-fit template spectrum. In the bottom panel we show the result after including the SWIMS  $K_1/K_2$  bands. In both cases, we exhibit the redshift probability distribution functions  $p(z)$  in the right panels. A much better constraint is obtained after including the SWIMS  $K_1/K_2$  data.

# Appendix B

## SFR vs. $EW_{[\text{OIII}]}$ of massive HAEs

In this analysis, we have focused on galaxies in our sample with stellar masses greater than  $10^{9.5}M_{\odot}$ . This specific subset of galaxies allows us to obtain the best-fit result with a slope of  $-0.06$  in Figure B.1. We acknowledge that there is a discrepancy between this specific result and the bottom-left panel of Figure 4.1, which could potentially be attributed to sample selection biases.

It is important to consider that sample selection biases can introduce uncertainties and limitations to our findings. Factors such as the selection criteria, observational constraints, and data quality can influence the observed trends and correlations. Therefore, it is crucial to carefully assess and account for any potential biases when interpreting and comparing different results within the analysis.



**Figure B.1.** Relationship between the  $EW_{[\text{OIII}]}$  and SFR of the HAEs with stellar masses larger than  $10^{9.5}M_{\odot}$ . Outlines as in Figure 4.1.



# Appendix C

## The resolved structures of HAEs from JWST observations

### C.1 The Cutout images of remaining cross-matched HAEs

In Chapter 6, a total of 135 HAEs is cross-matched to the JADES catalog (Rieke et al. 2023). In Section 6.2, we also illustrate how we extract Green Seeds from the combination of resolved stellar continuum map and emission line map, and we present the cutout color images of a series of HAEs built combining the F115W, F150W, and F277W filters which contains Green Seeds in Figure 6.2. Still, from the whole sample of host galaxies, we have a number of 67 HAEs that do not contains Green Seeds and we present the RGB images of these rest HAEs in this Figure C.1. As is shown here, these sample do not reveal a prominent resolved compact green structures.

Considering the integrated galaxy properties of the parent HAEs, we found about two thirds of the HAEs exhibit strong [OIII] emission lines, which is slightly higher than the fraction of the HAEs which contains Green Seeds. One reason could be the errors on the photometric redshifts of the HAEs from the ZFOURGE survey (Straatman et al. 2016). Although ZFOURGE is able to constrain such errors at  $\sim 3\%$  at  $z > 2$ , some [OIII] emission lines may still drop out from the F150W filter of JWST as outliers.

We also mention one special object: ID:19033 in Figure C.1, which actually contains strong [OIII] emission lines. This object is also classified as an AGN in Cowley et al. (2016). Interestingly, the [OIII] emission on both sides exhibits a symmetric filament structure, indicating a warm ionized outflow originating from the galaxy centre and extending over 10 kpc. This structure is similar to the large ionized Hanny’s Voorwerp features in nearby AGN, which have strong associations with mergers and radio jets/outflows. Furthermore, the JWST imaging uncovers an ongoing merger system in the galaxy’s nuclear region, revealing two closely separated point sources by an distance of 4 pixels, equating to a physical distance of  $\sim 1.5$  kpc. These two new findings outstand the uniqueness of this AGN, especially at high redshift. However, a detailed analysis on this object is beyond the scope of this

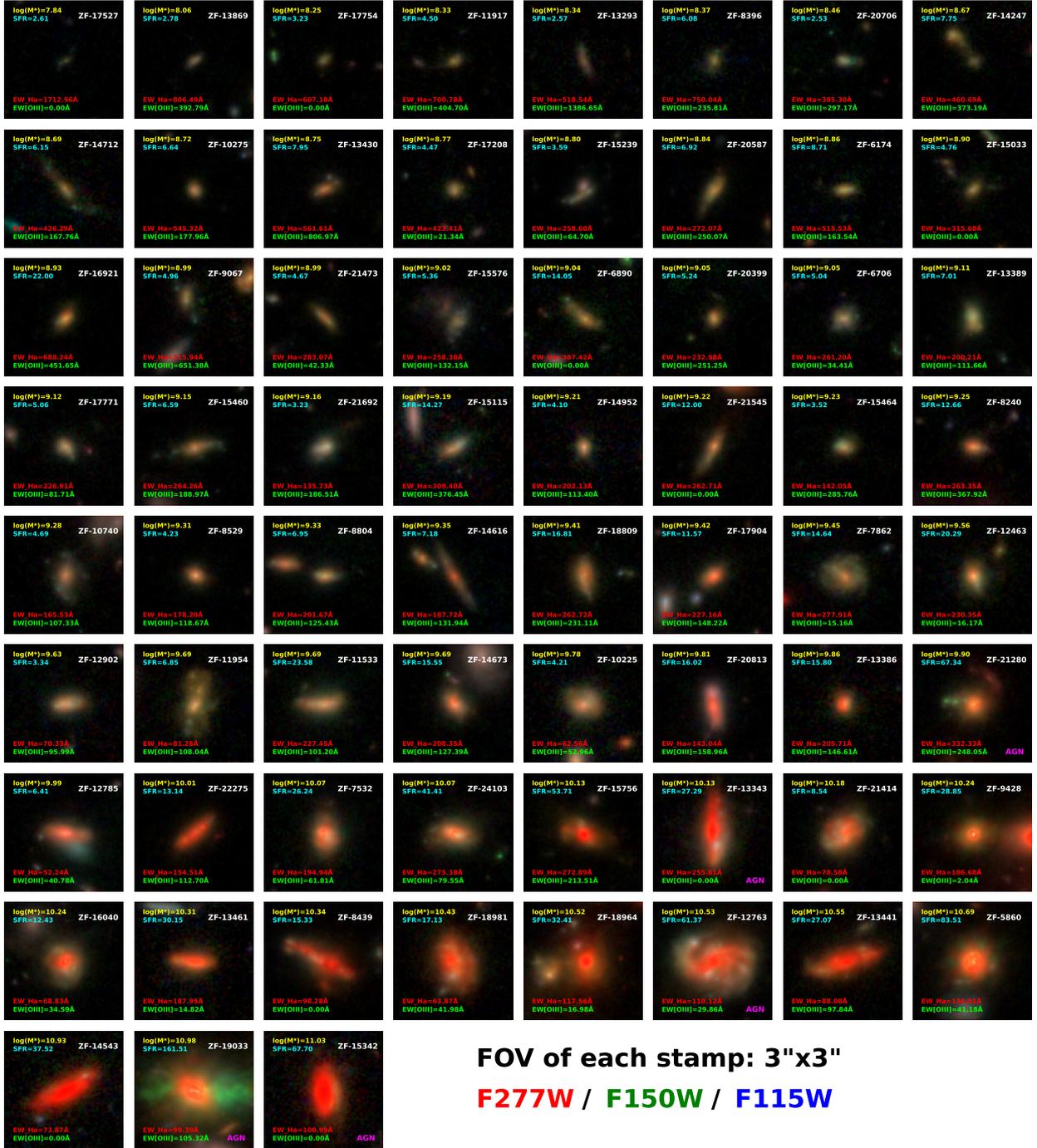
study.

## **C.2 The Continuum and Emission-line Map for the remaining HAEs with Green Seeds**

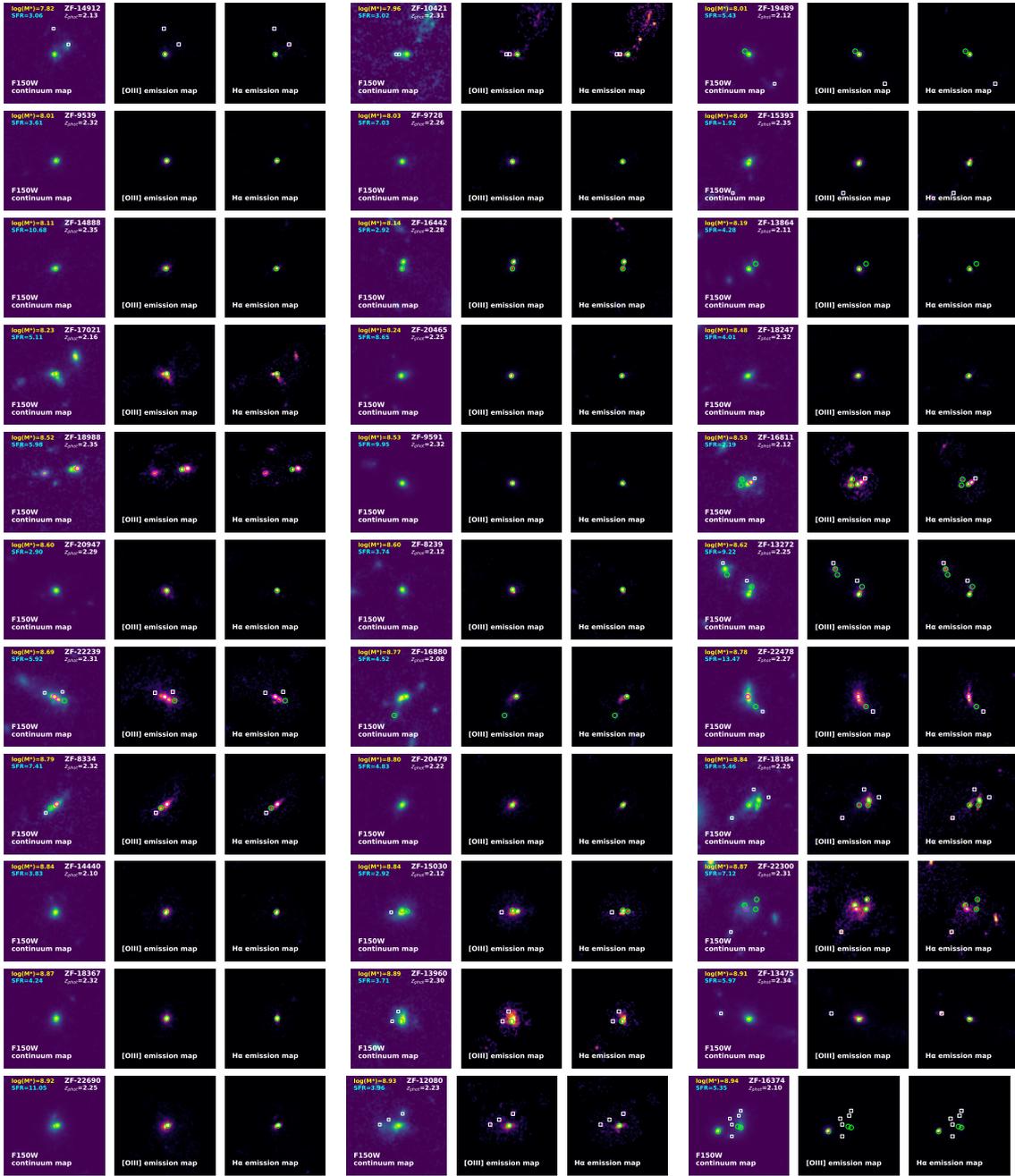
In Figure 6.7, we present the stellar continuum map at  $1.5\ \mu m$ , the [OIII]+H $\beta$  emission line map, and the H $\alpha$  emission line map of just six samples. Here we provide the extended versions in Figure C.2 and Figure C.3, which include the results of the remaining 62 HAEs (all 68 HAEs) that contain Green Seeds in this study.

## **C.3 The remaining UV clumps in combination with Green Seeds**

Similarly, in Figure 6.11, we illustrate the UV star-forming clumps of only six HAEs identical to the sample in Figure 6.7. Here we provide the extended versions in Figure C.4, which include the clump-detection of all 68 HAEs that contain Green Seeds in this work.



**Figure C.1.** JWST/NIRCam cutout images of the other 67 HAEs with size of  $3'' \times 3''$ , which are not exhibited in Figure 3. Outlines as in Figure 3. We do not extract any Green Seeds from these HAEs. This sub-sample of HAEs have a higher average stellar mass than that contains Green Seeds. The AGN identified in Cowley et al. (2016) are labeled at the bottom-right corner. The HAEs in Figure 3 contain no AGN. Notably, an AGN (ZF-19033) may exhibit a pair of strong warm ionized outflow traced by the [OIII] emissions, and another AGN (ZF-21280) also exhibit a possible ionized outflow. These two AGNs are beyond this work, thus we do not further discuss their galaxy properties.



**Figure C.2.** The stellar continuum map at  $1.5 \mu\text{m}$ , the  $[\text{OIII}]+\text{H}\beta$  emission line map, and the  $\text{H}\alpha$  emission line map of the rest HAEs which contain Green Seeds. Outlines as in Figure 6.7. Those sample detected with only one Green Seeds may also be considered as “Green Pea” galaxies at  $z \sim 2$ .

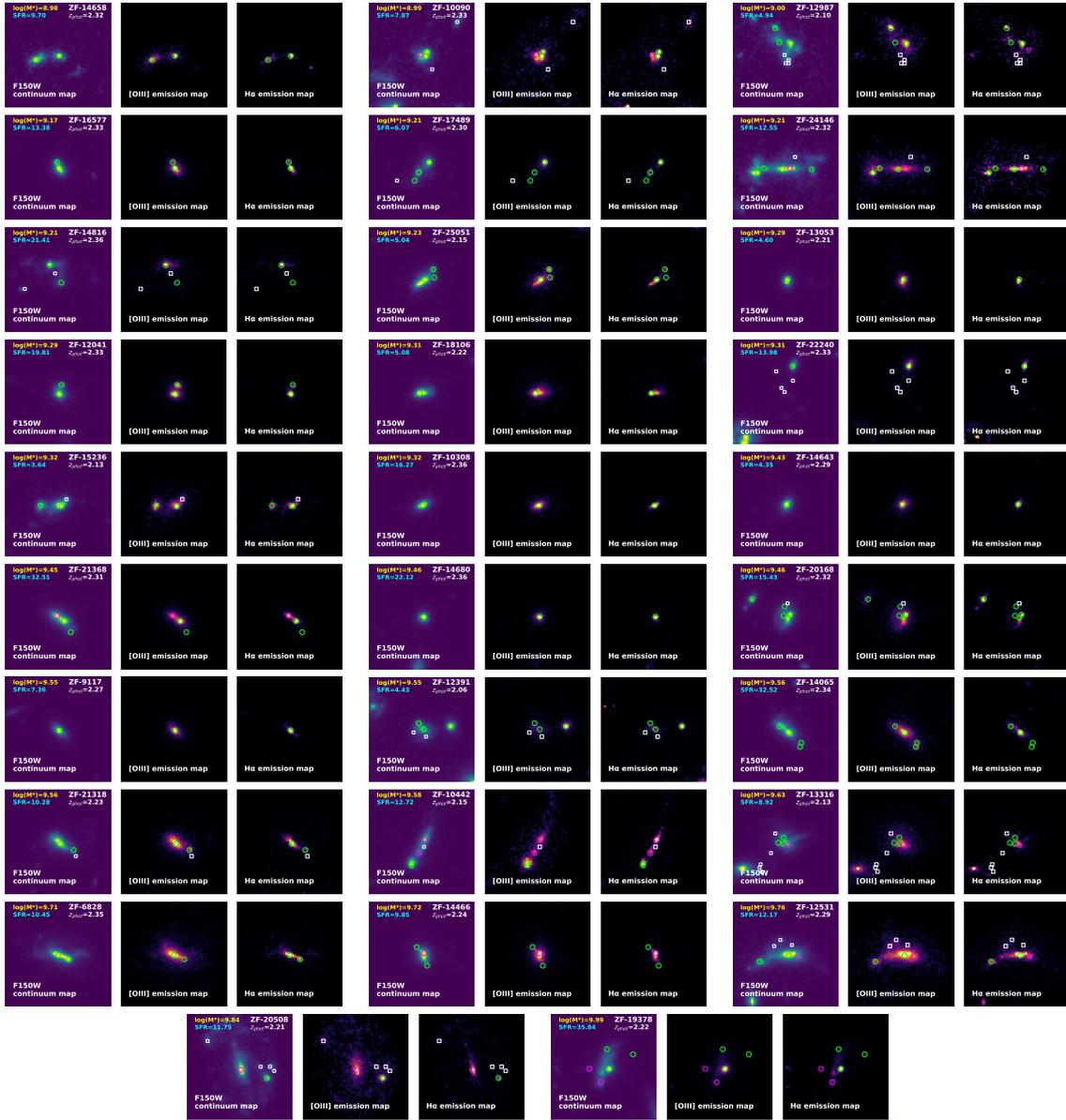
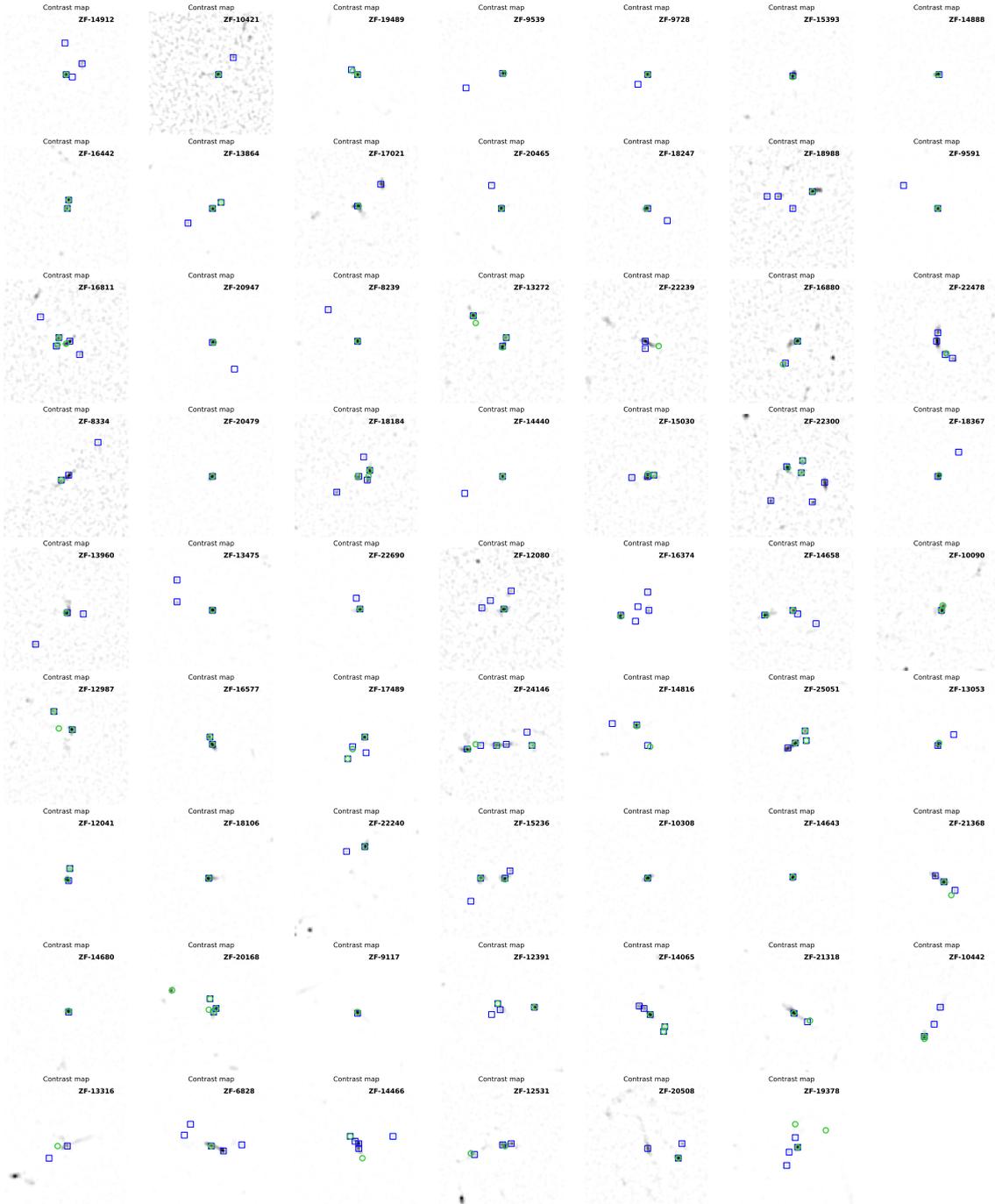


Figure C.3. Continue of Figure C.2. Outlines as in Figure 6.7.



**Figure C.4.** The rest contrast maps and the UV star-forming clumps of HAEs which contain Green Seeds. Outlines as in Figure 6.11.

# Bibliography

- Amorín, R. O., Pérez-Montero, E., & Vílchez, J. M. 2010, *ApJ*, 715, L128, doi: [10.1088/2041-8205/715/2/L128](https://doi.org/10.1088/2041-8205/715/2/L128)
- Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, *ApJ*, 935, 167, doi: [10.3847/1538-4357/ac7c74](https://doi.org/10.3847/1538-4357/ac7c74)
- Atek, H., Furtak, L. J., Oesch, P., et al. 2022, *MNRAS*, 511, 4464, doi: [10.1093/mnras/stac360](https://doi.org/10.1093/mnras/stac360)
- Atek, H., Malkan, M., McCarthy, P., et al. 2010, *ApJ*, 723, 104, doi: [10.1088/0004-637X/723/1/104](https://doi.org/10.1088/0004-637X/723/1/104)
- Atek, H., Siana, B., Scarlata, C., et al. 2011, *ApJ*, 743, 121, doi: [10.1088/0004-637X/743/2/121](https://doi.org/10.1088/0004-637X/743/2/121)
- Atek, H., Richard, J., Jauzac, M., et al. 2015, *ApJ*, 814, 69, doi: [10.1088/0004-637X/814/1/69](https://doi.org/10.1088/0004-637X/814/1/69)
- Atek, H., Labbé, I., Furtak, L. J., et al. 2024, *Nature*, 626, 975, doi: [10.1038/s41586-024-07043-6](https://doi.org/10.1038/s41586-024-07043-6)
- Bagley, M. B., Finkelstein, S. L., Koekemoer, A. M., et al. 2023, *ApJ*, 946, L12, doi: [10.3847/2041-8213/acbb08](https://doi.org/10.3847/2041-8213/acbb08)
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5, doi: [10.1086/130766](https://doi.org/10.1086/130766)
- Barbary, K. 2016, *Journal of Open Source Software*, 1, 58, doi: [10.21105/joss.00058](https://doi.org/10.21105/joss.00058)
- Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., et al. 2006, *AJ*, 132, 1729, doi: [10.1086/507302](https://doi.org/10.1086/507302)
- Bell, E. F., & de Jong, R. S. 2001, *ApJ*, 550, 212, doi: [10.1086/319728](https://doi.org/10.1086/319728)
- Bertin, E. 2010, *SWarp: Resampling and Co-adding FITS Images Together*. <http://ascl.net/1010.068>
- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393, doi: [10.1051/aas:1996164](https://doi.org/10.1051/aas:1996164)
- Bian, F., Kewley, L. J., & Dopita, M. A. 2018, *ApJ*, 859, 175, doi: [10.3847/1538-4357/aabd74](https://doi.org/10.3847/1538-4357/aabd74)
- Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, *Nature*, 311, 517, doi: [10.1038/311517a0](https://doi.org/10.1038/311517a0)
- Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, *A&A*, 622, A103, doi: [10.1051/0004-6361/201834156](https://doi.org/10.1051/0004-6361/201834156)
- Bournaud, F., Chapon, D., Teyssier, R., et al. 2011, *ApJ*, 730, 4, doi: [10.1088/0004-637X/730/1/4](https://doi.org/10.1088/0004-637X/730/1/4)
- Bournaud, F., Perret, V., Renaud, F., et al. 2014, *ApJ*, 780, 57, doi: [10.1088/0004-637X/780/1/57](https://doi.org/10.1088/0004-637X/780/1/57)
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, *ApJ*, 811, 140, doi: [10.1088/0004-637X/811/2/140](https://doi.org/10.1088/0004-637X/811/2/140)

- Bouwens, R. J., Smit, R., Labbé, I., et al. 2016, *ApJ*, 831, 176, doi: [10.3847/0004-637X/831/2/176](https://doi.org/10.3847/0004-637X/831/2/176)
- Boyett, K., Bunker, A. J., Curtis-Lake, E., et al. 2024, arXiv e-prints, arXiv:2401.16934, doi: [10.48550/arXiv.2401.16934](https://doi.org/10.48550/arXiv.2401.16934)
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503, doi: [10.1086/591786](https://doi.org/10.1086/591786)
- Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, *ApJS*, 200, 13, doi: [10.1088/0067-0049/200/2/13](https://doi.org/10.1088/0067-0049/200/2/13)
- Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, *MNRAS*, 351, 1151, doi: [10.1111/j.1365-2966.2004.07881.x](https://doi.org/10.1111/j.1365-2966.2004.07881.x)
- Brodie, J. P., & Strader, J. 2006, *ARA&A*, 44, 193, doi: [10.1146/annurev.astro.44.051905.092441](https://doi.org/10.1146/annurev.astro.44.051905.092441)
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000, doi: [10.1046/j.1365-8711.2003.06897.x](https://doi.org/10.1046/j.1365-8711.2003.06897.x)
- Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, *MNRAS*, 360, 1413, doi: [10.1111/j.1365-2966.2005.09131.x](https://doi.org/10.1111/j.1365-2966.2005.09131.x)
- Burgarella, D., Heinis, S., Magdis, G., et al. 2011, *ApJ*, 734, L12, doi: [10.1088/2041-8205/734/1/L12](https://doi.org/10.1088/2041-8205/734/1/L12)
- Cacciato, M., Dekel, A., & Genel, S. 2012, *MNRAS*, 421, 818, doi: [10.1111/j.1365-2966.2011.20359.x](https://doi.org/10.1111/j.1365-2966.2011.20359.x)
- Calabrò, A., Daddi, E., Fensch, J., et al. 2019, *A&A*, 632, A98, doi: [10.1051/0004-6361/201935778](https://doi.org/10.1051/0004-6361/201935778)
- Calzetti, D. 2013, *Star Formation Rate Indicators*, ed. J. Falcón-Barroso & J. H. Knapen, 419
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682, doi: [10.1086/308692](https://doi.org/10.1086/308692)
- Calzetti, D., Harris, J., Gallagher, John S., I., et al. 2004, *AJ*, 127, 1405, doi: [10.1086/382095](https://doi.org/10.1086/382095)
- Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, *ApJ*, 429, 582, doi: [10.1086/174346](https://doi.org/10.1086/174346)
- Caputi, K. I., Lagache, G., Yan, L., et al. 2007, *ApJ*, 660, 97, doi: [10.1086/512667](https://doi.org/10.1086/512667)
- Cardamone, C., Schawinski, K., Sarzi, M., et al. 2009, *MNRAS*, 399, 1191, doi: [10.1111/j.1365-2966.2009.15383.x](https://doi.org/10.1111/j.1365-2966.2009.15383.x)
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245, doi: [10.1086/167900](https://doi.org/10.1086/167900)
- Carilli, C. L., & Walter, F. 2013, *ARA&A*, 51, 105, doi: [10.1146/annurev-astro-082812-140953](https://doi.org/10.1146/annurev-astro-082812-140953)
- Carniani, S., Hainline, K., D'Eugenio, F., et al. 2024, *Nature*, 633, 318, doi: [10.1038/s41586-024-07860-9](https://doi.org/10.1038/s41586-024-07860-9)
- Carroll, S. M. 2001, *Living Reviews in Relativity*, 4, 1, doi: [10.12942/lrr-2001-1](https://doi.org/10.12942/lrr-2001-1)
- Castellano, M., Fontana, A., Treu, T., et al. 2022, *ApJ*, 938, L15, doi: [10.3847/2041-8213/ac94d0](https://doi.org/10.3847/2041-8213/ac94d0)
- Castellano, M., Belfiori, D., Pentericci, L., et al. 2023, *A&A*, 675, A121, doi: [10.1051/0004-6361/202346069](https://doi.org/10.1051/0004-6361/202346069)
- Chabrier, G. 2003, *PASP*, 115, 763, doi: [10.1086/376392](https://doi.org/10.1086/376392)
- Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv e-prints, arXiv:1612.05560. <https://arxiv.org/abs/1612.05560>

- Charlot, S., & Fall, S. M. 2000, *ApJ*, 539, 718, doi: [10.1086/309250](https://doi.org/10.1086/309250)
- Chevallard, J., Charlot, S., Senchyna, P., et al. 2018, *MNRAS*, 479, 3264, doi: [10.1093/mnras/sty1461](https://doi.org/10.1093/mnras/sty1461)
- Claeysens, A., Adamo, A., Richard, J., et al. 2023, *MNRAS*, 520, 2180, doi: [10.1093/mnras/stac3791](https://doi.org/10.1093/mnras/stac3791)
- Cohn, J. H., Leja, J., Tran, K.-V. H., et al. 2018, *ApJ*, 869, 141, doi: [10.3847/1538-4357/aaed3d](https://doi.org/10.3847/1538-4357/aaed3d)
- Combes, F. 2018, *A&A Rev.*, 26, 5, doi: [10.1007/s00159-018-0110-4](https://doi.org/10.1007/s00159-018-0110-4)
- Conselice, C. J. 2003, *ApJS*, 147, 1, doi: [10.1086/375001](https://doi.org/10.1086/375001)
- . 2014, *ARA&A*, 52, 291, doi: [10.1146/annurev-astro-081913-040037](https://doi.org/10.1146/annurev-astro-081913-040037)
- Conselice, C. J., Blackburne, J. A., & Papovich, C. 2005, *ApJ*, 620, 564, doi: [10.1086/426102](https://doi.org/10.1086/426102)
- Conselice, C. J., Rajgor, S., & Myers, R. 2008, *MNRAS*, 386, 909, doi: [10.1111/j.1365-2966.2008.13069.x](https://doi.org/10.1111/j.1365-2966.2008.13069.x)
- Conselice, C. J., Grogin, N. A., Jogee, S., et al. 2004, *ApJ*, 600, L139, doi: [10.1086/378556](https://doi.org/10.1086/378556)
- Cowie, L. L., & Hu, E. M. 1998, *AJ*, 115, 1319, doi: [10.1086/300309](https://doi.org/10.1086/300309)
- Cowley, M. J., Spitler, L. R., Tran, K.-V. H., et al. 2016, *MNRAS*, 457, 629, doi: [10.1093/mnras/stv2992](https://doi.org/10.1093/mnras/stv2992)
- Cresci, G., Vanzi, L., Telles, E., et al. 2017, *A&A*, 604, A101, doi: [10.1051/0004-6361/201730876](https://doi.org/10.1051/0004-6361/201730876)
- Curti, M., Cresci, G., Mannucci, F., et al. 2017, *MNRAS*, 465, 1384, doi: [10.1093/mnras/stw2766](https://doi.org/10.1093/mnras/stw2766)
- Daddi, E., Dickinson, M., Morrison, G., et al. 2007, *ApJ*, 670, 156, doi: [10.1086/521818](https://doi.org/10.1086/521818)
- Davies, R. L., Förster Schreiber, N. M., Genzel, R., et al. 2021, *ApJ*, 909, 78, doi: [10.3847/1538-4357/abd551](https://doi.org/10.3847/1538-4357/abd551)
- de Lapparent, V., Geller, M. J., & Huchra, J. P. 1986, *ApJ*, 302, L1, doi: [10.1086/184625](https://doi.org/10.1086/184625)
- Dekel, A., Mandelker, N., Bournaud, F., et al. 2022, *MNRAS*, 511, 316, doi: [10.1093/mnras/stab3810](https://doi.org/10.1093/mnras/stab3810)
- Dekel, A., Sari, R., & Ceverino, D. 2009, *ApJ*, 703, 785, doi: [10.1088/0004-637X/703/1/785](https://doi.org/10.1088/0004-637X/703/1/785)
- Di Matteo, P., Bournaud, F., Martig, M., et al. 2008, *A&A*, 492, 31, doi: [10.1051/0004-6361:200809480](https://doi.org/10.1051/0004-6361:200809480)
- Domínguez, A., Siana, B., Brooks, A. M., et al. 2015, *MNRAS*, 451, 839, doi: [10.1093/mnras/stv1001](https://doi.org/10.1093/mnras/stv1001)
- Dopita, M. A., & Evans, I. N. 1986, *ApJ*, 307, 431, doi: [10.1086/164432](https://doi.org/10.1086/164432)
- Drinkwater, M. J., Jones, J. B., Gregg, M. D., & Phillipps, S. 2000, *PASA*, 17, 227, doi: [10.1071/AS00034](https://doi.org/10.1071/AS00034)
- Eikenberry, S., Bandyopadhyay, R., Bennett, J. G., et al. 2012, in *SPIE*, Vol. 8446, , 84460I, doi: [10.1117/12.925679](https://doi.org/10.1117/12.925679)
- Eisenstein, D. J., Willott, C., Alberts, S., et al. 2023a, arXiv e-prints, arXiv:2306.02465, doi: [10.48550/arXiv.2306.02465](https://doi.org/10.48550/arXiv.2306.02465)

- Eisenstein, D. J., Johnson, B. D., Robertson, B., et al. 2023b, arXiv e-prints, arXiv:2310.12340, doi: [10.48550/arXiv.2310.12340](https://doi.org/10.48550/arXiv.2310.12340)
- Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, *A&A*, 533, A119, doi: [10.1051/0004-6361/201117239](https://doi.org/10.1051/0004-6361/201117239)
- Eldridge, J. J., & Stanway, E. R. 2022, *ARA&A*, 60, 455, doi: [10.1146/annurev-astro-052920-100646](https://doi.org/10.1146/annurev-astro-052920-100646)
- Elmegreen, B. G., Bournaud, F., & Elmegreen, D. M. 2008, *ApJ*, 688, 67, doi: [10.1086/592190](https://doi.org/10.1086/592190)
- Elmegreen, B. G., & Elmegreen, D. M. 2005, *ApJ*, 627, 632, doi: [10.1086/430514](https://doi.org/10.1086/430514)
- Elmegreen, D. M., Elmegreen, B. G., Ravindranath, S., & Coe, D. A. 2007, *ApJ*, 658, 763, doi: [10.1086/511667](https://doi.org/10.1086/511667)
- Emami, N., Siana, B., Alavi, A., et al. 2020, *ApJ*, 895, 116, doi: [10.3847/1538-4357/ab8f97](https://doi.org/10.3847/1538-4357/ab8f97)
- Emami, N., Siana, B., Weisz, D. R., et al. 2019, *ApJ*, 881, 71, doi: [10.3847/1538-4357/ab211a](https://doi.org/10.3847/1538-4357/ab211a)
- Emsellem, E., Schinnerer, E., Santoro, F., et al. 2022, *A&A*, 659, A191, doi: [10.1051/0004-6361/202141727](https://doi.org/10.1051/0004-6361/202141727)
- Erb, D. K. 2008, *ApJ*, 674, 151, doi: [10.1086/524727](https://doi.org/10.1086/524727)
- Erb, D. K., Pettini, M., Shapley, A. E., et al. 2010, *ApJ*, 719, 1168, doi: [10.1088/0004-637X/719/2/1168](https://doi.org/10.1088/0004-637X/719/2/1168)
- Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, *ApJ*, 644, 813, doi: [10.1086/503623](https://doi.org/10.1086/503623)
- Faisst, A. L., Capak, P. L., Emami, N., Tacchella, S., & Larson, K. L. 2019, *ApJ*, 884, 133, doi: [10.3847/1538-4357/ab425b](https://doi.org/10.3847/1538-4357/ab425b)
- Faisst, A. L., Masters, D., Wang, Y., et al. 2018, *ApJ*, 855, 132, doi: [10.3847/1538-4357/aab1fc](https://doi.org/10.3847/1538-4357/aab1fc)
- Faisst, A. L., Brinch, M., Casey, C. M., et al. 2024, arXiv e-prints, arXiv:2405.09619, doi: [10.48550/arXiv.2405.09619](https://doi.org/10.48550/arXiv.2405.09619)
- Fang, J. J., Faber, S. M., Koo, D. C., et al. 2018, *ApJ*, 858, 100, doi: [10.3847/1538-4357/aabcba](https://doi.org/10.3847/1538-4357/aabcba)
- Feltre, A., Charlot, S., & Gutkin, J. 2016, *MNRAS*, 456, 3354, doi: [10.1093/mnras/stv2794](https://doi.org/10.1093/mnras/stv2794)
- Ferland, G. J., Korista, K. T., Verner, D. A., et al. 1998, *PASP*, 110, 761, doi: [10.1086/316190](https://doi.org/10.1086/316190)
- Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, *RMxAA*, 49, 137. <https://arxiv.org/abs/1302.4485>
- Ferreira, L., Adams, N., Conselice, C. J., et al. 2022, *ApJ*, 938, L2, doi: [10.3847/2041-8213/ac947c](https://doi.org/10.3847/2041-8213/ac947c)
- Ferreira, L., Conselice, C. J., Sazonova, E., et al. 2023, *ApJ*, 955, 94, doi: [10.3847/1538-4357/acec76](https://doi.org/10.3847/1538-4357/acec76)
- Ferrière, K. M. 2001, *Reviews of Modern Physics*, 73, 1031, doi: [10.1103/RevModPhys.73.1031](https://doi.org/10.1103/RevModPhys.73.1031)
- Finkelstein, S. L., Papovich, C., Dickinson, M., et al. 2013, *Nature*, 502, 524, doi: [10.1038/nature12657](https://doi.org/10.1038/nature12657)
- Finkelstein, S. L., Ryan, Russell E., J., Papovich, C., et al. 2015, *ApJ*, 810, 71, doi: [10.1088/0004-637X/810/1/71](https://doi.org/10.1088/0004-637X/810/1/71)

- Finkelstein, S. L., D'Aloisio, A., Paardekooper, J.-P., et al. 2019, *ApJ*, 879, 36, doi: [10.3847/1538-4357/ab1ea8](https://doi.org/10.3847/1538-4357/ab1ea8)
- Forbes, D. A., Brodie, J. P., & Larsen, S. S. 2001, *ApJ*, 556, L83, doi: [10.1086/323006](https://doi.org/10.1086/323006)
- Forbes, D. A., Lasky, P., Graham, A. W., & Spitler, L. 2008, *MNRAS*, 389, 1924, doi: [10.1111/j.1365-2966.2008.13739.x](https://doi.org/10.1111/j.1365-2966.2008.13739.x)
- Forbes, D. A., Bastian, N., Gieles, M., et al. 2018, *Proceedings of the Royal Society of London Series A*, 474, 20170616, doi: [10.1098/rspa.2017.0616](https://doi.org/10.1098/rspa.2017.0616)
- Forrest, B., Tran, K.-V. H., Broussard, A., et al. 2017, *ApJ*, 838, L12, doi: [10.3847/2041-8213/aa653b](https://doi.org/10.3847/2041-8213/aa653b)
- . 2018, *ApJ*, 863, 131, doi: [10.3847/1538-4357/aad232](https://doi.org/10.3847/1538-4357/aad232)
- Förster Schreiber, N. M., Genzel, R., Lutz, D., Kunze, D., & Sternberg, A. 2001, *ApJ*, 552, 544, doi: [10.1086/320546](https://doi.org/10.1086/320546)
- Förster Schreiber, N. M., Shapley, A. E., Erb, D. K., et al. 2011a, *ApJ*, 731, 65, doi: [10.1088/0004-637X/731/1/65](https://doi.org/10.1088/0004-637X/731/1/65)
- Förster Schreiber, N. M., Genzel, R., Bouché, N., et al. 2009, *ApJ*, 706, 1364, doi: [10.1088/0004-637X/706/2/1364](https://doi.org/10.1088/0004-637X/706/2/1364)
- Förster Schreiber, N. M., Shapley, A. E., Genzel, R., et al. 2011b, *ApJ*, 739, 45, doi: [10.1088/0004-637X/739/1/45](https://doi.org/10.1088/0004-637X/739/1/45)
- Förster Schreiber, N. M., Übler, H., Davies, R. L., et al. 2019, *ApJ*, 875, 21, doi: [10.3847/1538-4357/ab0ca2](https://doi.org/10.3847/1538-4357/ab0ca2)
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, 616, A1, doi: [10.1051/0004-6361/201833051](https://doi.org/10.1051/0004-6361/201833051)
- Garland, C. A., Pisano, D. J., Mac Low, M. M., et al. 2015, *ApJ*, 807, 134, doi: [10.1088/0004-637X/807/2/134](https://doi.org/10.1088/0004-637X/807/2/134)
- Geach, J. E., Smail, I., Best, P. N., et al. 2008, *MNRAS*, 388, 1473, doi: [10.1111/j.1365-2966.2008.13481.x](https://doi.org/10.1111/j.1365-2966.2008.13481.x)
- Geach, J. E., Smail, I., Moran, S. M., et al. 2011, *ApJ*, 730, L19, doi: [10.1088/2041-8205/730/2/L19](https://doi.org/10.1088/2041-8205/730/2/L19)
- Geller, M. J., & Huchra, J. P. 1989, *Science*, 246, 897, doi: [10.1126/science.246.4932.897](https://doi.org/10.1126/science.246.4932.897)
- Giacconi, R., Zirm, A., Wang, J., et al. 2002, *ApJS*, 139, 369, doi: [10.1086/338927](https://doi.org/10.1086/338927)
- Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJ*, 600, L93, doi: [10.1086/379232](https://doi.org/10.1086/379232)
- Gil de Paz, A., Madore, B. F., & Pevunova, O. 2003, *ApJS*, 147, 29, doi: [10.1086/374737](https://doi.org/10.1086/374737)
- Giménez-Arteaga, C., Oesch, P. A., Brammer, G. B., et al. 2023, *ApJ*, 948, 126, doi: [10.3847/1538-4357/acc5ea](https://doi.org/10.3847/1538-4357/acc5ea)
- Giménez-Arteaga, C., Fujimoto, S., Valentino, F., et al. 2024, arXiv e-prints, arXiv:2402.17875, doi: [10.48550/arXiv.2402.17875](https://doi.org/10.48550/arXiv.2402.17875)
- Grazian, A., Giallongo, E., Gerbasi, R., et al. 2016, *A&A*, 585, A48, doi: [10.1051/0004-6361/201526396](https://doi.org/10.1051/0004-6361/201526396)

- Grazian, A., Giallongo, E., Paris, D., et al. 2017, *A&A*, 602, A18, doi: [10.1051/0004-6361/201730447](https://doi.org/10.1051/0004-6361/201730447)
- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, 197, 35, doi: [10.1088/0067-0049/197/2/35](https://doi.org/10.1088/0067-0049/197/2/35)
- Groves, B., Kreckel, K., Santoro, F., et al. 2023, *MNRAS*, 520, 4902, doi: [10.1093/mnras/stad114](https://doi.org/10.1093/mnras/stad114)
- Gruen, D., Seitz, S., & Bernstein, G. M. 2014, *PASP*, 126, 158, doi: [10.1086/675080](https://doi.org/10.1086/675080)
- Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, *MNRAS*, 432, 23, doi: [10.1093/mnras/stt308](https://doi.org/10.1093/mnras/stt308)
- Guo, Y., Giavalisco, M., Ferguson, H. C., Cassata, P., & Koekemoer, A. M. 2012, *ApJ*, 757, 120, doi: [10.1088/0004-637X/757/2/120](https://doi.org/10.1088/0004-637X/757/2/120)
- Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *ApJS*, 207, 24, doi: [10.1088/0067-0049/207/2/24](https://doi.org/10.1088/0067-0049/207/2/24)
- Guo, Y., Ferguson, H. C., Bell, E. F., et al. 2015, *ApJ*, 800, 39, doi: [10.1088/0004-637X/800/1/39](https://doi.org/10.1088/0004-637X/800/1/39)
- Guo, Y., Rafelski, M., Bell, E. F., et al. 2018, *ApJ*, 853, 108, doi: [10.3847/1538-4357/aaa018](https://doi.org/10.3847/1538-4357/aaa018)
- Gupta, A., Jaiswar, R., Rodriguez-Gomez, V., et al. 2023, *ApJ*, 957, L35, doi: [10.3847/2041-8213/ad0788](https://doi.org/10.3847/2041-8213/ad0788)
- Gutkin, J., Charlot, S., & Bruzual, G. 2016, *MNRAS*, 462, 1757, doi: [10.1093/mnras/stw1716](https://doi.org/10.1093/mnras/stw1716)
- Hagen, A., Zeimann, G. R., Behrens, C., et al. 2016, *ApJ*, 817, 79, doi: [10.3847/0004-637X/817/1/79](https://doi.org/10.3847/0004-637X/817/1/79)
- Harikane, Y., Ouchi, M., Oguri, M., et al. 2023, *ApJS*, 265, 5, doi: [10.3847/1538-4365/acaaa9](https://doi.org/10.3847/1538-4365/acaaa9)
- Hayashi, M., Kodama, T., Tanaka, I., et al. 2016, *ApJL*, 826, L28, doi: [10.3847/2041-8205/826/2/L28](https://doi.org/10.3847/2041-8205/826/2/L28)
- Hayes, M., Schaerer, D., & Östlin, G. 2010a, *A&A*, 509, L5, doi: [10.1051/0004-6361/200913217](https://doi.org/10.1051/0004-6361/200913217)
- Hayes, M., Östlin, G., Schaerer, D., et al. 2010b, *Nature*, 464, 562, doi: [10.1038/nature08881](https://doi.org/10.1038/nature08881)
- Heckman, T. M., Alexandroff, R. M., Borthakur, S., Overzier, R., & Leitherer, C. 2015, *ApJ*, 809, 147, doi: [10.1088/0004-637X/809/2/147](https://doi.org/10.1088/0004-637X/809/2/147)
- Heckman, T. M., & Thompson, T. A. 2017, arXiv e-prints, arXiv:1701.09062, doi: [10.48550/arXiv.1701.09062](https://doi.org/10.48550/arXiv.1701.09062)
- Hopkins, P. F., Quataert, E., & Murray, N. 2012, *MNRAS*, 421, 3522, doi: [10.1111/j.1365-2966.2012.20593.x](https://doi.org/10.1111/j.1365-2966.2012.20593.x)
- Hubble, E. P. 1936, *Realm of the Nebulae*
- Huchra, J., Davis, M., Latham, D., & Tonry, J. 1983, *ApJS*, 52, 89, doi: [10.1086/190860](https://doi.org/10.1086/190860)
- Inoue, A. K. 2011, *MNRAS*, 415, 2920, doi: [10.1111/j.1365-2966.2011.18906.x](https://doi.org/10.1111/j.1365-2966.2011.18906.x)
- Inoue, A. K., Iwata, I., & Deharveng, J.-M. 2006, *MNRAS*, 371, L1, doi: [10.1111/j.1745-3933.2006.00195.x](https://doi.org/10.1111/j.1745-3933.2006.00195.x)
- Izotov, Y. I., Guseva, N. G., & Thuan, T. X. 2011, *ApJ*, 728, 161, doi: [10.1088/0004-637X/728/2/161](https://doi.org/10.1088/0004-637X/728/2/161)

- Izotov, Y. I., Stasińska, G., Meynet, G., Guseva, N. G., & Thuan, T. X. 2006, *A&A*, 448, 955, doi: [10.1051/0004-6361:20053763](https://doi.org/10.1051/0004-6361:20053763)
- Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2021, *MNRAS*, 503, 1734, doi: [10.1093/mnras/stab612](https://doi.org/10.1093/mnras/stab612)
- Jacobs, C., Glazebrook, K., Calabrò, A., et al. 2023, *ApJ*, 948, L13, doi: [10.3847/2041-8213/accd6d](https://doi.org/10.3847/2041-8213/accd6d)
- Janowiecki, S., Salzer, J. J., van Zee, L., Rosenberg, J. L., & Skillman, E. 2017, *ApJ*, 836, 128, doi: [10.3847/1538-4357/836/1/128](https://doi.org/10.3847/1538-4357/836/1/128)
- Jaskot, A. E., Dowd, T., Oey, M. S., Scarlata, C., & McKinney, J. 2019, *ApJ*, 885, 96, doi: [10.3847/1538-4357/ab3d3b](https://doi.org/10.3847/1538-4357/ab3d3b)
- Ji, Z., Williams, C. C., Tacchella, S., et al. 2024, *ApJ*, 974, 135, doi: [10.3847/1538-4357/ad6e7f](https://doi.org/10.3847/1538-4357/ad6e7f)
- Kaasinen, M., Bian, F., Groves, B., Kewley, L. J., & Gupta, A. 2017, *MNRAS*, 465, 3220, doi: [10.1093/mnras/stw2827](https://doi.org/10.1093/mnras/stw2827)
- Kaasinen, M., Kewley, L., Bian, F., et al. 2018, *MNRAS*, 477, 5568, doi: [10.1093/mnras/sty1012](https://doi.org/10.1093/mnras/sty1012)
- Kalita, B. S., Silverman, J. D., Daddi, E., et al. 2024, *ApJ*, 960, 25, doi: [10.3847/1538-4357/acfee4](https://doi.org/10.3847/1538-4357/acfee4)
- Kartaltepe, J. S., Rose, C., Vanderhoof, B. N., et al. 2023, *ApJ*, 946, L15, doi: [10.3847/2041-8213/acad01](https://doi.org/10.3847/2041-8213/acad01)
- Kashino, D., Lilly, S. J., Matthee, J., et al. 2023, *ApJ*, 950, 66, doi: [10.3847/1538-4357/acc588](https://doi.org/10.3847/1538-4357/acc588)
- Kashino, D., Silverman, J. D., Rodighiero, G., et al. 2013, *ApJL*, 777, L8, doi: [10.1088/2041-8205/777/1/L8](https://doi.org/10.1088/2041-8205/777/1/L8)
- Kashino, D., Silverman, J. D., Sanders, D., et al. 2019, *ApJS*, 241, 10, doi: [10.3847/1538-4365/ab06c4](https://doi.org/10.3847/1538-4365/ab06c4)
- Katz, N., Hernquist, L., & Weinberg, D. H. 1992, *ApJ*, 399, L109, doi: [10.1086/186619](https://doi.org/10.1086/186619)
- Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, *MNRAS*, 303, 188, doi: [10.1046/j.1365-8711.1999.02202.x](https://doi.org/10.1046/j.1365-8711.1999.02202.x)
- Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, *MNRAS*, 346, 1055, doi: [10.1111/j.1365-2966.2003.07154.x](https://doi.org/10.1111/j.1365-2966.2003.07154.x)
- Kennicutt, Robert C., J. 1998, *ARA&A*, 36, 189, doi: [10.1146/annurev.astro.36.1.189](https://doi.org/10.1146/annurev.astro.36.1.189)
- Kennicutt, R. C., & Evans, N. J. 2012, *ARA&A*, 50, 531, doi: [10.1146/annurev-astro-081811-125610](https://doi.org/10.1146/annurev-astro-081811-125610)
- Kewley, L. J., & Dopita, M. A. 2002, *ApJS*, 142, 35, doi: [10.1086/341326](https://doi.org/10.1086/341326)
- Kewley, L. J., Dopita, M. A., Leitherer, C., et al. 2013, *ApJ*, 774, 100, doi: [10.1088/0004-637X/774/2/100](https://doi.org/10.1088/0004-637X/774/2/100)
- Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, *ApJ*, 556, 121, doi: [10.1086/321545](https://doi.org/10.1086/321545)
- Kewley, L. J., Nicholls, D. C., & Sutherland, R. S. 2019, *ARA&A*, 57, 511, doi: [10.1146/annurev-astro-081817-051832](https://doi.org/10.1146/annurev-astro-081817-051832)

- Kewley, L. J., Zahid, H. J., Geller, M. J., et al. 2015, *ApJL*, 812, L20, doi: [10.1088/2041-8205/812/2/L20](https://doi.org/10.1088/2041-8205/812/2/L20)
- Khostovan, A. A., Sobral, D., Mobasher, B., et al. 2015, *MNRAS*, 452, 3948, doi: [10.1093/mnras/stv1474](https://doi.org/10.1093/mnras/stv1474)
- . 2016, *MNRAS*, 463, 2363, doi: [10.1093/mnras/stw2174](https://doi.org/10.1093/mnras/stw2174)
- Kimura, M., Maihara, T., Iwamuro, F., et al. 2010, *PASJ*, 62, 1135, doi: [10.1093/pasj/62.5.1135](https://doi.org/10.1093/pasj/62.5.1135)
- Kobulnicky, H. A., & Kewley, L. J. 2004, *ApJ*, 617, 240, doi: [10.1086/425299](https://doi.org/10.1086/425299)
- Kodama, T., Hayashi, M., Koyama, Y., et al. 2013, in *IAU Symposium*, Vol. 295, *The Intriguing Life of Massive Galaxies*, ed. D. Thomas, A. Pasquali, & I. Ferreras, 74–77, doi: [10.1017/S1743921313004353](https://doi.org/10.1017/S1743921313004353)
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, 197, 36, doi: [10.1088/0067-0049/197/2/36](https://doi.org/10.1088/0067-0049/197/2/36)
- Konishi, M., Motohara, K., Takahashi, H., et al. 2012, in *SPIE*, Vol. 8446, , 84467P, doi: [10.1117/12.924259](https://doi.org/10.1117/12.924259)
- Konishi, M., Motohara, K., Takahashi, H., et al. 2020, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 11447, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 114475H, doi: [10.1117/12.2560422](https://doi.org/10.1117/12.2560422)
- Kouroumpatzakis, K., Zezas, A., Kyritsis, E., Salim, S., & Svoboda, J. 2023, *A&A*, 673, A16, doi: [10.1051/0004-6361/202346054](https://doi.org/10.1051/0004-6361/202346054)
- Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, *ApJ*, 700, 221, doi: [10.1088/0004-637X/700/1/221](https://doi.org/10.1088/0004-637X/700/1/221)
- Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, *ApJS*, 218, 15, doi: [10.1088/0067-0049/218/2/15](https://doi.org/10.1088/0067-0049/218/2/15)
- Kruijssen, J. M. D. 2014, *Classical and Quantum Gravity*, 31, 244006, doi: [10.1088/0264-9381/31/24/244006](https://doi.org/10.1088/0264-9381/31/24/244006)
- Kunth, D., & Östlin, G. 2000, *A&A Rev.*, 10, 1, doi: [10.1007/s001590000005](https://doi.org/10.1007/s001590000005)
- Labbé, I., van Dokkum, P., Nelson, E., et al. 2023, *Nature*, 616, 266, doi: [10.1038/s41586-023-05786-2](https://doi.org/10.1038/s41586-023-05786-2)
- Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, *ApJS*, 224, 24, doi: [10.3847/0067-0049/224/2/24](https://doi.org/10.3847/0067-0049/224/2/24)
- Lam, D., Bouwens, R. J., Labbé, I., et al. 2019, *A&A*, 627, A164, doi: [10.1051/0004-6361/201935227](https://doi.org/10.1051/0004-6361/201935227)
- Lasker, B. M., Lattanzi, M. G., McLean, B. J., et al. 2008, *AJ*, 136, 735, doi: [10.1088/0004-6256/136/2/735](https://doi.org/10.1088/0004-6256/136/2/735)
- Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, *MNRAS*, 379, 1599, doi: [10.1111/j.1365-2966.2007.12040.x](https://doi.org/10.1111/j.1365-2966.2007.12040.x)
- Leitherer, C., & Heckman, T. M. 1995, *ApJS*, 96, 9, doi: [10.1086/192112](https://doi.org/10.1086/192112)
- Lequeux, J., Peimbert, M., Rayo, J. F., Serrano, A., & Torres-Peimbert, S. 1979, *A&A*, 500, 145
- Lin, X., Cai, Z., Zou, S., et al. 2023, *ApJ*, 944, L59, doi: [10.3847/2041-8213/acalc4](https://doi.org/10.3847/2041-8213/acalc4)

- Liu, C., Peng, E. W., Côté, P., et al. 2015, *ApJ*, 812, 34, doi: [10.1088/0004-637X/812/1/34](https://doi.org/10.1088/0004-637X/812/1/34)
- Liu, C., Côté, P., Peng, E. W., et al. 2020, *ApJS*, 250, 17, doi: [10.3847/1538-4365/abad91](https://doi.org/10.3847/1538-4365/abad91)
- Liu, X., Shapley, A. E., Coil, A. L., Brinchmann, J., & Ma, C.-P. 2008, *ApJ*, 678, 758, doi: [10.1086/529030](https://doi.org/10.1086/529030)
- Livermore, R. C., Jones, T. A., Richard, J., et al. 2015, *MNRAS*, 450, 1812, doi: [10.1093/mnras/stv686](https://doi.org/10.1093/mnras/stv686)
- Lotz, J. M., Primack, J., & Madau, P. 2004, *AJ*, 128, 163, doi: [10.1086/421849](https://doi.org/10.1086/421849)
- Madau, P., & Dickinson, M. 2014, *ARA&A*, 52, 415, doi: [10.1146/annurev-astro-081811-125615](https://doi.org/10.1146/annurev-astro-081811-125615)
- Madau, P., & Haardt, F. 2015, *ApJ*, 813, L8, doi: [10.1088/2041-8205/813/1/L8](https://doi.org/10.1088/2041-8205/813/1/L8)
- Maihara, T., Iwamuro, F., Tanabe, H., et al. 2001, *PASJ*, 53, 25, doi: [10.1093/pasj/53.1.25](https://doi.org/10.1093/pasj/53.1.25)
- Maller, A. H., Katz, N., Kereš, D., Davé, R., & Weinberg, D. H. 2006, *ApJ*, 647, 763, doi: [10.1086/503319](https://doi.org/10.1086/503319)
- Mandelker, N., Dekel, A., Ceverino, D., et al. 2017, *MNRAS*, 464, 635, doi: [10.1093/mnras/stw2358](https://doi.org/10.1093/mnras/stw2358)
- . 2014, *MNRAS*, 443, 3675, doi: [10.1093/mnras/stu1340](https://doi.org/10.1093/mnras/stu1340)
- Mandelker, N., van Dokkum, P. G., Brodie, J. P., van den Bosch, F. C., & Ceverino, D. 2018, *ApJ*, 861, 148, doi: [10.3847/1538-4357/aaca98](https://doi.org/10.3847/1538-4357/aaca98)
- Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., & Gnerucci, A. 2010, *MNRAS*, 408, 2115, doi: [10.1111/j.1365-2966.2010.17291.x](https://doi.org/10.1111/j.1365-2966.2010.17291.x)
- Marchi, F., Pentericci, L., Guaita, L., et al. 2017, *A&A*, 601, A73, doi: [10.1051/0004-6361/201630054](https://doi.org/10.1051/0004-6361/201630054)
- Marino, R. A., Rosales-Ortega, F. F., Sánchez, S. F., et al. 2013, *A&A*, 559, A114, doi: [10.1051/0004-6361/201321956](https://doi.org/10.1051/0004-6361/201321956)
- Marques-Chaves, R., Schaerer, D., Álvarez-Márquez, J., et al. 2022, *MNRAS*, 517, 2972, doi: [10.1093/mnras/stac2893](https://doi.org/10.1093/mnras/stac2893)
- Maschietto, F., Hatch, N. A., Venemans, B. P., et al. 2008, *MNRAS*, 389, 1223, doi: [10.1111/j.1365-2966.2008.13571.x](https://doi.org/10.1111/j.1365-2966.2008.13571.x)
- Matthee, J., Mackenzie, R., Simcoe, R. A., et al. 2022, doi: [10.48550/ARXIV.2211.08255](https://doi.org/10.48550/ARXIV.2211.08255)
- Matthee, J., Sobral, D., Best, P., et al. 2017, *MNRAS*, 465, 3637, doi: [10.1093/mnras/stw2973](https://doi.org/10.1093/mnras/stw2973)
- Matthee, J., Sobral, D., Oteo, I., et al. 2016, *MNRAS*, 458, 449, doi: [10.1093/mnras/stw322](https://doi.org/10.1093/mnras/stw322)
- McGaugh, S. S., & Schombert, J. M. 2014, *AJ*, 148, 77, doi: [10.1088/0004-6256/148/5/77](https://doi.org/10.1088/0004-6256/148/5/77)
- McLean, I. S., Steidel, C. C., Epps, H. W., et al. 2012, in *SPIE*, Vol. 8446, , 84460J, doi: [10.1117/12.924794](https://doi.org/10.1117/12.924794)
- Mehta, V., Scarlata, C., Colbert, J. W., et al. 2015, *ApJ*, 811, 141, doi: [10.1088/0004-637X/811/2/141](https://doi.org/10.1088/0004-637X/811/2/141)
- Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, *ApJ*, 521, 64, doi: [10.1086/307523](https://doi.org/10.1086/307523)

- Moffat, A. F. J. 1969, *A&A*, 3, 455
- Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, *ApJS*, 225, 27, doi: [10.3847/0067-0049/225/2/27](https://doi.org/10.3847/0067-0049/225/2/27)
- Momose, R., Shimasaku, K., Kashikawa, N., et al. 2021, *ApJ*, 909, 117, doi: [10.3847/1538-4357/abd2af](https://doi.org/10.3847/1538-4357/abd2af)
- Moreno, J., Torrey, P., Ellison, S. L., et al. 2019, *MNRAS*, 485, 1320, doi: [10.1093/mnras/stz417](https://doi.org/10.1093/mnras/stz417)
- Motohara, K., Konishi, M., Takahashi, H., et al. 2016, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9908, *Ground-based and Airborne Instrumentation for Astronomy VI*, ed. C. J. Evans, L. Simard, & H. Takami, 99083U, doi: [10.1117/12.2231386](https://doi.org/10.1117/12.2231386)
- Naidu, R. P., Forrest, B., Oesch, P. A., Tran, K.-V. H., & Holden, B. P. 2018, *MNRAS*, 478, 791, doi: [10.1093/mnras/sty961](https://doi.org/10.1093/mnras/sty961)
- Nakajima, K., Ellis, R. S., Iwata, I., et al. 2016, *ApJ*, 831, L9, doi: [10.3847/2041-8205/831/1/L9](https://doi.org/10.3847/2041-8205/831/1/L9)
- Nakajima, K., Ellis, R. S., Robertson, B. E., Tang, M., & Stark, D. P. 2020, *ApJ*, 889, 161, doi: [10.3847/1538-4357/ab6604](https://doi.org/10.3847/1538-4357/ab6604)
- Nakajima, K., Fletcher, T., Ellis, R. S., Robertson, B. E., & Iwata, I. 2018, *MNRAS*, 477, 2098, doi: [10.1093/mnras/sty750](https://doi.org/10.1093/mnras/sty750)
- Nakajima, K., & Maiolino, R. 2022, *MNRAS*, 513, 5134, doi: [10.1093/mnras/stac1242](https://doi.org/10.1093/mnras/stac1242)
- Nakajima, K., & Ouchi, M. 2014, *MNRAS*, 442, 900, doi: [10.1093/mnras/stu902](https://doi.org/10.1093/mnras/stu902)
- Nakajima, K., Ouchi, M., Isobe, Y., et al. 2023, *arXiv e-prints*, arXiv:2301.12825, doi: [10.48550/arXiv.2301.12825](https://doi.org/10.48550/arXiv.2301.12825)
- Nakajima, K., Ouchi, M., Shimasaku, K., et al. 2012, *ApJ*, 745, 12, doi: [10.1088/0004-637X/745/1/12](https://doi.org/10.1088/0004-637X/745/1/12)
- Nanayakkara, T., Glazebrook, K., Kacprzak, G. G., et al. 2016, *ApJ*, 828, 21, doi: [10.3847/0004-637X/828/1/21](https://doi.org/10.3847/0004-637X/828/1/21)
- Nanayakkara, T., Brinchmann, J., Glazebrook, K., et al. 2020, *ApJ*, 889, 180, doi: [10.3847/1538-4357/ab65eb](https://doi.org/10.3847/1538-4357/ab65eb)
- Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, *ApJ*, 660, L43, doi: [10.1086/517926](https://doi.org/10.1086/517926)
- Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, *A&A*, 507, 1793, doi: [10.1051/0004-6361/200912497](https://doi.org/10.1051/0004-6361/200912497)
- Oesch, P. A., Bouwens, R. J., Carollo, C. M., et al. 2010, *ApJ*, 725, L150, doi: [10.1088/2041-8205/725/2/L150](https://doi.org/10.1088/2041-8205/725/2/L150)
- Oesch, P. A., Brammer, G., van Dokkum, P. G., et al. 2016, *ApJ*, 819, 129, doi: [10.3847/0004-637X/819/2/129](https://doi.org/10.3847/0004-637X/819/2/129)
- Oke, J. B., & Gunn, J. E. 1983, *ApJ*, 266, 713, doi: [10.1086/160817](https://doi.org/10.1086/160817)
- Ono, Y., Harikane, Y., Ouchi, M., et al. 2023, *ApJ*, 951, 72, doi: [10.3847/1538-4357/acd44a](https://doi.org/10.3847/1538-4357/acd44a)
- Onodera, M., Shimakawa, R., Suzuki, T. L., et al. 2020, *ApJ*, 904, 180, doi: [10.3847/1538-4357/abc174](https://doi.org/10.3847/1538-4357/abc174)
- Osterbrock, D. E. 1989, *Astrophysics of gaseous nebulae and active galactic nuclei*

- Oteo, I., Sobral, D., Ivison, R. J., et al. 2015, MNRAS, 452, 2018, doi: [10.1093/mnras/stv1284](https://doi.org/10.1093/mnras/stv1284)
- Ouchi, M., Shimasaku, K., Furusawa, H., et al. 2003, ApJ, 582, 60, doi: [10.1086/344476](https://doi.org/10.1086/344476)
- Ouchi, M., Mobasher, B., Shimasaku, K., et al. 2009, ApJ, 706, 1136, doi: [10.1088/0004-637X/706/2/1136](https://doi.org/10.1088/0004-637X/706/2/1136)
- Pagel, B. E. J., Edmunds, M. G., Blackwell, D. E., Chun, M. S., & Smith, G. 1979, MNRAS, 189, 95, doi: [10.1093/mnras/189.1.95](https://doi.org/10.1093/mnras/189.1.95)
- Pagel, B. E. J., Simonson, E. A., Terlevich, R. J., & Edmunds, M. G. 1992, MNRAS, 255, 325, doi: [10.1093/mnras/255.2.325](https://doi.org/10.1093/mnras/255.2.325)
- Pahl, A. J., Shapley, A., Steidel, C. C., Chen, Y., & Reddy, N. A. 2021, MNRAS, 505, 2447, doi: [10.1093/mnras/stab1374](https://doi.org/10.1093/mnras/stab1374)
- Peebles, P. J., & Ratra, B. 2003, Reviews of Modern Physics, 75, 559, doi: [10.1103/RevModPhys.75.559](https://doi.org/10.1103/RevModPhys.75.559)
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266, doi: [10.1086/340952](https://doi.org/10.1086/340952)
- . 2010, AJ, 139, 2097, doi: [10.1088/0004-6256/139/6/2097](https://doi.org/10.1088/0004-6256/139/6/2097)
- Pentericci, L., Vanzella, E., Fontana, A., et al. 2014, ApJ, 793, 113, doi: [10.1088/0004-637X/793/2/113](https://doi.org/10.1088/0004-637X/793/2/113)
- Perrin, M. D., Sivaramakrishnan, A., Lajoie, C.-P., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9143, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave, ed. J. Oschmann, Jacobus M., M. Clampin, G. G. Fazio, & H. A. MacEwen, 91433X, doi: [10.1117/12.2056689](https://doi.org/10.1117/12.2056689)
- Perrin, M. D., Soummer, R., Elliott, E. M., Lallo, M. D., & Sivaramakrishnan, A. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave, ed. M. C. Clampin, G. G. Fazio, H. A. MacEwen, & J. Oschmann, Jacobus M., 84423D, doi: [10.1117/12.925230](https://doi.org/10.1117/12.925230)
- Persson, S. E., Murphy, D. C., Smee, S., et al. 2013, PASP, 125, 654, doi: [10.1086/671164](https://doi.org/10.1086/671164)
- Pettini, M., & Pagel, B. E. J. 2004, MNRAS, 348, L59, doi: [10.1111/j.1365-2966.2004.07591.x](https://doi.org/10.1111/j.1365-2966.2004.07591.x)
- Phillipps, S., Drinkwater, M. J., Gregg, M. D., & Jones, J. B. 2001, ApJ, 560, 201, doi: [10.1086/322517](https://doi.org/10.1086/322517)
- Price, S. H., Kriek, M., Brammer, G. B., et al. 2014, ApJ, 788, 86, doi: [10.1088/0004-637X/788/1/86](https://doi.org/10.1088/0004-637X/788/1/86)
- Reddy, N. A., Kriek, M., Shapley, A. E., et al. 2015, ApJ, 806, 259, doi: [10.1088/0004-637X/806/2/259](https://doi.org/10.1088/0004-637X/806/2/259)
- Reddy, N. A., Shapley, A. E., Sanders, R. L., et al. 2018, ApJ, 869, 92, doi: [10.3847/1538-4357/aad1e](https://doi.org/10.3847/1538-4357/aad1e)
- Reddy, N. A., Shapley, A. E., Kriek, M., et al. 2020, ApJ, 902, 123, doi: [10.3847/1538-4357/abb674](https://doi.org/10.3847/1538-4357/abb674)
- Renaud, F., Bournaud, F., & Duc, P.-A. 2015, MNRAS, 446, 2038, doi: [10.1093/mnras/stu2208](https://doi.org/10.1093/mnras/stu2208)
- Renaud, F., Bournaud, F., Kraljic, K., & Duc, P. A. 2014, MNRAS, 442, L33, doi: [10.1093/mnrasl/slu050](https://doi.org/10.1093/mnrasl/slu050)

- Rhoads, J. E., Malhotra, S., Dey, A., et al. 2000, *ApJ*, 545, L85, doi: [10.1086/317874](https://doi.org/10.1086/317874)
- Rich, J. A., Dopita, M. A., Kewley, L. J., & Rupke, D. S. N. 2010, *ApJ*, 721, 505, doi: [10.1088/0004-637X/721/1/505](https://doi.org/10.1088/0004-637X/721/1/505)
- Richard, J., Jones, T., Ellis, R., et al. 2011, *MNRAS*, 413, 643, doi: [10.1111/j.1365-2966.2010.18161.x](https://doi.org/10.1111/j.1365-2966.2010.18161.x)
- Rieke, M. J., Robertson, B., Tacchella, S., et al. 2023, *ApJS*, 269, 16, doi: [10.3847/1538-4365/acf44d](https://doi.org/10.3847/1538-4365/acf44d)
- Roberts, M. S., & Haynes, M. P. 1994, *ARA&A*, 32, 115, doi: [10.1146/annurev.aa.32.090194.000555](https://doi.org/10.1146/annurev.aa.32.090194.000555)
- Robertson, B. E., Ellis, R. S., Dunlop, J. S., McLure, R. J., & Stark, D. P. 2010, *Nature*, 468, 49, doi: [10.1038/nature09527](https://doi.org/10.1038/nature09527)
- Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, *ApJ*, 802, L19, doi: [10.1088/2041-8205/802/2/L19](https://doi.org/10.1088/2041-8205/802/2/L19)
- Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, *ApJ*, 768, 71, doi: [10.1088/0004-637X/768/1/71](https://doi.org/10.1088/0004-637X/768/1/71)
- Robertson, B. E., Tacchella, S., Johnson, B. D., et al. 2023, *Nature Astronomy*, 7, 611, doi: [10.1038/s41550-023-01921-1](https://doi.org/10.1038/s41550-023-01921-1)
- Rodighiero, G., Vaccari, M., Franceschini, A., et al. 2010, *A&A*, 515, A8, doi: [10.1051/0004-6361/200912058](https://doi.org/10.1051/0004-6361/200912058)
- Rodriguez-Gomez, V., Snyder, G. F., Lotz, J. M., et al. 2019, *MNRAS*, 483, 4140, doi: [10.1093/mnras/sty3345](https://doi.org/10.1093/mnras/sty3345)
- Ronayne, K., Papovich, C., Yang, G., et al. 2024, *ApJ*, 970, 61, doi: [10.3847/1538-4357/ad5006](https://doi.org/10.3847/1538-4357/ad5006)
- Rubin, V. C., Ford, W. K., J., & Thonnard, N. 1980, *ApJ*, 238, 471, doi: [10.1086/158003](https://doi.org/10.1086/158003)
- Rubin, V. C., & Ford, W. Kent, J. 1970, *ApJ*, 159, 379, doi: [10.1086/150317](https://doi.org/10.1086/150317)
- Saito, S., de la Torre, S., Ilbert, O., et al. 2020, *MNRAS*, 494, 199, doi: [10.1093/mnras/staa727](https://doi.org/10.1093/mnras/staa727)
- Saldana-Lopez, A., Schaerer, D., Chisholm, J., et al. 2023, *MNRAS*, 522, 6295, doi: [10.1093/mnras/stad1283](https://doi.org/10.1093/mnras/stad1283)
- Salpeter, E. E. 1955, *ApJ*, 121, 161, doi: [10.1086/145971](https://doi.org/10.1086/145971)
- Sanders, D. B., Mazzarella, J. M., Kim, D. C., Surace, J. A., & Soifer, B. T. 2003, *AJ*, 126, 1607, doi: [10.1086/376841](https://doi.org/10.1086/376841)
- Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2015, *ApJ*, 799, 138, doi: [10.1088/0004-637X/799/2/138](https://doi.org/10.1088/0004-637X/799/2/138)
- . 2016, *ApJ*, 816, 23, doi: [10.3847/0004-637X/816/1/23](https://doi.org/10.3847/0004-637X/816/1/23)
- . 2018, *ApJ*, 858, 99, doi: [10.3847/1538-4357/aabcdb](https://doi.org/10.3847/1538-4357/aabcdb)
- Sanders, R. L., Shapley, A. E., Reddy, N. A., et al. 2020, *MNRAS*, 491, 1427, doi: [10.1093/mnras/stz3032](https://doi.org/10.1093/mnras/stz3032)
- Sanders, R. L., Shapley, A. E., Jones, T., et al. 2021, *ApJ*, 914, 19, doi: [10.3847/1538-4357/abf4c1](https://doi.org/10.3847/1538-4357/abf4c1)

- Santoro, F., Kreckel, K., Belfiore, F., et al. 2022, *A&A*, 658, A188, doi: [10.1051/0004-6361/202141907](https://doi.org/10.1051/0004-6361/202141907)
- Saxena, A., Robertson, B. E., Bunker, A. J., et al. 2023, *A&A*, 678, A68, doi: [10.1051/0004-6361/202346245](https://doi.org/10.1051/0004-6361/202346245)
- Schechter, P. 1976, *ApJ*, 203, 297, doi: [10.1086/154079](https://doi.org/10.1086/154079)
- Schenker, M. A., Ellis, R. S., Konidaris, N. P., & Stark, D. P. 2014, *ApJ*, 795, 20, doi: [10.1088/0004-637X/795/1/20](https://doi.org/10.1088/0004-637X/795/1/20)
- Scoville, N., Abraham, R. G., Aussel, H., et al. 2007, *ApJS*, 172, 38, doi: [10.1086/516580](https://doi.org/10.1086/516580)
- Shapiro, K. L., Genzel, R., & Förster Schreiber, N. M. 2010, *MNRAS*, 403, L36, doi: [10.1111/j.1745-3933.2010.00810.x](https://doi.org/10.1111/j.1745-3933.2010.00810.x)
- Shapley, A. E., Reddy, N. A., Kriek, M., et al. 2015, *ApJ*, 801, 88, doi: [10.1088/0004-637X/801/2/88](https://doi.org/10.1088/0004-637X/801/2/88)
- Shapley, A. E., Sanders, R. L., Topping, M. W., et al. 2024, arXiv e-prints, arXiv:2407.00157, doi: [10.48550/arXiv.2407.00157](https://doi.org/10.48550/arXiv.2407.00157)
- Shen, L., Papovich, C., Matharu, J., et al. 2024, *ApJ*, 963, L49, doi: [10.3847/2041-8213/ad28bd](https://doi.org/10.3847/2041-8213/ad28bd)
- Shibuya, T., Ouchi, M., & Harikane, Y. 2015, *ApJS*, 219, 15, doi: [10.1088/0067-0049/219/2/15](https://doi.org/10.1088/0067-0049/219/2/15)
- Shibuya, T., Ouchi, M., Kubo, M., & Harikane, Y. 2016, *ApJ*, 821, 72, doi: [10.3847/0004-637X/821/2/72](https://doi.org/10.3847/0004-637X/821/2/72)
- Shimakawa, R., Kodama, T., Hayashi, M., et al. 2017, *MNRAS*, 468, L21, doi: [10.1093/mnras1/slx019](https://doi.org/10.1093/mnras1/slx019)
- . 2018a, *MNRAS*, 473, 1977, doi: [10.1093/mnras/stx2494](https://doi.org/10.1093/mnras/stx2494)
- Shimakawa, R., Koyama, Y., Röttgering, H. J. A., et al. 2018b, *MNRAS*, 481, 5630, doi: [10.1093/mnras/sty2618](https://doi.org/10.1093/mnras/sty2618)
- Shivaei, I., Reddy, N. A., Shapley, A. E., et al. 2015, *ApJ*, 815, 98, doi: [10.1088/0004-637X/815/2/98](https://doi.org/10.1088/0004-637X/815/2/98)
- Shivaei, I., Reddy, N. A., Siana, B., et al. 2018, *ApJ*, 855, 42, doi: [10.3847/1538-4357/aaad62](https://doi.org/10.3847/1538-4357/aaad62)
- Silverman, J. D., Kashino, D., Sanders, D., et al. 2015, *ApJS*, 220, 12, doi: [10.1088/0067-0049/220/1/12](https://doi.org/10.1088/0067-0049/220/1/12)
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, 131, 1163, doi: [10.1086/498708](https://doi.org/10.1086/498708)
- Smit, R., Bouwens, R. J., Labbé, I., et al. 2014, *ApJ*, 784, 58, doi: [10.1088/0004-637X/784/1/58](https://doi.org/10.1088/0004-637X/784/1/58)
- Smith, L. J., Westmoquette, M. S., Gallagher, J. S., et al. 2006, *MNRAS*, 370, 513, doi: [10.1111/j.1365-2966.2006.10507.x](https://doi.org/10.1111/j.1365-2966.2006.10507.x)
- Sobral, D., Santos, S., Matthee, J., et al. 2018, *MNRAS*, 476, 4725, doi: [10.1093/mnras/sty378](https://doi.org/10.1093/mnras/sty378)
- Sobral, D., Smail, I., Best, P. N., et al. 2013, *MNRAS*, 428, 1128, doi: [10.1093/mnras/sts096](https://doi.org/10.1093/mnras/sts096)
- Sobral, D., Best, P. N., Geach, J. E., et al. 2009, *MNRAS*, 398, 75, doi: [10.1111/j.1365-2966.2009.15129.x](https://doi.org/10.1111/j.1365-2966.2009.15129.x)
- Sparke, L. S., & Gallagher, John S., I. 2007, *Galaxies in the Universe: An Introduction*

- Sparre, M., Hayward, C. C., Feldmann, R., et al. 2017, MNRAS, 466, 88, doi: [10.1093/mnras/stw3011](https://doi.org/10.1093/mnras/stw3011)
- Sparre, M., Whittingham, J., Damle, M., et al. 2022, MNRAS, 509, 2720, doi: [10.1093/mnras/stab3171](https://doi.org/10.1093/mnras/stab3171)
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15, doi: [10.1088/0067-0049/214/2/15](https://doi.org/10.1088/0067-0049/214/2/15)
- Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175, doi: [10.1086/377226](https://doi.org/10.1086/377226)
- Spitler, L. R., Labbé, I., Glazebrook, K., et al. 2012, ApJL, 748, L21, doi: [10.1088/2041-8205/748/2/L21](https://doi.org/10.1088/2041-8205/748/2/L21)
- Stanway, E. R., & Eldridge, J. J. 2018, MNRAS, 479, 75, doi: [10.1093/mnras/sty1353](https://doi.org/10.1093/mnras/sty1353)
- Stark, D. P., Schenker, M. A., Ellis, R., et al. 2013, ApJ, 763, 129, doi: [10.1088/0004-637X/763/2/129](https://doi.org/10.1088/0004-637X/763/2/129)
- Stark, D. P., Walth, G., Charlot, S., et al. 2015, MNRAS, 454, 1393, doi: [10.1093/mnras/stv1907](https://doi.org/10.1093/mnras/stv1907)
- Stark, D. P., Ellis, R. S., Charlot, S., et al. 2017, MNRAS, 464, 469, doi: [10.1093/mnras/stw2233](https://doi.org/10.1093/mnras/stw2233)
- Stasińska, G., & Leitherer, C. 1996, ApJS, 107, 661, doi: [10.1086/192377](https://doi.org/10.1086/192377)
- Stefanon, M., Bouwens, R. J., Illingworth, G. D., et al. 2022, ApJ, 935, 94, doi: [10.3847/1538-4357/ac7e44](https://doi.org/10.3847/1538-4357/ac7e44)
- Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, ApJ, 592, 728, doi: [10.1086/375772](https://doi.org/10.1086/375772)
- Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2018, ApJ, 869, 123, doi: [10.3847/1538-4357/aaed28](https://doi.org/10.3847/1538-4357/aaed28)
- Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, ApJ, 462, L17, doi: [10.1086/310029](https://doi.org/10.1086/310029)
- Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, ApJ, 795, 165, doi: [10.1088/0004-637X/795/2/165](https://doi.org/10.1088/0004-637X/795/2/165)
- Steinmetz, M., & Navarro, J. F. 2002, New A, 7, 155, doi: [10.1016/S1384-1076\(02\)00102-1](https://doi.org/10.1016/S1384-1076(02)00102-1)
- Straatman, C. M. S., Spitler, L. R., Quadri, R. F., et al. 2016, ApJ, 830, 51, doi: [10.3847/0004-637X/830/1/51](https://doi.org/10.3847/0004-637X/830/1/51)
- Suzuki, R., Tokoku, C., Ichikawa, T., et al. 2008, PASJ, 60, 1347, doi: [10.1093/pasj/60.6.1347](https://doi.org/10.1093/pasj/60.6.1347)
- Suzuki, T. L., Kodama, T., Tadaki, K.-i., et al. 2015, ApJ, 806, 208, doi: [10.1088/0004-637X/806/2/208](https://doi.org/10.1088/0004-637X/806/2/208)
- Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, MNRAS, 457, 2790, doi: [10.1093/mnras/stw131](https://doi.org/10.1093/mnras/stw131)
- Tacchella, S., Eisenstein, D. J., Hainline, K., et al. 2023, ApJ, 952, 74, doi: [10.3847/1538-4357/acdbc6](https://doi.org/10.3847/1538-4357/acdbc6)
- Tacconi, L. J., Genzel, R., & Sternberg, A. 2020, ARA&A, 58, 157, doi: [10.1146/annurev-astro-082812-141034](https://doi.org/10.1146/annurev-astro-082812-141034)
- Tadaki, K.-i., Kodama, T., Tanaka, I., et al. 2014, ApJ, 780, 77, doi: [10.1088/0004-637X/780/1/77](https://doi.org/10.1088/0004-637X/780/1/77)

- Tamburello, V., Mayer, L., Shen, S., & Wadsley, J. 2015, MNRAS, 453, 2490, doi: [10.1093/mnras/stv1695](https://doi.org/10.1093/mnras/stv1695)
- Tang, M., Stark, D. P., Chevallard, J., & Charlot, S. 2019, MNRAS, 489, 2572, doi: [10.1093/mnras/stz2236](https://doi.org/10.1093/mnras/stz2236)
- Terao, Y., Spitler, L. R., Motohara, K., & Chen, N. 2022, ApJ, 941, 70, doi: [10.3847/1538-4357/ac9fce](https://doi.org/10.3847/1538-4357/ac9fce)
- Theios, R. L., Steidel, C. C., Strom, A. L., et al. 2019, ApJ, 871, 128, doi: [10.3847/1538-4357/aaf386](https://doi.org/10.3847/1538-4357/aaf386)
- Toomre, A. 1964, ApJ, 139, 1217, doi: [10.1086/147861](https://doi.org/10.1086/147861)
- Topping, M. W., Stark, D. P., Endsley, R., et al. 2022, ApJ, 941, 153, doi: [10.3847/1538-4357/aca522](https://doi.org/10.3847/1538-4357/aca522)
- Topping, M. W., Shapley, A. E., Sanders, R. L., et al. 2021, MNRAS, 506, 1237, doi: [10.1093/mnras/stab1793](https://doi.org/10.1093/mnras/stab1793)
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898, doi: [10.1086/423264](https://doi.org/10.1086/423264)
- Van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., et al. 2014, PeerJ, 2, e453
- van der Wel, A., Straughn, A. N., Rix, H. W., et al. 2011, ApJ, 742, 111, doi: [10.1088/0004-637X/742/2/111](https://doi.org/10.1088/0004-637X/742/2/111)
- van Dokkum, P. G., Labbé, I., Marchesini, D., et al. 2009, PASP, 121, 2, doi: [10.1086/597138](https://doi.org/10.1086/597138)
- Vanzella, E., Giavalisco, M., Inoue, A. K., et al. 2010, ApJ, 725, 1011, doi: [10.1088/0004-637X/725/1/1011](https://doi.org/10.1088/0004-637X/725/1/1011)
- Vanzella, E., Castellano, M., Meneghetti, M., et al. 2017, ApJ, 842, 47, doi: [10.3847/1538-4357/aa74ae](https://doi.org/10.3847/1538-4357/aa74ae)
- Vanzella, E., Castellano, M., Bergamini, P., et al. 2022, A&A, 659, A2, doi: [10.1051/0004-6361/202141590](https://doi.org/10.1051/0004-6361/202141590)
- Vanzella, E., Claeysens, A., Welch, B., et al. 2023, ApJ, 945, 53, doi: [10.3847/1538-4357/acb59a](https://doi.org/10.3847/1538-4357/acb59a)
- Wang, K., Peng, E. W., Liu, C., et al. 2023, Nature, 623, 296, doi: [10.1038/s41586-023-06650-z](https://doi.org/10.1038/s41586-023-06650-z)
- Weisz, D. R., Johnson, B. D., Johnson, L. C., et al. 2012, ApJ, 744, 44, doi: [10.1088/0004-637X/744/1/44](https://doi.org/10.1088/0004-637X/744/1/44)
- Whitaker, K. E., Franx, M., Leja, J., et al. 2014, ApJ, 795, 104, doi: [10.1088/0004-637X/795/2/104](https://doi.org/10.1088/0004-637X/795/2/104)
- Whitler, L., Stark, D. P., Endsley, R., et al. 2023, arXiv e-prints, arXiv:2305.16670, doi: [10.48550/arXiv.2305.16670](https://doi.org/10.48550/arXiv.2305.16670)
- Wisnioski, E., Förster Schreiber, N. M., Fossati, M., et al. 2019, ApJ, 886, 124, doi: [10.3847/1538-4357/ab4db8](https://doi.org/10.3847/1538-4357/ab4db8)
- Wuyts, E., Wisnioski, E., Fossati, M., et al. 2016, ApJ, 827, 74, doi: [10.3847/0004-637X/827/1/74](https://doi.org/10.3847/0004-637X/827/1/74)
- Wuyts, S., Förster Schreiber, N. M., Genzel, R., et al. 2012, ApJ, 753, 114, doi: [10.1088/0004-637X/753/2/114](https://doi.org/10.1088/0004-637X/753/2/114)
- Yabe, K., Ohta, K., Iwamuro, F., et al. 2014, MNRAS, 437, 3647, doi: [10.1093/mnras/stt2185](https://doi.org/10.1093/mnras/stt2185)

- Yang, H., Malhotra, S., Rhoads, J. E., & Wang, J. 2017a, ApJ, 847, 38, doi: [10.3847/1538-4357/aa8809](https://doi.org/10.3847/1538-4357/aa8809)
- Yang, H., Malhotra, S., Gronke, M., et al. 2017b, ApJ, 844, 171, doi: [10.3847/1538-4357/aa7d4d](https://doi.org/10.3847/1538-4357/aa7d4d)
- Yeh, S. C. C., & Matzner, C. D. 2012, ApJ, 757, 108, doi: [10.1088/0004-637X/757/2/108](https://doi.org/10.1088/0004-637X/757/2/108)
- York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579, doi: [10.1086/301513](https://doi.org/10.1086/301513)
- Yoshii, Y., Aoki, T., Doi, M., et al. 2010, in SPIE, Vol. 7733, , 773308, doi: [10.1117/12.856680](https://doi.org/10.1117/12.856680)
- Zanella, A., Daddi, E., Le Floch, E., et al. 2015, Nature, 521, 54, doi: [10.1038/nature14409](https://doi.org/10.1038/nature14409)
- Zitrin, A., Broadhurst, T., Barkana, R., Rephaeli, Y., & Benítez, N. 2011, MNRAS, 410, 1939, doi: [10.1111/j.1365-2966.2010.17574.x](https://doi.org/10.1111/j.1365-2966.2010.17574.x)