

NEW INSIGHTS INTO PRIMEVAL GALAXIES IN THE ERA OF JWST



Callum Edward Charles Witten

Supervisors: Prof. Debora Sijacki
Dr. Nicolas Laporte

Institute of Astronomy
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

WOLFSON COLLEGE

September 2024

For Grandad, who seemed to know years before anybody else that I would end up here.

DECLARATION

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted, or is being concurrently submitted, for any degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

The use of 'we' instead of 'I' throughout various parts of this thesis merely reflects a stylistic choice. In particular, the following chapters are based on published or submitted articles resulting from research carried out in collaboration:

i. Chapter 2 is based on work completed in collaboration with N. Laporte and H. Katz, published in Witten et al. (2023). All analysis in this chapter was perfomed by me unless otherwise stated in the text.

ii. Chapter 3 is based on work completed in collaboration with N. Laporte, S. Martin-Alvarez, D. Sijacki, Y. Yuan, M. G. Haehnelt, W. M. Baker, J. S. Dunlop, R. S. Ellis, N. A. Grogin, G. Illingworth, H. Katz, A. M. Koekemoer, D. Magee, R. Maiolino, W. McClymont, P. Pérez-González, D. Puskás, G. Roberts-Borsani, P. Santini and C. Simmonds, published in Witten et al. (2024b). The observational analysis and data reduction was performed by me unless otherwise stated in the text. The Simulation analysis and post-processing was led by Sergio Martin-Alvarez and Yuxuan Yuan.

iii. Chapter 4 is based on work completed in collaboration with J. Scholtz, N. Laporte, H. Übler, M. Perna, R. Maiolino, S. Arribas, W. M. Baker, J. S. Bennett, F. D'Eugenio, C. Simmonds, S. Tacchella, J. Witstok, A. J. Bunker, S. Carniani, S. Charlot, G. Cresci, E. Curtis-Lake, D. J. Eisenstein, B. Robertson B. Rodríguez Del Pino, R. Smit, G. Venturi, C. C. Williams and C. N. A. Willmer, published as the dual first author paper Scholtz et al. (2024). The identification of a potential protocluster core was done by me. The

analysis of the Lyα halo was joint work with Jan Scholtz. The data reduction was performed by Hannah Übler and Michele Perna.

iv. Chapter 5 is based on work completed in collaboration with W. McClymont, N. Laporte, G. Roberts-Borsani, D. Sijacki, S. Tacchella, C. Simmonds, H. Katz, R. S. Ellis, J. Witstok, R. Maiolino, X. Ji, B. R. Hayes, T. J. Looser and F. D'Eugenio, submitted to MNRAS, in preprint as Witten et al. (2024a). All analysis in this chapter was performed by me unless otherwise stated in the text.

Callum Edward Charles Witten
September 2024

ACKNOWLEDGEMENTS

I owe a great debt to many people who have made this thesis possible and made the previous three years not just bearable, but thoroughly enjoyable.

First and foremost, I want to thank *the pioneer*, Nicolas Laporte, for his incredible support, encouragement and generosity. It has been a privilege to work with someone so able to put their finger on the pulse of the high-redshift community and move towards new areas of interest before the rest of the field gets there. I have never met someone who cares so deeply for the wellbeing of PhD students, or a faculty member who pushes students so much *not* to work, but to find a healthy work-life balance. You have instilled in me a love of hikes, wine and the country of France, and I think few British PhD students can say *that* is the outcome of their PhD studies.

In a very close second, I want to thank Debora Sijacki. Your support and kindness in allowing me to freely stumble into the world of *JWST*, even when that was not your easiest option, is something for which I will forever be grateful. Your advice has been invaluable in both creating the work in this thesis and navigating the last three years. I have never met someone who places such a high value on scientific integrity and I know that I am not the only person who hopes they become half the scientist you are.

My thanks must also go to the wider Cambridge (and Cambridge adjacent) astrophysics community, to name a few: Martin Haehnelt, Sergio Martin-Alvarez, Harley Katz, Jake Bennett, Guido Roberts-Borsani, Richard Ellis, Adam Dillamore, Elliot Davies, Frances Rigby, Adam Ormondroyd, Wyn Evans, David Aguado, Jason Sanders and (last, but arguably most importantly) Vasily Belokurov. Vasily's support in not just my first year, but his advice at multiple points throughout my PhD has been crucial in pointing me in the right direction. I would also like to thank the administrative staff at both the IoA and the Kavli. To name only the tip of an iceberg – Debbie Peterson, Steve Brereton and Alison Wilson.

Next, I must thank my office mates in K31 (the best office in the Kavli): Lola, Will, William and Dily. They have not only helped me think through some of my (sometimes

unhinged) scientific ideas, they have also supported me through some of the more challenging parts of the PhD. Significant recognition must go to Will M, our many long discussions have radically improved me as an astrophysicist. I like to think this has been a mutually beneficial relationship, but I fear, while you have made me skeptical about most observations (and simulations for that matter), I may have made you trust them too much.

The extra-galactic group at the Kavli is nothing without its outstanding leadership team: Roberto Maiolino and Sandro Tacchella. I am incredibly grateful for their generosity in allowing a student, who is technically not at all their responsibility, to attend both their student and group meetings and to work closely with them, especially in the competitive field that is high-redshift galaxies. Learning about *JWST* from the many experts in the group has inevitably significantly improved this thesis. On that note, I should thank the many postdocs and fellows in the group who have helped me along the way: Charlotte Simmonds, Joris Witstok, Jan Scholtz, Hannah Übler, Gareth Jones and Francesco D'Eugenio. Roberto and Sandro have sculpted an incredible group of students, to name just a few past and present that I will dearly miss: Gabriel Maheson, Dávid Puskás, Tobias Looser, Natalia Villanueva, Roberta Tripodi and Giovanni Mazzolari.

I want to especially thank the "Extra-Gals" (and Zach and Olly). You have made every day spent at the Kavli over the last two years such incredible fun. It is impressive to think we found time to work between May Balls, formals, bops, brunches, lunches and trips for iced coffees, dumplings and the cinema. I love you all so much and I cannot wait to see you all again.

I also would like to acknowledge my colleagues at Laboratoire d'Astrophysique de Marseille, most notably Ilias, Roser and Meriam, for their insightful conversations, help with postdoc applications and their hospitality.

My thanks must also go to my housemates and college friends – Yorick, Josephien, Matt, Sarah, Ras, Nathan, Hamideh, Sula, Cody and Veronica. Without you all, the last three years would have been a vastly different experience. The many trips to the pub, chips after bops and cups of tea we have shared, I believe kept us all somewhat more sane than we may have become.

To Liam, Libby, Ed, Catherine, Ollie, Megan and Kieran, you have all shaped such a large fraction of my life, I love you all so much. While I must thank you for so many things, two things come to the forefront of my mind – thank you for spending half of your lives driving me around and the other half waiting for me to be ready.

To Ben Garland, who believed in me when I wanted to make a very late academic switch to study physics, without your trust I would never have been able to do a physics

degree, let alone a PhD in astronomy. My thanks also go to Louise Evans and Julie Axtell-Powell without whom I may not have had the courage or belief to aim high enough to reach this point. I must also acknowledge Will Bowdery, Lydia Mulligan and Samira Kelly for their ever present friendship and support throughout our undergraduate degrees and Patricia Schady and Stijn Wuyts for their incredible supervision, advice and recommendations, all of which helped put me in this position.

Without the endless love and support of my incredible parents, Emma and Mark, the annoyance of my siblings, James, Lissy and Heath, and the enthusiasm of my grandparents, Tricia, Clive, Kerry and Pauline, I would not have become the person I am today, and in turn this thesis would not exist. I do not fail to note the privileged position that I am in having such a supportive and encouraging family, something many will never experience – I love you all to the moon and back.

To my in-laws, Karen, Dave and Georgia, thank you for the many escapes from work and for accepting me into your family – I am incredibly grateful for the love and support you have offered both me and Lewis over the last few years.

Finally, I would like to thank my incredible husband, Lewis. We decided that the perfect time to get married was during the middle of writing this thesis. Your generosity and willingness to fill in for me when I was struggling to juggle this thesis and the wedding saved me on multiple occasions, and in a very literal sense, neither the wedding nor this thesis would have happened without you. On a more practical note, your daily love and support have made the PhD experience a joyful one. I love you dearly.

SUMMARY

The properties of galaxies that existed in the early Universe hold the key to understanding not only their formation and evolution but also the broader reionisation process. While the *Hubble* Space Telescope (*HST*) was able to identify a small sample of candidate $z > 7$ galaxies, the *James Webb* Space Telescope (*JWST*) has revolutionised the field. The sensitivity of the near infrared camera (NIRCam) and spectrograph (NIRSpec) instruments have facilitated the detection and spectroscopic confirmation of galaxies out to $z \sim 14$. Moreover, the study of the properties of not just individual, but samples of high-redshift galaxies is now possible.

As astronomers utilise the capabilities of this telescope multiple key questions, left over from the *HST*-era, require answering. Here I present the results of several studies that aim to answer some of these important unanswered questions. I utilise ground-based observations of the intrinsically strongest hydrogen emission line, Lyman-α (Lyα), to constrain the escape fraction of UV-bright galaxies in the early Universe. This shows that UV-bright galaxies appear to have low escape fractions of ionising photons, setting the scene for *JWST* studies of the ionising capabilities of different galaxy populations. I then use NIRCam observations of Lyα-emitting galaxies to explain the escape of Lyα photons from galaxies by showing that all high-redshift Lyα emitters (LAEs) have close companions, in contrast to typical high-redshift galaxies. The combination of observations and simulations, in this work, shows how interactions between close companion galaxies can both enhance Lyα emission, but also facilitate its escape. The following work I present focuses on observations of one of the highest redshift LAEs with the integral field unit on *JWST*. These observations reveal both a Lyα halo around this galaxy as well as an abundance of candidate UV-faint emission line galaxies, evidencing the complex escape of Lyα but also the environments within which high-redshift galaxies exist. Following this, I identify a galaxy that shows both strong emission lines and a Balmer break indicating the presence of both a young and old stellar population, and hence a rejuvenating star-formation history (SFH), revealing the complex SFHs of early galaxies. Finally, I discuss

these results and future avenues for this work, especially in the context of next generation instruments and simulations.

I also point the reader to some additional research conducted as part of my PhD, but not included in this thesis, in Witten et al. (2022). This work focuses on the low-resolution spectrographs on-board the *Gaia* satellite. This work, completed before the release of *Gaia* data release 3, addresses the potential information available in these low-resolution spectra to measure the temperature, surface gravity and metallicity of observed stars. I also show that these spectra can be used to identify candidate carbon-enriched metal-poor stars for follow-up high-resolution observations.

TABLE OF CONTENTS

List of figures **xvii**

List of tables **xix**

1 Introduction **1**

1.1 The known history of the early Universe 1

1.2 The Formation of Structure 3

1.2.1 Linear regime 3

1.2.2 Non-linear regime 5

1.2.3 Dark matter 6

1.2.4 First galaxies 7

1.2.5 First stars 10

1.2.6 Initial mass function 12

1.2.7 Early galaxy evolution: Dwarf galaxies to overdense structures 14

1.3 Models of Reionisation 15

1.3.1 When did reionisation occur? 17

1.3.2 Which astrophysical sources drove reionisation? 19

1.3.3 What is the morphology of reionisation? 23

1.4 Galaxy spectral properties in the early Universe 25

1.4.1 The Lyman-break 25

1.4.2 The UV-slope 26

1.4.3 Nebular emission 28

1.4.4 The Balmer break 30

1.5 Galaxy photometric properties in the early Universe 31

1.6 The James Webb Space Telescope 32

1.7 Open questions 34

2 Evidence for a low Lyman Continuum Escape fraction in three Massive, UV-bright galaxies at $z > 7$ 37
2.1 Introduction . . . 37
2.2 Targets selection . . . 38
2.3 Method . . . 41
2.4 Results . . . 42
2.4.1 Constraints from SED . . . 42
2.4.2 Lyα profile . . . 46
2.5 Discussion . . . 46
2.5.1 Interpretation of the velocity offset . . . 46
2.5.2 Interpretation of the red-peak asymmetry . . . 47
2.5.3 Effects of IGM attenuation . . . 48
2.6 Summary . . . 51

3 Deciphering Lyman-α Emission Deep into the Epoch of Reionisation 53
3.1 Introduction . . . 53
3.2 Data . . . 55
3.2.1 Data Reduction . . . 56
3.3 Results . . . 57
3.3.1 Redshift Determination . . . 59
3.3.2 Galaxy properties . . . 68
3.3.3 Close companion fraction of non-LAEs . . . 70
3.3.4 Close companions are necessary but not sufficient for Lyα emission . 70
3.4 *Azahar* simulation . . . 71
3.4.1 RASCAS post-processing . . . 72
3.4.2 Galaxy selection and measurements in the simulations . . . 74
3.4.3 Mergers in the *Azahar* simulations . . . 76
3.4.4 Simulated ionised bubble evolution in the neutral IGM . . . 81
3.5 Conclusion . . . 81

4 The environment of an AGN at z =10.603: New insights on the most distant Lyα detection 83
4.1 Introduction . . . 83
4.2 Observations and data reduction . . . 85
4.3 Properties of the most distant Lyα halo . . . 86
4.4 Search for companions in the vicinity of GN-z11 . . . 90
4.4.1 Lyα emitting blobs . . . 91

4.4.2 A candidate $z = 10.6$ protocluster core 94
4.5 Conclusions 95
4.6 Appendix 97
4.6.1 Tentative detection of CIV in NE blob 97

5 Rising from the ashes: evidence of old stellar populations and rejuvenation events in the very early Universe 99
5.1 Introduction 99
5.2 Data 102
5.3 A2744-YD4 104
5.4 Stack 108
5.5 Modelling 111
5.5.1 Simple stellar population analysis 111
5.5.2 SED fitting 114
5.5.3 SPHINX 121
5.6 Discussion and conclusion 126

6 Conclusions 131
6.1 Conclusions 131
6.2 Future outlook 133
6.2.1 Wide surveys 133
6.2.2 *JWST* 135
6.2.3 Simulations 135
6.3 Final Remarks 136

References 139

LIST OF FIGURES

1.1 The CMB temperature fluctuation map as seen by Planck. 3
1.2 Atomic cooling diagram and rates for different temperatures, densities and metallicities of the gas. 8
1.3 The mass distribution of stellar populations assuming different initial mass functions. 13
1.4 The neutral hydrogen fraction as a function of redshift. 18
1.5 The evolution of the UV luminosity function as a function of redshift. 21
1.6 An example of a high-redshift galaxy spectrum that shows evidence of strong nebular continuum and models of each component of the nebular continuum. 25
1.7 The sensitivity of NIRCam and NIRSpec relative to other telescopes. 33

2.1 Lyα profiles seen in high-redshift galaxies. 43
2.2 Escape fraction of Lyman-continuum diagnostics. 45
2.3 The potential intrinsic Lyα profile of high-redshift galaxies. 48
2.4 The effects of absorption from neutral hydrogen in the intergalactic medium on the profile of the Lyα emission line. 49

3.1 Modelling of the structural components of a high-redshift galaxy. 57
3.2 An abundance of galaxy mergers seen with NIRCam. 58
3.3 SED-fits of all Lyα-emitting galaxies and their companions. 60
3.4 Figure 3.3 continued. 61
3.5 Constraints on the Lyα emission from companion galaxies. 65
3.6 NIRCam wide field slitless spectroscopy of Lyα-emitters and their companions. 66
3.7 The effects of dust attenuation on the Hα and Lyα emission from a simulated galaxy. 75
3.8 Comparison of *Azahar* to an observed galaxy merger. 77
3.9 The evolution of galaxy properties with an ongoing merger. 78

4.1 Spectrum and map of the Lyα emission from a $z = 10.6$ galaxy. 87
4.2 The Lyα halo surface brightness of GN-z11. 91
4.3 Detection of further Lyα-emitters around GN-z11. 92
4.4 The evolution of the dark matter halo mass of protoclusters with redshift. . 96
4.5 Tentative detection of C IV from a companion galaxy 97

5.1 The spectrum of a $z \sim 7.9$ galaxy that shows both a Balmer break and strong nebular emission lines. 104
5.2 The Balmer break strengths of high-redshift galaxies. 105
5.3 The excitation-ionisation diagram. 106
5.4 The stacked spectrum of evolved galaxies. 110
5.5 Images of the galaxies selected from our samples. 113
5.6 The star-formation history derived from spectro-photometric fitting. 117
5.7 The relative contribution of different stellar ages to the spectrum of the rejuvenating galaxy. 118
5.8 The Balmer break strength of rejuvenating galaxies as a function of their star-formation histories and dust. 120
5.9 The relationship between the ionisation-excitation plot and the burstiness of star-formation. 122
5.10 The spectral features of simulated galaxies that exhibit different star-formation histories 123
5.11 The Balmer break strength as a function of stellar-mass-weighted age in simulated galaxies. 125

LIST OF TABLES

2.1 The properties of the sample of Lyα-emitters. 39
2.2 The Lyα escape fraction estimated by different diagnostics. 41

3.1 Properties inferred from NIRCam Grism spectroscopy. 67
3.2 The properties of Lyα-emitting galaxies. 69

4.1 Lyα emission properties of the regions around GN-z11. 89

5.1 The emission line properties of YD4. 105

CHAPTER 1

INTRODUCTION

"Embrace the mess. That's where the good stuff lives."

Cinda Canning, *New York, United States*

1.1. THE KNOWN HISTORY OF THE EARLY UNIVERSE

Our understanding of modern cosmology stems back to the discoveries of Georges Lemaître and Edwin Hubble in the 1920's that galaxies, outside of our own, are receding from us at speeds proportional to their distance from ourselves (Lemaître, 1927; Hubble, 1929). This in turn implies that the Universe was once considerably smaller, forming the foundations of Big Bang cosmology.

Assuming the Universe is isotropic and homogeneous, this expansion is only a function of cosmic time and is given by the scale factor, $a(t)$. We can define the cosmological redshift, z, of a given cosmic time, $t_{\rm emit}$, as

$$1 + z = \frac{1}{a(t_{\rm emit})} = \frac{\lambda_{\rm obs}}{\lambda_{\rm emit}}, \tag{1.1}$$

where the expansion of the Universe acts to stretch the emitted wavelength, $\lambda_{\rm emit}$, to longer observed wavelengths, $\lambda_{\rm obs}$.

The density at earlier cosmic times is proportional to $(1+z)^3$ as a result of this expansion. This increase in density also implies an increase in temperature and as such, the early Universe was considerably hotter and denser than it is at current times. At the very earliest times ($\lesssim 380,000$ years after the Big Bang), the Universe consists of a plasma of

baryons, electrons and photons. A few minutes post Big Bang, the Universe is cool enough that the first atomic nuclei can form as the temperature falls below their binding energies. This process is entitled Big Bang Nucleosynthesis (BBN) and was postulated by Alpher & Herman (1949); Gamow (1949). Expansion causes a lower density, reducing the frequency of collisions, combined with the lack of stable nuclei with 5 or 8 nucleons meaning only the lightest elements can exist in non-negligible numbers. BBN predicts that the Universe at this early time was 75% hydrogen and 25% helium by mass, with small amounts of lithium (Peebles, 1966; Wagoner et al., 1967).

Enough of the photons at this early epoch have energies greater than the ionisation potential of neutral hydrogen $h\nu > 13.6$ eV as to keep the Universe ionised. Thomson scattering of electromagnetic radiation by these free electrons ultimately means the Universe is opaque at these early times (Thomson, 1906).

In 1965, the detection of the Cosmic Microwave Background (CMB; Penzias & Wilson, 1965; Dicke et al., 1965) facilitated a vast improvement in the understanding of the temperature evolution of the early Universe. Assuming that the cooling of the CMB is driven purely by the expansion of the Universe with no energy input, i.e. adiabatic cooling, the CMB temperature at a given redshift is $T = (1+z)T_0$, where T_0 is the temperature of the CMB at $z = 0$, which was found to be 2.725K.

After 372,000 years the temperature of the CMB cools enough ($T \sim 4000$K; Peebles, 1968) that the number of photons with energies great enough to ionise neutral hydrogen becomes insufficient to combat the recombination of ionised hydrogen and the Universe becomes neutral. Following recombination, photons are able to permeate through the early Universe and radiation and baryons decouple and are no longer in thermal equilibrium (Mo et al., 2010).

The CMB maps the distribution of photon energies at the surface of last scattering. The temperature map of the CMB (see Planck Collaboration et al., 2016), shown in Figure 1.1 is uniform (on the order of 10^{-4}K) on large scales in all directions, i.e. *isotropic*. If this is also true in all locations in the Universe, i.e. *homogeneous*, these two features compose the cosmological principle. While the cosmological principle governs large scales, on small scales, the cosmological principle is broken. The small temperature fluctuations seen in the CMB evidence this, where matter perturbations drive deeper potential wells, which in turn gravitationally redshift photons, lowering their temperatures. These small density fluctuations grow through gravitational instabilities, eventually forming the "cosmic web".

The period following recombination, where CMB photons first permeate through the early Universe (the photons that we observed today), ironically, is entitled the Dark Ages. Throughout this period, the CMB radiation is redshifted out of the visible wavelength

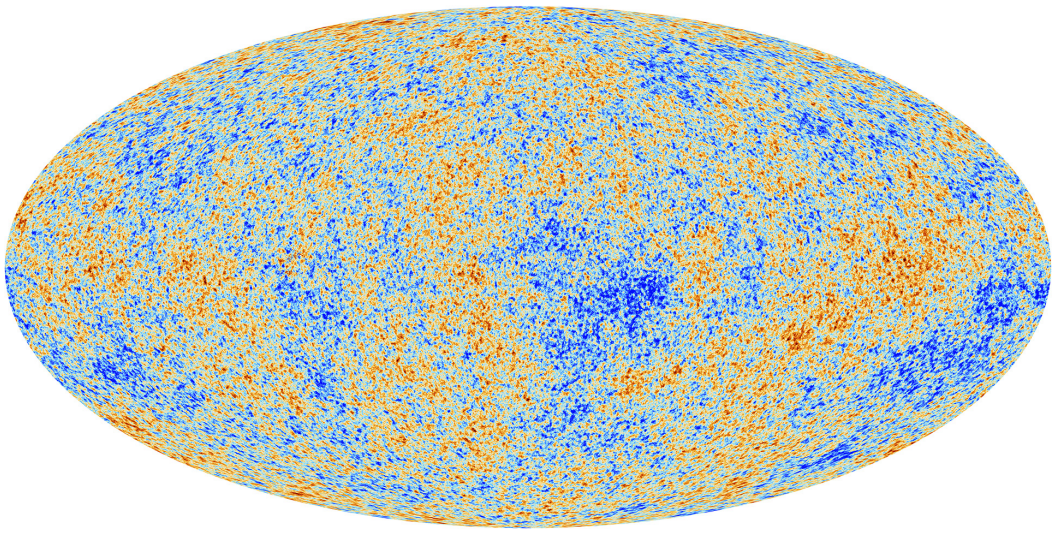


Fig. 1.1 The CMB temperature fluctuation map as seen by Planck, with the colorscale indicating temperature changes from $+300\mu$K (red) to -300μK (blue). Figure credit: ESA and the Planck Collaboration.

range, as such the Universe would indeed have been dark to the human eye. Astronomers have struggled to observe the "dark ages", filling the gap between the observation of the CMB and the light from the first luminous sources. The dark ages remain 'dark' until the first luminous sources emerge, occurring at most 300 million years after the Big Bang (Carniani et al., 2024). The formation of these first luminous sources is a result of the gravitational collapse of the early CMB density fluctuations.

1.2. THE FORMATION OF STRUCTURE

The regions of the density field, ρ, with slightly increased density relative to the background mean density, $\bar{\rho}$, seen in the CMB (the density contrast, $\delta = (\rho - \bar{\rho})/\bar{\rho} \approx 10^{-5}$), will attract their surrounding slightly more than average, increasing their density. While matter flows away from slightly under-dense regions making them even more under-dense. At early times, these perturbations remain in the linear regime as long as perturbations remain relatively small such that $\delta \ll 1$.

1.2.1 Linear regime

Given that the growth of these perturbations is driven by massive, non-relativistic particles a Newtonian treatment is sufficient (however, for a more rigorous derivation utilising

general relativity see Lifshitz, 1946). With this in hand we can model this evolution assuming the equations that govern ideal, non-relativistic, self-gravitating, baryonic fluids.

We begin with the mass conservation equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \left(\rho \mathbf{u}\right) = 0, \tag{1.2}$$

and the Euler equation,

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{1}{\rho} \nabla P - \nabla \Phi, \tag{1.3}$$

where t is time, $\mathbf{u}$ is the velocity field, P is the pressure, ρ is the density of the fluid and Φ is the gravitational potential. The mass conservation equation states that without sources or sinks, the rate of change of mass within a volume is given by the rate of flow of mass through a surface area. The Euler equation says that the acceleration of our massive, non-relativistic particles is governed by the pressure gradient and the gravitational potential.

The gravitational potential is described by the Poisson equation,

$$\nabla^2 \Phi = 4\pi G \rho, \tag{1.4}$$

where G is the gravitational constant. We complete the set of fluid equations by including the barotropic equation of state $P = P(\rho)$. We move these equation to comoving coordinates, $\mathbf{x}$, related by $\mathbf{r} = a\mathbf{x}$, where a is the scale factor, and after linearly perturbing these equations, we take the gradient of Equation 1.3 and substitute into the results of Equations 1.2 & 1.4 resulting in a second order differential equation. By Fourier transforming this equation, with $\delta(\mathbf{x}, t) = \sum_{\mathbf{k}} \delta_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}}$, gives

$$\ddot{\delta}_{\mathbf{k}} + 2H\dot{\delta}_{\mathbf{k}} + \left(\frac{c_s^2}{a^2} k^2 - 4\pi G \bar{\rho}\right) \delta_{\mathbf{k}} = 0, \tag{1.5}$$

where $c_s = \sqrt{\partial P / \partial \rho}$ is the adiabatic sound speed, k is the wavenumber and $H = \dot{a}/a$ is the Hubble parameter. This equation is analogous to that of a damped harmonic oscillator, with a frequency

$$\omega = \sqrt{\frac{c_s^2}{a^2} k^2 - 4\pi G \bar{\rho}}, \tag{1.6}$$

where damping is provided by the expansion of the Universe.

The Jeans wavenumber, in proper units,

$$k_{\mathrm{J}} \geq \sqrt{\frac{4\pi G \bar{\rho}}{c_s^2}}, \tag{1.7}$$

represents that transition from $\omega^2 > 0$, where perturbations oscillate back and forth, and pressure overcomes gravity, and $\omega^2 < 0$, where gravity overcomes pressure and gravitational collapse is induced. This occurs as the wavelength of these perturbations,

$$\lambda > \lambda_{\mathrm{J}} = \frac{2\pi}{k_{\mathrm{J}}}, \tag{1.8}$$

where λ_{J} is the Jeans length, and the system becomes unstable. In this regime the wavenumber term in Equation 1.6 becomes negligible and the solutions become $\omega = \pm i\sqrt{4\pi G\bar{\rho}}$ and hence the overdense regions collapse, referred to as the Jeans instability (Jeans, 1902).

1.2.2 Non-linear regime

As linear collapse continues, the density of these overdense regions grows until $\delta \sim 1$, at which point the overdensity starts to evolve non-linearly. Analytical descriptions of this phase are not possible, instead simulations are required. However, simple analytical models, e.g. the spherical collapse model introduced by (Gunn & Gott, 1972), can give an insight into this evolution.

The spherical collapse model assumes an spherical region with uniform density. As the Universe expands, so does the spherical region, however gravitational forces means that this is at a slower rate than the background. As such, the density grows with time. At some time, t_{turn}, the radius of the spherical region stops growing and instead begins to collapse. This non-linear collapse, according to the spherical region model, is expected to begin at $\delta = 1.686$. Accordingly, the spherical region collapses indefinitely, forming a point mass.

However, this neglects the fact that the gravitational collapse of these perturbations are not expected to be spherical, nor are they expected to have uniform density. As the matter distribution is non-symmetric, as matter infalls it is impacted by a varying gravitational potential. These variations act to redistribute the orbital energies of this matter, through a process called violent relaxation (Lynden-Bell, 1967), stopping infall onto a point mass and virialising (a point at which the kinetic energy of the system balances the gravitational potential energy) the system at some time, t_{vir}.

We can also calculate the virial mass,

$$M_{\mathrm{vir}} = \frac{4}{3}\pi R_{\mathrm{vir}}^3 \Delta_{\mathrm{vir}} \bar{\rho}, \tag{1.9}$$

where $\Delta_{\rm vir}$ is the density contrast and $R_{\rm vir}$ is the virial radius of the virialised structure. This density contrast has been shown to be approximately $\Delta_{\rm vir} \equiv \delta(t_{\rm vir}) \sim 200$ (Bryan & Norman, 1998). By setting the density contrast one can estimate the mass of the collapsed perturbation from its virial radius or visa versa.

Press & Schechter (1974) modelled the linear and non-linear regimes of the evolution of these perturbations to estimate the number of structures with a given mass. They find a power law at low masses, with an exponential cut-off at high mass that increases with time. While this modelling requires many simplifications and assumptions, it does somewhat accurately reproduce the shape of the galaxy luminosity function (Schechter, 1976).

1.2.3 Dark matter

The current favoured model of the Universe is the Λ cold dark matter (ΛCDM) model (Peebles, 1982; Blumenthal et al., 1984), however, until now we have largely ignored the effects of "dark matter". Observations of the motions of stars in galaxies, galaxies within galaxy clusters and galaxy clusters themselves (Zwicky, 1933; Roberts & Rots, 1973; Einasto et al., 1974; Ostriker et al., 1974; Rubin et al., 1978, 1980), led to the discovery of matter that appears not to interact with electromagnetic radiation, the so-called "dunkle materie" (Zwicky, 1933) or dark matter. This along with the discovery that, using supernovae as standard candles, the Universe is undergoing accelerated expansion (Riess et al., 1998; Schmidt et al., 1998; Perlmutter et al., 1999), implies the presence of dark energy, making up the two main constituents (26% and 69% of the mass and energy of the Universe, respectively; Planck Collaboration et al., 2020) of the Universe within the ΛCDM model. These two components contribute the 'DM' (dark matter) and Λ (dark energy, which acts as a cosmological constant, Λ) to ΛCDM. The 'C' refers to the fact the dark matter particles are "cold", a theory proposed by Peebles (1982).

While Equations 1.7 & 1.8 do not hold for a collisionless fluid, like dark matter, the Jeans length of dark matter can be found to be (Jeans, 1902),

$$\lambda_{DM} = \sigma \sqrt{\frac{\pi}{G\bar{\rho}}}\,, \tag{1.10}$$

where σ is the velocity dispersion of the dark matter particles. This velocity dispersion can wash out perturbations on scales smaller than λ_{DM}, a process called free streaming. Therefore, in a CDM dominated Universe, this sets the lower mass limit of structure formation. If, instead DM is relativistic, i.e. hot, the lower mass limit grows to be the mass of galaxy clusters (Bond et al., 1980), enforcing a "top-down" growth model. This would be

at odds with observational evidence that individual galaxies form first then merge to form larger galaxies, and galaxy groups, clusters and superclusters, in turn – the hierarchical structure growth.

Dark matter decouples from the radiation field at a much earlier time than baryonic matter, becoming non-relativistic as it does so, owing to its lack of interaction with electromagnetic radiation. As a result, dark matter forms perturbations long before recombination occurs. Baryons account for just a sixth of the matter content of the Universe, and as such, after recombination, the baryons fall into the dark matter potential wells that are already in place.

1.2.4 First galaxies

In order to form a galaxy, gravitational collapse needs to occur down to scales significantly smaller than that of the virial radius discussed in Section 1.2.2. As discussed earlier, gas will infall into dark matter halos. As it passes the virial radius it is shock heated to the virial temperature, $T_{\rm vir}$, defined as,

$$k_{\rm B} T_{\rm vir} = \frac{G m M_{\rm vir}}{2 R_{\rm vir}}, \tag{1.11}$$

where $k_{\rm B}$ is the Boltzmann constant and m is the particle mass. One can assume a mean particle mass, for a fully ionised gas with hydrogen and helium at primordial abundances (as discussed in Section 1.1) of $m = 0.59 m_{\rm p}$, where m_p is the mass of a proton. Taking the Milky Way as an example, with a dark matter halo mass of $\sim 10^{12}$ $\mathrm{M}_\odot$ and virial radius of $R_{\rm vir} \sim 200$ kpc (Battaglia et al., 2005), the virial temperature is of order 10^6 K. As the baryonic matter reaches thermal equilibrium further gravitational collapse is halted.

The above statement assumes that the baryonic matter, like dark matter, is unable to lose energy via dissipative processes. However, while dark matter does not interact with electromagnetic radiation, baryonic matter does. As such, the baryonic matter that infalls into the gravitational potential produced by the virialised dark matter halos is able to radiate away some of its gravitational potential energy. As it does so, it reduces the radius of the cloud of baryonic matter within the dark matter halo, allowing collapse onto much smaller scales.

There are three key timescales for the cooling of gas in dark matter halos: the cooling time of the gas, $t_{\rm cool}$, the free-fall time of the galaxy $t_{\rm ff}$ and the age of the Universe, $t_{\rm H}$ (Silk, 1977; Rees & Ostriker, 1977; Binney, 1977; White & Rees, 1978). When we compare these timescales there are three regimes of cooling:

- $t_{\rm cool} > t_{\rm H}$: the shock heated gas remains hot and in hydrostatic equilibrium, collapse onto the core is not possible.

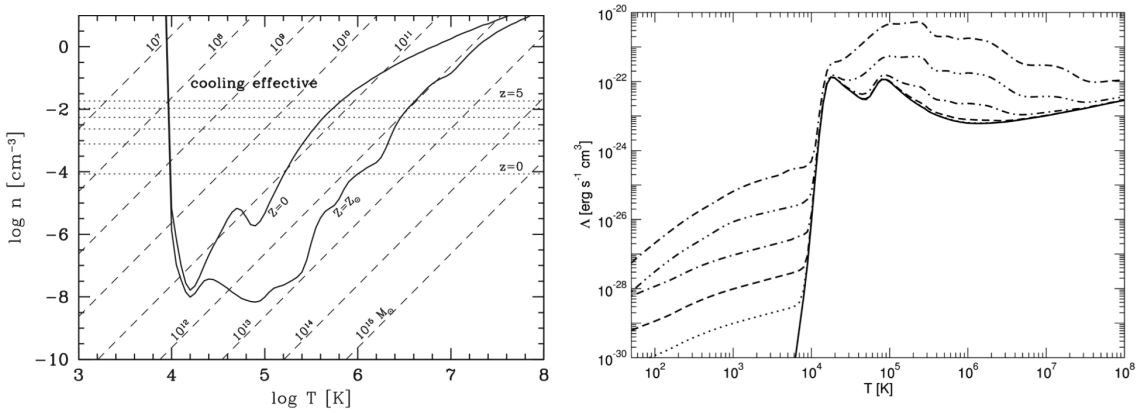


Fig. 1.2 (Left:) A cooling diagram, taken from (Mo et al., 2010), illustrating the boundary of where cooling becomes effective for varying temperatures and densities of the gas, for two different metallicities – zero and solar, the upper and lower curves, respectively. The diagonal dashed lines indicate the density and temperature of constant gas masses (in $M_\odot$) and the horizontal dashed lines show expected densities of virialised halos at different redshifts. (Right:) The atomic cooling rate as a function of gas temperature and metallicity (different dashed lines), taken from (Smith et al., 2008). The upper curve represents gas with metallicty of 10 $Z_\odot$, moving down the following curves are for 1 $Z_\odot$, 0.1 $Z_\odot$, 0.01 $Z_\odot$, 0.001 $Z_\odot$ and 0 (solid line) $Z_\odot$. Figure credit: Mo et al. (2010)

- $t_{\rm H} > t_{\rm cool} > t_{\rm ff}$: the accreted gas is able to cool over long timescales.
- $t_{\rm cool} < t_{\rm ff}$: the gas can rapidly cool, pressure no longer supports the gas and collapse onto the core can be triggered.

The cooling timescale can be defined as,

$$t_{\rm cool} \equiv \frac{\varepsilon}{|{\rm d}\varepsilon/{\rm dt}|} \propto \frac{nT}{\Lambda(T)}, \tag{1.12}$$

where ε is the gas internal energy per unit volume, n is the gas number density, T is its temperature and $\Lambda(T)$ is the cooling function.

The efficiency of cooling is determined by the temperature, density and metallicity of atomic gas, as well as the molecular cooling rate. For the formation of the very first galaxies, we can set the metallicity to be zero. At $T_{\rm vir} \gtrsim 10^7$K, the gas is entirely collisionally ionised and hence the key cooling process is bremsstrahlung emission from free electrons. However, this is not an efficient cooling mechanism at high temperatures ($\Lambda \propto T^{0.5}$; Rybicki & Lightman, 1979), ultimately, $t_{\rm cool} \propto T^{1/2}$. The temperature range $10^4 \lesssim T \lesssim 10^6$K collisional excitation proves to be the most efficient form of cooling. At these high temperatures, collisions between free electrons and atoms can excited a bound electron

to a higher energy level, as this electron relaxes to its ground state, it emits a photon, allowing the gas to radiate away its gravitational potential energy. The cooling efficiency within this temperature range can be increased further when metals are present in the gas, as shown in Figure 1.2. The first peak in the metal-free cooling function, at $T \approx 2 \times 10^4$K, corresponds to the temperature at which collisional excitation and ionisation are most effective, while the second peak corresponds to the equivalent for helium. Above these respective temperatures hydrogen and helium are ionised and the cooling function decreases until bremsstrahlung emission becomes the dominant cooling mechanism. The presence of metals acts to increase the cooling efficiency via collisional excitation – with increasing atomic mass comes increasing line transitions that allow energy to be radiated away.

At gas temperatures below $T < 10^4$K the gas is entirely neutral and hence the above cooling mechanisms that largely require plentiful free electrons fail to cool efficiently. As seen in Figure 1.2, the atomic cooling rate drops rapidly at $T < 10^4$K [1]. However, in this temperature regime, the small molecular component of primordial gas becomes the dominant coolant (Peebles & Dicke, 1968; Benson et al., 2006; Wise & Abel, 2008). This cooling occurs through collisional excitation of their rotational and vibrational transitions and the cooling rates of H_2 and others molecules can be found in Galli & Palla (1998). The presence of metals act to enhance both collisional excitation and molecular cooling, and hence the cooling function is significantly greater at these cooler temperatures for metal enriched gas than for the primordial gas that the first galaxies will form from.

The final key process in the infall of gas onto dark matter halos is the effects of photoionisation by the ultraviolet background (UVB). As the first stars form, they emit photons with energies great enough to ionise hydrogen ($h\nu > 13.6$eV). As these photons leak into surrounding star-forming regions, they ionise the neutral hydrogen required for collisional excitation, suppressing the first peak in the cooling function. However, for the earliest galaxies, where very few stars have formed and the neutral (intergalactic medium) IGM absorbs the transmission of the UVB, this affect is less important. The combination of photoionisation and decreasing efficiency of cooling at temperatures below 10^4K enforce a lower limit ($M \sim 10^6 M_\odot$ Bromm, 2013) on the dark matter halo masses that galaxies can form within.

These mechanisms cool gas that has fallen into the dark matter halos present in the early Universe. As they do so, the gas emits photons in an isotropic matter meaning the infalling gas retains its initial angular momentum. This angular momentum is inherited

[1] Even with the presence of metals, some of which have low-energy transitions, the cooling rate declines significantly.

from the dark matter halo that aquires rotation thanks to tidal forces exerted on it by neighbouring structures (Peebles, 1969). The fraction of specific angular momentum inherited by the gas is referred to as the retained fraction of specific angular momentum. If this fraction is too low, galaxies would fail to reproduce the observed size distribution of local spiral galaxies (e.g. Steinmetz & Navarro, 1999), which is referred to as the angular momentum catastrophe, and is instead close to unity (Mo et al., 1998; Fall & Efstathiou, 1980; Peebles, 1969; Governato et al., 2007).

1.2.5 First stars

The formation of these first stars and galaxies in the Universe marks the end of the dark ages. The Universe enters a period where it is gradually illuminated by the radiation released from these galaxies as they form and evolve, marking the epoch called the "cosmic dawn".

Within the first primordial gaseous disks, dense, cool regions are able to overcome their pressure support and collapse into protostellar cores. This occurs in the event that the mass enclosed in the region surpasses the Jeans mass. The Jeans mass, $M_{\rm J}$, can be defined as the mass enclosed within a sphere of diameter equal to the Jeans length, ($\lambda_{\rm J}$, Equation 1.8):

$$M_{\rm J} = \frac{4}{3}\pi\rho\left(\frac{\lambda_{\rm J}}{2}\right)^3 = \frac{\pi^{5/2}c_{\rm s}^3}{6G^{3/2}\rho^{1/2}} \propto \left(\frac{T^3}{\rho}\right)^{1/2}, \tag{1.13}$$

given that the sound speed, $c_{\rm s} \propto T^{1/2}$ (given that $P \propto \rho T$), where T and ρ are the temperature and density of the gas, respectively. Therefore, with increasing density and decreasing temperature, the Jeans mass reduces rapidly until it is surpassed by the mass of the region and runaway collapse is induced.

However, for primordial gas, the cooling mechanisms at $T < 100$K are inefficient, as such the collapse can be approximated as being adiabatic (pV^γ = constant i.e. $p\rho^{-\gamma}$ = constant, where γ is the adiabatic index; $\gamma = 5/3$ for a monatomic gas[2]). Combining this with the fact that $p \propto \rho T$, it is clear that the temperature is a function of ρ, $T \propto \rho^{\gamma-1}$. As such the Jeans mass is proportional $\rho^{1/2}$, and as the collapse continues and density increases, so does the Jeans mass, stopping fragmentation. As the density rises, so does the temperature, and as it surpasses $T \sim 100$ K, molecular hydrogen cooling rates increase and $M_{\rm J}$ returns to the relation shown in Equation 1.13 allowing collapse. This occurs when masses are of order 10^2 $M_\odot$ (Haiman et al., 1996; Tan & McKee, 2004), forming a single massive protostar, in a "monolithic" process. The result is the formation of the very

[2]Molecular hydrogen makes up a negligible proportion of the gaseous disk.

first population of stars, these are metal-free and massive and are titled Population III (Pop. III) stars (for a review see Schneider et al., 2006). Recent works have attempted to find diagnostics of the presence of Pop. III stars (Trussler et al., 2023; Nakajima & Maiolino, 2022), such as strong hydrogen and helium emission lines and very low inferred metallicities. Moreover, Vanzella et al. (2023); Maiolino et al. (2024b); Wang et al. (2024a) have claimed to identify galaxies that satisfy some of these diagnostics, and hence may host Pop. III stars.

As discussed earlier, the presence of metals allows for more efficient cooling. Therefore, metal enriched gas is able to effectively remain isothermal in its collapse. As such, $M_{\rm J} \propto \rho^{-1/2}$, and initial perturbations in the gas cloud can pass the Jeans mass resulting in multiple collapses into less massive protostars, in a process called "fragmentation" (McKee & Ostriker, 2007).

Both of these avenues of star formation lead to the formation of a gaseous disk which accretes onto the protostar as collisions cause the gas to lose angular momentum. Both the increasing density and cooling processes facilitate further gravitational contraction. Eventually the pressure within the protostar becomes great enough to ignite nuclear fusion. This fusion drives a radiation pressure that counteracts the gravitational forces and leaves the new star in hydrostatic equilibrium.

The exact masses of the stars that form out of these processes dictate their future evolution. The luminosity of a star, L, is strongly dependent on its mass, M: $L \propto M^{\alpha}$, where $\alpha \sim 4$ can be assumed for main sequence stars (Kuiper, 1938). As such the mass of a star dictates the rate at which the star radiates, however a star can only convert a fixed fraction of its mass into radiation, the energy release from the fusion of hydrogen to helium is 0.7% of the rest mass, and as such, the time a star spends on the main sequence, $t_{\rm MS}$, is governed primarily by its mass, i.e.

$$t_{\rm MS} \propto M/L \propto M^{-3}. \tag{1.14}$$

Therefore, with increasing mass, a star utilises the fuel available for nuclear reactions at a faster rate and hence spends considerably less time on the main sequence. For a star like our Sun $t_{\rm MS} \sim 10^{10}$ yr, but as we head towards the regime of the most massive, metal-free stars $t_{\rm MS} \lesssim 10^{7}$ yr (Schaller et al., 1992).

The mass of a star also dictates its evolutionary pathway once it moves off the main sequence. The least massive stars ($M \lesssim 2{\rm M}_{\odot}$) have main sequence lifetimes comparable to the Hubble time, $t_{\rm H}$, making them somewhat irrelevant for the chemical enrichment of galaxies. As they exhaust their hydrogen supply, the star contracts under gravity, however the temperature remains too low to ignite helium burning. Instead, the core is

supported by the electron degeneracy pressure, so long as the core mass remains below $1.4M_{\odot}$ (Chandrasekhar, 1931). Fusion in the outer hydrogen shell continues to increase the mass of the helium core until the temperature in the core is great enough to ignite helium burning and the "helium flash" occurs (a thermal runaway fusion reaction as a result of the degenerate state of the core; Mestel, 1952). Their evolution continues in a similar manner to intermediate mass stars ($2\mathrm{M}_{\odot} < \mathrm{M} < 8\mathrm{M}_{\odot}$) after this, where these stars upon contraction will be massive and hence hot enough ($\sim 10^8$K) to ignite helium burning without the subsequent steps that produce the "helium flash". Helium burning within these low and intermediate mass stellar cores produces carbon through the triple-α process. As some of these carbon nuclei capture additional helium nuclei oxygen is also produced. Throughout this process the core's temperature remains too low to ignite carbon burning and hence, when the remaining helium is synthesised into carbon and oxygen, fusion is halted, leaving a carbon and oxygen core supported again by the electron degeneracy pressure – a white dwarf.

The early evolutionary phases of massive stars ($M > 8\mathrm{M}_{\odot}$) are the same, however, their masses and hence temperatures mean they are able to fuse heavier elements; that is until an iron and nickel core is produced. Any further nuclear reactions do not release energy as the binding energy of all nuclei heavier than iron is lower than iron itself, and hence the core is no longer supported against gravity by radiation pressure and the outer shells collapse inwards. As they do so a shock wave ignites further nuclear reactions in the shells producing a core-collapse supernova (type-II SN), leaving behind a neutron star or black hole, depending on whether the Tolman–Oppenheimer–Volkoff limit (Tolman, 1939; Oppenheimer & Volkoff, 1939) of $M_{\mathrm{core}} = 2.0 - 2.2\mathrm{M}_{\odot}$ is passed (Rezzolla et al., 2018). The stellar lifetime is of order ~ 100 Myr allowing efficient enrichment of the interstellar medium (ISM) with heavier elements than hydrogen and helium. Furthermore, these SN can compress gas in the ISM that can induce further star formation (Elmegreen & Lada, 1977; Walborn et al., 2002).

Given that the mass of a star drives both its potential to enrich the ISM with metals, but also the timescales over which it lives, constraining the distribution of stellar masses within a galaxy is clearly of great importance (Maiolino & Mannucci, 2019).

1.2.6 Initial mass function

The relative distribution of masses with which stars are born, the initial mass function (IMF), $\phi(\mathrm{m})$, where $m = M/\mathrm{M}_{\odot}$, must be understood in order to fundamentally understand galaxy formation and evolution. In order to measure the IMF one must measure the absolute magnitude of individual stars, and hence their distance must be known and we

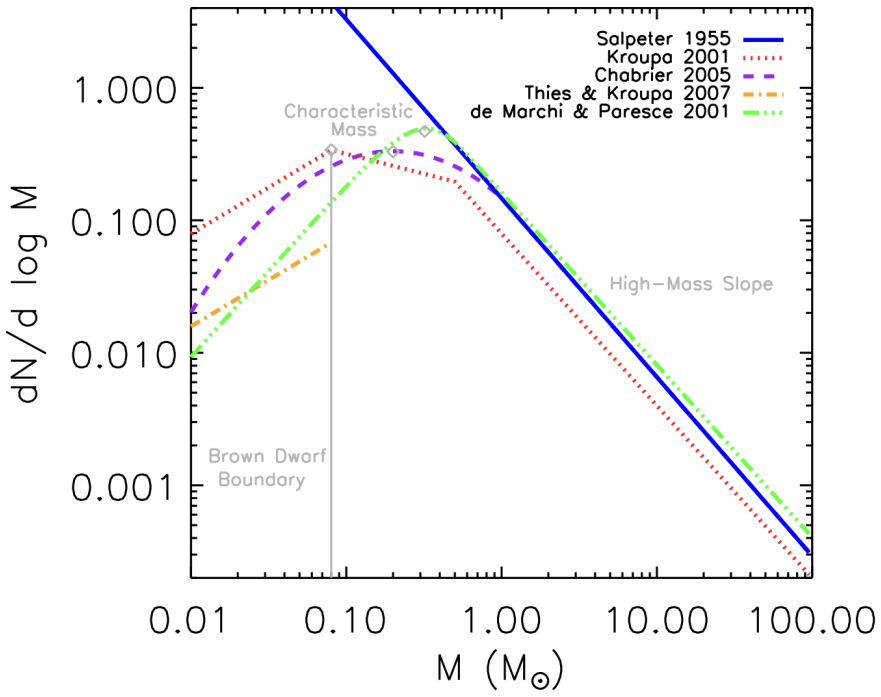


Fig. 1.3 The number of stars per stellar mass bin – the IMF – based on fits to Galactic stellar data (Salpeter, 1955; Kroupa, 2001; Chabrier, 2005; Thies & Kroupa, 2007; de Marchi & Paresce, 2001). The figure indicates the varying characteristic masses and high-mass slopes of these IMFs. Figure credit: Offner et al. (2014)

must also be able to resolve each star. This means that the only direct measurements of the IMF that are possible are from either the Milky Way (Salpeter, 1955; Kroupa, 2001; Chabrier, 2003) or nearby galaxies (Mateo, 1988; Elson et al., 1989; Hunter et al., 1995), however, the applicability of this IMF to galaxies in the high-redshift Universe can be considered dubious (Cameron et al., 2024). Observations of the solar neighbourhood suggest that $\phi(m)$ can be well described by a power law,

$$\phi(m) = dN/dm \propto m^{-b}, \tag{1.15}$$

where N is the number of stars within the mass range dm and $b = 2.35$ (Salpeter, 1955). However, further works have suggested that b varies with stellar mass (Kroupa, 2001; Chabrier, 2003), as illustrated in Figure 1.3.

What is clear from the Salpeter (1955) IMF is that it is "bottom-heavy" i.e. the low-mass stars ($M < 1\mathrm{M}_{\odot}$) dominate the total stellar mass,

$$M_{\mathrm{tot}} = \int_{M_{\mathrm{lower}}}^{M_{\mathrm{upper}}} M\phi(m)\,dm = \left[\frac{M^{2-b}}{2-b}\right]_{M_{\mathrm{lower}}}^{M_{\mathrm{upper}}} \approx \frac{M_{\mathrm{lower}}^{-0.35}}{0.35}, \tag{1.16}$$

where we assume the upper and lower mass limits differ by orders of magnitudes[3]. Conversely, the Kroupa (2001) and Chabrier (2003) IMFs show a turnover at some characteristic mass that make them "top-heavy" relative to the Salpeter (1955) IMF. There has been long-standing debate regarding the universality of the IMF; its slope and characteristic mass (Larson, 1985). This becomes especially problematic when studying high-redshift galaxies, who evolve in a vastly different Universe, are relatively metal-poor and have likely recently hosted (or currently do host) Pop. III stars, at significant odds with the Milky Way that is used to derive the Salpeter (1955), Kroupa (2001) and Chabrier (2003) IMFs. While observational evidence for top-heavy IMFs (Cameron et al., 2024) and alternative top-heavy IMF models for Pop. III stars have been suggested (Larson, 1998; Zackrisson et al., 2011; Klessen & Glover, 2023), the weak observational constraints on the characteristic mass and slope that should be assumed mean these are rarely implemented in modelling of high-redshift galaxies. However, when IMF variations are included in semi-analytical modelling they can begin to reproduce the observed properties of high-redshift galaxies (Trinca et al., 2024).

As these stellar populations form and evolve they produce ionising photons. Some of these photons escape from the galaxy without being absorbed and ionise the IGM, thus beginning a new phase in the Universe's history – the Epoch of Reionisation (EoR).

1.2.7 Early galaxy evolution: Dwarf galaxies to overdense structures

The power spectrum of the CMB shows significantly more small-scale fluctuations than large-scale, resulting in the formation of more low mass than high mass galaxies – the hierarchical growth model. Within this model it is the low mass galaxies that form first in the early Universe. Numerical simulations of galaxy formation show that in these low mass galaxies, supernovae feedback may be able to remove most of the gas and heavy elements before they are able to form more stars, resulting in the low brightness and metallicity of dwarf galaxies (Larson, 1974; Navarro et al., 1996; Governato et al., 2010). Slightly more massive perturbations produce galaxies that have large enough stellar masses and hence gravitational potentials to retain their gas and metals and become star-forming spiral

[3] M_{lower} is typically of order $0.1M_{\odot}$, while M_{upper} is typically of order $100M_{\odot}$

galaxies (e.g. Silk, 1977; Rees & Ostriker, 1977; White & Rees, 1978; Frenk et al., 1985; Hopkins et al., 2023). However, observations also find massive, elliptical galaxies (Hubble, 1936) whose formation pathways need to be explained. Toomre (1977) used numerical simulations to produce some of the first disc-galaxy merging models, where spheroidal galaxies are the remnants of disk galaxy merger events. Further simulation work shows that these merger events can result in elliptical galaxies with density profiles that are in agreement with observations (e.g. White & Rees, 1978; Springel et al., 2005), in agreement with the hierarchical formation model. While these models can explain basic galaxy properties, one of the most significant conundrums is in addressing the properties of the rarest galaxies with masses above $\sim 10^{12} M_{\odot}$. Semi-analytical models that include AGN feedback have shown that this significant decrease in the number of the most massive galaxies (seen in the decrease in star-formation efficiency at high halo masses – Zheng et al. 2007) can be explained by quenching via AGN activity (Bower et al., 2006; Croton et al., 2006).

The power spectrum of the CMB shows that density perturbations are present across wide angular scales. While the most overdense perturbations will result in galaxy formation, wider scale perturbations will result in loose clusters of dark matter halos hosting galaxies. The gravitational attraction between these dark matter halos results in a slower than average expansion of this region with cosmic time, further increasing the overdensity, until, like we discussed in Section 1.2.2 these become a large-scale virialised structure. The most overdense structures in the early Universe that go on to form the most massive clusters of galaxies observed at current day (e.g. the Coma cluster; Zwicky, 1933) are suitably entitled protoclusters and are seen to be in place by at least $z \sim 8$ (Laporte et al., 2022; Hashimoto et al., 2023; Morishita et al., 2023) and possibly as early as $z = 10.6$ (Scholtz et al., 2024) (see Section 4). By this cosmic epoch the central cores of these protoclusters (with scales of 10's kpc) are expected to dominate the star formation of the protocluster undergoing an inside-out growth phase, while protoclusters themselves may contribute up to 50% of the cosmic star-formation rate density (Chiang et al., 2017). These overdense large structures, hosting an order of magnitude more galaxies than comparable regions of the Universe, are therefore thought to be important agents in the reionisation process.

1.3. MODELS OF REIONISATION

After recombination the Universe is neutral. As cosmic dawn commences, the first stars and galaxies form in the early Universe, emitting Lyman-continuum (LyC) photons (those with energies $h\nu \geq 13.6$eV) which are capable of ionising neutral hydrogen. These photons

gradually ionise the neutral hydrogen present in the ISM before they then leak into the IGM starting the ionisation of the neutral hydrogen present in the early Universe, over the Epoch of Reionisation (EoR)[4]. As LyC photons ionise neutral hydrogen they get absorbed ($\lambda_{\rm rest} = 0.09113\mu$m). This absorption by neutral hydrogen also occurs for the resonant Lyman emission lines, most notably the $n = 2$ to $n = 1$ Lyman-α (Lyα) line found at $\lambda_{\rm rest} = 0.1216\mu$m, (discussed in more detail below). The absorption from neutral hydrogen in the IGM is characterised by the optical depth,

$$\tau_{\rm GP} = 2 \times 10^5 x_{\rm HI}(z) \frac{(1+z)^{3/2}}{4}, \tag{1.17}$$

where $x_{\rm HI}(z) = n_{\rm HI}/n_{\rm H}$ is the fraction of hydrogen at a given redshift that is neutral (Gunn & Peterson, 1965). As such, at high redshift, where the Universe is thought to be neutral after recombination, continuum emission at $\lambda_{\rm rest} \leq 0.1216\mu$m is eventually redshifted onto the rest-frame wavelength of neutral hydrogen and is effectively absorbed[5]. The non-detection of continuum emission at $\lambda \lesssim 0.1216\mu$m in the early Universe (Becker et al., 2001) confirms the high density of neutral hydrogen in the IGM. Conversely, the detection of LyC photons at lower redshifts confirms that the neutral hydrogen in the Universe has undergone a process of ionisation(Gunn & Peterson, 1965).

This process of ionisation can be characterised by a relatively simple model (Madau et al., 1999):

$$\frac{{\rm d}Q_{\rm HII}}{{\rm d}t}[{\rm s}^{-1}] = \frac{\dot{n}_{ion}[{\rm cm}^{-3}\ {\rm s}^{-1}]}{\langle n_{\rm H}\rangle\,[{\rm cm}^{-3}]} - \frac{Q_{\rm HII}}{\langle t_{\rm rec}\rangle\,[{\rm s}]}. \tag{1.18}$$

The evolution of the volume-filling factor of ionised hydrogen, $Q_{\rm HII}$, is governed by a "source" and "sink" term. The ionisation of hydrogen by the ionising photons produced by early stars, galaxies and AGN contribute the "source" term – the ratio of the rate of ionising photon production, $\dot{n}_{ion}$, to the neutral hydrogen density at a given cosmic time, $n_{\rm H}$. The recombination of hydrogen, the "sink" term, is governed by the amount of ionised hydrogen present in the Universe and the volume-averaged time before ionised hydrogen recombines, $t_{\rm rec}$. When the rate of photoionisations exceeds that of recombinations the fraction of the Universe that is ionised grows.

The recombination time can be given by,

$$t_{\rm rec}[{\rm s}] = \left[(1 + Y/4X)\,\langle n_{\rm H}[{\rm cm}^{-3}]\rangle\,\alpha(T)[{\rm cm}^3\ {\rm s}^{-1}]{\rm C}\left(1+z\right)^3)\right]^{-1}, \tag{1.19}$$

[4]The EoR refers exclusively to hydrogen ionisation, however, helium does also get ionised later in the Universe's history.

[5]Continuum emission at $\lambda_{\rm rest} \ll 0.1216\mu$m can be absorbed by lower wavelength Lyman transitions or at $\lambda_{\rm rest} < 0.09113\mu$m all photons can be absorbed by the $n = \infty$ to $n = 1$ transition.

as shown in Madau et al. (1999), where $C \equiv \langle n_{\rm HII}^2 \rangle / \langle n_{\rm HII} \rangle^2$ is the clumping factor of ionised hydrogen, where $n_{\rm HII}$ is the ionised hydrogen density, Y and X are the helium and hydrogen abundance, respectively, and $\alpha(T)$ is the recombination coefficient, which, assuming Case A recombination and a temperature of 20,000K, as suggested by Kaurov & Gnedin (2014)[6] is $\alpha = 2.52 \times 10^{-13} {\rm cm}^3\ {\rm s}^{-1}$ (Osterbrock & Ferland, 2006).

The rate of ionising photon production from galaxies can be approximated as,

$$\dot{n}_{ion}(z)[{\rm cm}^{-3}\ {\rm s}^{-1}] = f_{\rm esc}(z)\rho_{\rm UV}(z)[{\rm erg\ cm}^{-3}\ {\rm Hz}^{-1}\ {\rm s}^{-1}]\xi_{ion}(z)[{\rm Hz\ erg}^{-1}], \tag{1.20}$$

where the $f_{\rm esc}$ is the average LyC escape fraction (the fraction of LyC photons escaping from a galaxy into the IGM), $\rho_{\rm UV}$ is the UV luminosity density and ξ_{ion} is the average LyC photon production efficiency (the ratio of LyC photons produced per UV photon) (Robertson, 2022). While the UV luminosity density can be measured by integrating the UV luminosity function (e.g. Bouwens et al., 2022) over all luminosities, the LyC escape fraction and LyC photon production efficiency are not directly observable in the EoR given that LyC photons are absorbed by the neutral IGM (Madau, 1995; Inoue et al., 2014). While the UV luminosity density is well constrained, $f_{\rm esc}$ and ξ_{ion} may change the impact of individual galaxies towards reionisation, resulting in a significant debate about the properties and hence relative contributions of different populations towards reionisation (e.g. Madau et al., 1999; Finkelstein et al., 2019; Robertson, 2022; Simmonds et al., 2024b).

The exact timings of *when* this process began and ended are important in both a cosmological context, but also in understanding *which* astrophysical sources were responsible for it. Whether reionisation occurred in a democratic or oligarchical structure will also impact *how* these early ionised regions grew.

1.3.1 When did reionisation occur?

In order to probe the absorption effects on galaxies in the EoR, spectroscopic observations are required. As such, in the era of ground-based spectroscopy, observations had to focus on the brightest sources at high redshifts for which continuum measurements are possible, i.e. quasars and gamma-ray bursts. These observations suggest that more than 90% of the IGM was ionised by $z \sim 5.5-6$ (Fan et al., 2006; McGreer et al., 2015; Boera et al., 2019). With increasing observations of high-redshift sources and improving observational facilities, further probes of the neutral fraction of the IGM have become available. The Lyα forest opacity currently offers the most robust constraint on the end of reionisation

[6] Changing from Case A to Case B recombinations, at such high temperatures, has negligible impact on the recombination coefficient. These different regimes of recombination are discussed further in Section 1.3.2

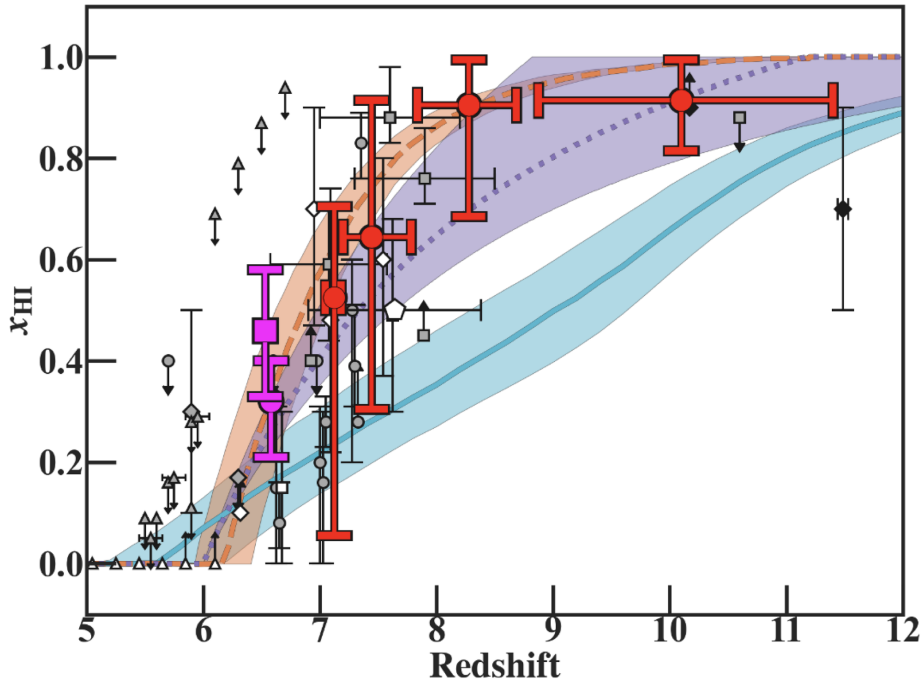


Fig. 1.4 The redshift evolution of $x_{\rm HI}$ inferred from a variety of probes. Three models are presented: blue solid (Finkelstein et al., 2019), orange dashed (Naidu et al., 2020) and purple dotted (Ishigaki et al., 2018) curves. The red circles and black diamonds are measurements of Lyman-α (Lyα) damping wing absorption from (Umeda et al., 2023) and (Curtis-Lake et al., 2023; Hsiao et al., 2023), respectively. The grey circles and white squares are estimated from the Lyα luminosity function (Ouchi et al., 2010; Konno et al., 2014, 2018; Inoue et al., 2018; Goto et al., 2021; Morales et al., 2021) and Lyα-emitter (LAE) clustering analysis (Ouchi et al., 2018), respectively. The white and grey triangles are measured using quasar Gunn-Peterson troughs (Fan et al., 2006) and the Lyα +Lyβ dark fraction/gaps (Mesinger et al., 2015; Zhu et al., 2022; Jin et al., 2023), respectively. The white and grey diamonds show the measurements using quasar (Schroeder et al., 2013; Davies et al., 2018; Greig et al., 2019; Wang et al., 2020) and GRB (Totani et al., 2006, 2014) Lyα damping wing absorption, respectively. The grey squares shows the estimate from Lyα emission from Lyman break galaxies (Mesinger et al., 2015; Hoag et al., 2019; Mason et al., 2019; Jung et al., 2020; Whitler et al., 2020; Bruton et al., 2023; Morishita et al., 2023). Finally, the white pentagon is an estimate from CMB observations (Planck Collaboration et al., 2020). Figure credit: Umeda et al. (2023).

($z \sim 5.3$) (Kulkarni et al., 2019; Keating et al., 2020; Bosman et al., 2022). The mean flux in the Lyα forest is dictated by the optical depth, which in turn gives the neutral hydrogen fraction. However, at relatively low neutral hydrogen fractions ($x_{\rm HI} \sim 10^{-5}$) the transmitted flux will be negligible, as such, this method is poor at constraining anything more than the end of the EoR. Studying the red wing of individual Gunn-Peterson (Gunn & Peterson, 1965) troughs around high-redshift galaxies offers a method of constraining the neutral hydrogen fraction in the early Universe (Miralda-Escudé, 1998; Mortlock et al., 2011; Umeda et al., 2023; Keating et al., 2024). Increasing neutral hydrogen density increases the absorption at the wings of the profile. As such, a decreasing fraction of the emitted Lyα photons are able to pass through the IGM, making the equivalent width (EW) of Lyα another measure of the neutral hydrogen fraction. While the fraction of Lyman-α-emitters (LAEs) increases up to $z \sim 6$, thought to be linked to their decreasing dust content, the fraction rapidly drops at $z > 6$, evidencing the increasing neutral hydrogen in the IGM (Stark et al., 2011; Pentericci et al., 2011; Schenker et al., 2014). However, as can be seen from Figure 1.4, even these improved constraints fail to discern between an early and late start of reionisation.

1.3.2 Which astrophysical sources drove reionisation?

The sources that drove reionisation are heavily debated. Three parameters drive the contribution of individual galaxies towards reionisation (as shown in Equation 1.20): $f_{\rm esc}$, ξ_{ion} and $L_{\rm UV}$, the LyC escape fraction, the ionising photon production efficiency and the UV luminosity, respectively. While UV-bright galaxies will intrinsically produce the most LyC photons, the most UV-bright galaxies are often the most evolved and enriched with metals and dust (Cullen et al., 2024; Austin et al., 2024), leading to the absorption of some LyC photons. Moreover, based on the hierarchical growth of structure in the Universe, these most massive galaxies will be much rarer than their UV-faint counterparts. This interplay of galaxy properties with UV brightness has led to significant debate over whether the most UV-bright galaxies dominate reionisation, in an "oligarchical" process (Naidu et al., 2020), or whether it is driven by the abundant UV-faint galaxies in a "democratic" process (Simmonds et al., 2024b).

The UV luminosity function (LF) of galaxies is constrained by counting the number of galaxies in a given UV luminosity bin and volume, at a given redshift. Under the assumption of constant star formation efficiency, the LF will trace the halo mass function. In reality, star formation efficiency is not constant, and where the UV luminosity deviates from the halo mass function gives insight into the feedback processes at play (Van Den Bosch et al., 2003; Vale & Ostriker, 2004; Behroozi et al., 2013). The LF can typically be

characterised by the Schechter function (Schechter, 1976),

$$\Phi(L)\mathrm{d}L = \Phi^* \left(\frac{L}{L^*}\right)^{\alpha} \exp\left(\frac{L}{L^*}\right) \mathrm{d}\left(\frac{L}{L^*}\right), \tag{1.21}$$

which is a combination of a power law, with a faint end slope α, which at characteristic luminosity, L^*, changes to an exponential regime, and is normalised by the normalisation density, Φ^*. Bouwens et al. (2006) quantified the evolution of the LF while fitting all three parameters, finding that the characteristic luminosity is reduced from $z \sim 3$ to $z \sim 6$, suggesting a higher fraction of UV-faint galaxies in the early Universe. However, with the introduction of the Wide Field Camera 3 on the *Hubble* Space Telescope (*HST*) came a significant increase in the number of photometric candidate galaxies out to $z \sim 8$ (from ~ 20 to around $200-500$ galaxies at $z \sim 7-8$ identified via the Lyman-break method, discussed in more detail in Section 1.4) facilitating wider studies of the LF that actually suggested that the normalisation density and the faint end slope evolve with redshift (Ouchi et al., 2009; Castellano et al., 2010; Bouwens et al., 2011; Bradley et al., 2012; McLure et al., 2013). A detailed analysis of $\sim 10,000$ galaxies from $z \sim 4$ to $z \sim 10$ suggested a steepening of the faint-end slope and decreasing normalisation density, that can be seen in Figure 1.5, removes the need for an evolution in the characteristic luminosity (Bouwens et al., 2015).

However, with these increasing numbers of galaxies identified at $z > 8$, especially for most UV-luminous galaxies (Bowler et al., 2014, 2020) and eventually, with the advent of the *James Webb* Space Telescope (*JWST*) and its sensitive Near Infrared Imaging Camera (NIRCam) (e.g. Donnan et al., 2023) and Spectrograph (NIRSpec) (e.g. Harikane et al., 2024), new constraints on the LF at these highest redshifts become available. This unveiled that the LF at these high redshifts is likely no longer characterised by the Schechter function, but instead a double power-law, suggesting that the feedback mechanisms that inhibit star formation in the most massive galaxies is yet to impact on galaxies at these redshifts (Bowler et al., 2014). This is notable as it implies more UV-bright galaxies in the early Universe than had originally been expected. However, regardless of the exact shape of the LF, it is clear that past the characteristic luminosity comes a more rapid decline in UV-bright galaxies than occurs for UV-faint galaxies approaching the characteristic luminosity. By utilising these LF constraints it has been shown that one can reconstruct the reionisation process with a model that is dominated by the ionising photons produced by the most UV-bright, massive galaxies (Naidu et al., 2020) (an "oligarchicical" process). However, alternative works suggest that the quantity of extremely UV-faint galaxies is consistent with the amount necessary for a more "democratic" reionisation process.

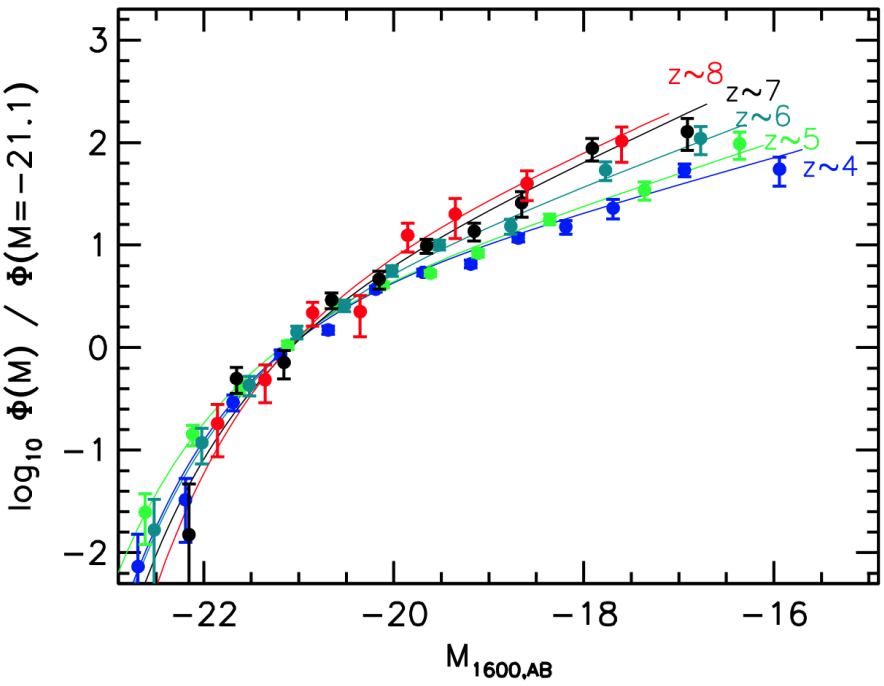


Fig. 1.5 The evolution of the UV luminosity function with redshift. The red, black, dark cyan, green and blue curves and data points represent $z \sim 8$, $z \sim 7$, $z \sim 6$, $z \sim 5$ and $z \sim 4$, respectively. This suggests an evolution in the slope of the LF with redshift. Figure credit: Bouwens et al. (2015).

Given that the LF of high-redshift galaxies fails to constrain the reionisation process, it is clear that understanding the evolution of $f_{\rm esc}$ and ξ_{ion} with both redshift and galaxy properties is crucial. Direct observations of hydrogen ionising photons in the EoR are impossible and hence the escape fraction of LyC photons, $f_{\rm esc}$, has to be estimated using calibrations built on either simulations (e.g. Choustikov et al., 2024) or low-redshift observations (Mascia et al., 2023). At lower redshifts, where LyC photons can be observed (Leitet et al., 2013; Borthakur et al., 2014), samples of LyC emitting galaxies (LyC leakers) have been collated (e.g. Izotov et al., 2016, 2018b; Flury et al., 2022). $f_{\rm esc}$ is seen to scale with a variety of parameters – the ratio of the [O III]/[O II] nebular emission lines, the UV slope β and the absolute UV magnitude M_{UV}. An alternative probe of LyC escape is Lyα escape; in the presence of neutral hydrogen LyC photons will be absorbed, while Lyα will be scattered. While Lyα is often ultimately absorbed, either by infinite scattering, or by dust, the escape fraction of Lyα photons, $f_{\rm esc}$(Lyα), can be seen as an upper bound of the LyC escape fraction. However, where significant Lyα photons are able to escape, it appears likely that LyC photons may also be able to escape. As such the study of LAEs allows for further understanding of LyC escape. Both Lyα emission and known $f_{\rm esc}$ diagnostics can be used to show that UV-bright galaxies have $f_{\rm esc} < 10\%$ (discussed in detail in Section 2) which is significantly below what is expected for an "oligarchical" process (Naidu et al., 2020).

If UV-bright galaxies are dominating reionisation, they must have very high production efficiencies of LyC photons. While original works utilised indirect diagnostics of ξ_{ion}, such as the UV slope β (e.g. Robertson et al., 2013), the advent of the *Spitzer* space telescope facilitated photometric observations of the rest-optical spectrum, notably the Hα emission line, of galaxies at $z \sim 4-5$ (e.g. Schaerer & de Barros, 2009). Measurements on the Hα emission line can be used to estimate the LyC production assuming a recombination model. Case B recombination is assumed to be a fair approximation for galaxies in the early Universe[7], where the gas is optically thick to Lyman line emission (Menzel & Baker, 1937; Baker & Menzel, 1938). Under the assumption of optically thick gas, direct recombinations lead to the release of a hydrogen ionising photon, which immediately ionises a surrounding neutral hydrogen atom, with the net result of no recombination. Therefore, we can assume that instead, the hydrogen atom recombines indirectly to an excited state ($n \geq 2$). Under this Case B recombination, assuming a temperature of 10^4 K, the amount of ionising photons is directly related to the dust-corrected luminosity of Hα,

[7]This approximation may break down in certain high-redshift galaxies which have experienced an extreme quenching event disrupting their dense gas producing density bounded nebulae (McClymont et al., 2024b)

(Osterbrock & Ferland, 2006),

$$N(H^0)[\mathrm{Hz}] = 7.28 \times 10^{11}[\mathrm{erg}^{-1}]L(\mathrm{H}\alpha)[\mathrm{erg\ s}^{-1}]. \tag{1.22}$$

With $N(H^0)$ in hand one can measure $\xi_{ion,0} = N(H^0)/L_{UV}$. Results with *Spitzer* suggest that ξ_{ion} does not evolve with UV magnitude (Bouwens et al., 2016). The sensitivity and wavelength range of *JWST* NIRSpec and medium-band filters of NIRCam has facilitated rest-optical photometry and spectroscopy, and hence measurements of the Hα emission line and in turn ξ_{ion} (e.g. Simmonds et al., 2024b; Saxena et al., 2024; Jung et al., 2024; Tang et al., 2023), which suggests that ξ_{ion} increases with decreasing UV magnitude. The results appear to suggest that burstiness of star formation in UV-faint galaxies (discussed in more detail in Section 1.4) drives this increasing production efficiency of LyC photons. However, this method of constraining ξ_{ion} utilising the Hα emission lines forces pre-selection of galaxies that show Hα emission in their spectral energy distribution (SED), thus neglecting galaxies that show no nebular emission, i.e. (mini-) quenched galaxies (e.g. Looser et al., 2024) or remnant leakers (galaxies that are likely density bounded, having depleted themselves of star-forming gas; Katz et al., 2023b). Alternatively, using spectral energy distribution (SED) fitting to derive ξ_{ion} (which has been shown to be consistent with measurements from Hα; Simmonds et al., 2024b) for all photometrically identified galaxies, regardless of whether emission lines are observed in said photometry, results in a flattening of this ξ_{ion} and M_{UV} relation (Simmonds et al., 2024a). Neither of these results show any rapid increase in ξ_{ion} with increasing UV magnitude and as such, the current paradigm of the reionisation process is that UV-faint galaxies dominate in a "democratic" process.

1.3.3 What is the morphology of reionisation?

The current paradigm of reionisation states that early LyC emitting galaxies formed the first ionised bubbles. These bubbles were however relatively small (Neyer et al., 2024), and recombination combated their growth. However, as galaxies evolve with time, their stellar populations increase, and as a result so does their intrinsic number of ionising photons. Moreover, as galaxy clusters and overdensities form, an abundance of galaxies can, in combination, drive large ionised bubbles. The EW distribution of LAEs suggest that the reionisation process is patchy (Pentericci et al., 2014), moreover, evidence from *JWST* suggests that LAEs are biased to regions of galaxy overdensities (Witstok et al., 2024c; Tang et al., 2023; Chen et al., 2024).

While complex Lyα radiation transfer codes exist (e.g. Verhamme et al., 2006; Duval et al., 2014), relatively simple models, where an intrinsic Lyα profile is assumed and passed through an IGM attenuation curve that begins attenuating at a wavelength corresponding to the size of the ionised bubble, can be utilised to predict the profile of the escaping Lyα emission (Mason & Gronke, 2020). Therefore, when the profile of the observed Lyα emission is known, we can reverse this model to estimate the ionised bubble required to produce the observed emission. Two key properties are required: the velocity offset (relative to the systemic redshift) and escape fraction of the Lyα emission. While ground-based spectroscopy has provided observations of the Lyα emission line at high redshifts (e.g. Zitrin et al., 2015), without other emission lines and hence a systemic redshift, the velocity offset cannot be accurately measured. Assuming Case B recombination and a temperature of 10^4 K, dust corrected Balmer emission line fluxes ($f(\mathrm{H}\alpha)$ and $f(\mathrm{H}\beta)$) can be used to estimate the intrinsic Lyα flux (Osterbrock & Ferland, 2006), and hence the escape fraction,

$$f_{\mathrm{esc}}(Ly\alpha) \equiv \frac{f(\mathrm{Ly}\alpha)}{f(\mathrm{Ly}\alpha_{\mathrm{int}})} = \frac{f(\mathrm{Ly}\alpha)}{8.7 \times f(\mathrm{H}\alpha)} = \frac{f(\mathrm{Ly}\alpha)}{24.8 \times f(\mathrm{H}\beta)}. \tag{1.23}$$

The spectral range and sensitivity of *JWST* facilitate the observation of strong UV and optical emission lines and constraints on the Balmer series, therefore allowing both the velocity offset and escape fraction to be measured and in turn an estimate of the required ionised bubble size, R_{req}, that allows the Lyα emission to escape. Conversely, the size of the ionised bubble that can feasibly be produced by an individual galaxy, assuming it has a constant ionising photon production rate of $N(H^0)$ for some time t_* with a constant LyC escape fraction, f_{esc}, is given by,

$$R_{\mathrm{pred}} \approx \left(\frac{3 f_{\mathrm{esc}} N(H^0)[\mathrm{Hz}]\, t_*[\mathrm{s}]}{4\pi \bar{n}_{\mathrm{H}}(z)[\mathrm{cm}^{-3}]} \right)^{1/3}, \tag{1.24}$$

where $\bar{n}_{\mathrm{H}}(z)$ is the mean hydrogen number density at a given redshift. By comparing the implied ionised bubble size of LAEs to the bubble size that they can individually produce, examples of galaxies that cannot self-ionise the bubble within which they must reside have been identified ($R_{\mathrm{pred}} < R_{\mathrm{req}}$; Saxena et al., 2023; Witstok et al., 2024c,a). The implication therefore is that instead the ionised bubbles, within which these galaxies reside, must have been driven by an abundance of galaxies. It is however notable that LyC escape is not necessarily isotropic and hence significant line-of-sight affects are expected to impact the observed Lyα emission (Zackrisson et al., 2017), and hence impacting bubble size estimates.

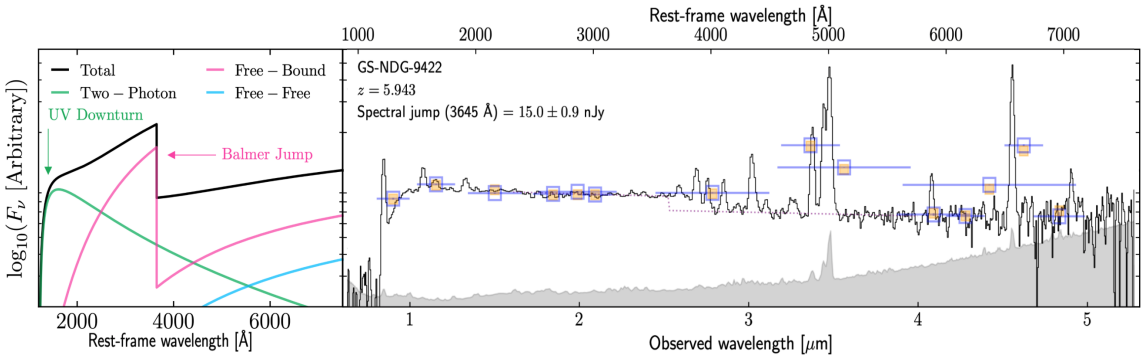


Fig. 1.6 (Left:) The three components of nebular continuum emission: two-photon (green), free-bound (pink) and free-free (blue). The total nebular emission in this model is indicated in black. The UV downturn and Balmer jump, the two key diagnostics of nebular continuum emission, are annotated. (Right:) The NIRSpec prism spectrum of a $z = 5.943$ galaxy that shows a strong Lyman break ($\lambda_{\rm rest} < 1215$ Å), Balmer jump ($\lambda_{\rm rest} \sim 3645$ Å), nebular emission lines and recombination lines. This galaxy also show a significant UV downturn at wavelengths longer than the Lyman break. This has been claimed to be either be a damped Lyα absorber (Tacchella et al., 2024) or two-photon emission (Cameron et al., 2024). The NIRCam photometry is indicated by orange squares. This figure is adapted from Katz et al. (2024) and Cameron et al. (2024).

While the driving forces behind reionisation appear to be relatively well understood the exact timescales and topology of reionisation remains uncertain.

1.4. Galaxy spectral properties in the early Universe

Galaxies that exist in the early Universe have characteristic spectral properties that are defined by the stars and gas that they host and the environment within which they exist. The most notable features include the Lyman-break, the Balmer break, the UV-continuum slope and strong nebular emission lines.

1.4.1 The Lyman-break

Absorption of photons with energies greater than the Lyα photon by neutral hydrogen in the IGM produces the Lyman-break (seen in Figure 1.6). By $z \sim 6$ this results in a complete break in the spectrum of all galaxies (Madau, 1995; Inoue et al., 2014). In the pre-*JWST* era, where high sensitivity continuum spectroscopy of high-redshift galaxies was not possible, the primary identification of high-redshift galaxies was utilising Lyman-break detections in ground and space-based photometric instruments (e.g. the Visible and Infrared Survey

Telescope for Astronomy (VISTA) and *HST*). Interlopers exist that can mimic high redshift Lyman-breaks in broad-band photometry (e.g. Balmer-breaks at lower redshifts and brown dwarf stars), however, the selection of high-redshift targets with the *HST* has been found to have been accurate by follow-up *JWST* observations (Bunker et al., 2023b). This method was often supported by exploiting a color-color selection which helps to remove dusty low-redshift interlopers. The number of robust $z \gtrsim 7$ galaxies identified in deep *HST* imaging totalled more than 700 (Bouwens et al., 2015). The sensitivity, wavelength range and medium-band filters of *JWST* photometry has facilitated the robust detection of UV-faint galaxies significantly expanding the number of photometric candidates. Moreover, *JWST* spectroscopy has shown that many spectroscopically confirmed galaxies have slightly lower spectroscopic redshift than implied from the Lyman-break observed with photometry, often driven by damped Lyα absorbers (DLAs; Heintz et al., 2024c,b; Hainline et al., 2024). This excess damping of the Lyman-break is caused by extremely high neutral hydrogen column densities ($N_{\rm HI} > 10^{20}\,{\rm cm}^{-2}$) in the circumgalactic medium (CGM) or ISM, and has been seen to be linked with overdense environments (Chen et al., 2024; Heintz et al., 2024a).

The strength of the continuum before attenuation by neutral gas, the UV luminosity, is dictated by the number of stars that have ages below 100 Myrs and as such can be used as a tracer of the star formation over the last 100 Myrs (Kennicutt, 1998). We expect that galaxies at higher redshifts have had less time to accrete gas and drive significant star formation and hence they become more UV-faint. In the *HST*-era there was significant debate as to whether the cosmic SFR density, inferred from the UV-luminosity function, evolved smoothly or rapidly at $z > 8$ (e.g. McLeod et al., 2016; Oesch et al., 2018). However, the relatively shallow depths and small surface area of these observations left significant uncertainty in these conclusions. *JWST* has facilitated significant increases in the area of deep imaging and hence vast improvements in UV-luminosity functions at high redshifts and hence the cosmic SFR density (e.g. Donnan et al., 2023). This has shown that the cosmic SFR density evolves more smoothly at $z > 8$ and appears to have an excess of UV-bright at $z > 12$ (e.g. Harikane et al., 2024). While there are a number of potential solutions to explain this excess (e.g. increased star formation efficiency, AGN activity, top-heavy IMFs, large scatters in the halo mass-SFR relation and cosmic variance; Harikane et al. 2024), which of these solutions are indeed responsible remains highly debated.

1.4.2 The UV-slope

An additional constraint on the photometric selection of high-redshift galaxies is the color redward of the Lyman-break. High-redshift galaxies have intrinsically blue colors after the

Lyman-break, which is a key discriminator with low-redshift red interlopers. This blue UV-slope is largely a result of stellar emission.

Stars can be modelled as a superposition of blackbody spectra - the emission from an object in thermodynamic equilibrium with its environment. This thermal radiation from a blackbody is governed only by its temperature, T (Planck, 1901),

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5}\frac{1}{e^{hc/\lambda k_B T} - 1}, \tag{1.25}$$

where B_λ is the spectral radiance in units of erg s^{-1} cm^{-2} Å^{-1} sr^{-1} at a given wavelength, λ, h is the Planck constant and c is the speed of light. The maximum of this function gives Wien's displacement law (Wien, 1896),

$$\lambda_{\rm max} T = 2.898 \text{ [m K]}, \tag{1.26}$$

where $\lambda_{\rm max}$ is the wavelength that corresponds to the peak of the spectral radiance. For a typical O-type star with a temperature of 30,000K, this corresponds to a peak wavelength of $\lambda_{\rm max} = 97\mu$m. As stellar populations age, lower temperature stars dominate the SED and as such the UV-slope becomes redder. However, at stellar ages of less than 10 Myr, the intense stellar ionising radiation output is capable of ionising significant amounts of the nebular gas around the stellar population. The relative effect of the nebular emission from this gas on the galaxy SED is to reduce the strength of the UV-slope (see Figure 1.6 from values of $\beta = -3$ to $\beta = -2$ (Bouwens et al., 2010).

Moreover, assuming dust attenuation curves at high-redshift follow those seen at lower redshifts, e.g. Calzetti et al. (2000); Gordon et al. (2003), the attenuation in the rest-UV far exceeds that in the rest-optical driving increasingly red colours with dust (Salim & Narayanan, 2020).

A combination of dust and nebular emission driving a reddening of the intrinsic UV slope of typical stellar populations enforces a relatively strict cut off of $\beta \gtrsim -2.6$ (Cullen et al., 2024). The wavelength range of *JWST* NIRCam filters combined with its sensitivity enables significantly stronger constraints on the UV slope for a large population of high-redshift galaxies at $z \gtrsim 10$ that were unseen by *HST*. Studies of the UV-slope at high redshifts show a decreasing UV-slope with increasing redshift, as expected from the decreasing time to build up significant dust masses (Cullen et al., 2024; Austin et al., 2024). However, the UV-slope has occasionally been seen fall to values of $\beta < -2.6$ at $z \gtrsim 10$, potentially indicating a period of time where either exotic stellar populations exist or $f_{\rm esc}$ tends to unity, significantly reducing the nebular emission from these galaxies, producing an extremely blue UV-slope (e.g. Witstok et al., 2024b).

1.4.3 Nebular emission

While photometry can provide strong constraints on the presence of a Lyman-break and blue UV-slope colours, in turn providing an estimate of the redshift, this estimate is uncertain. Alternatively the production of nebular emission lines at specific, known rest-frame wavelengths allows for robust spectroscopic confirmation of high-redshift galaxies. As discussed briefly earlier, stellar radiation can ionise the gas that surrounds it, the resulting free electrons as they interact and recombine produce characteristic continuum and line emission.

The majority of nebular emission lines seen in galaxies are collisionally-excited lines, where excitation to higher levels are caused by collisions with free electrons. The energy levels that electrons can be promoted to have excitation potentials of order kT, where T is the temperature of the gas. In order to understand the emissivity of individual transitions, one can consider the electron transition between energy level 1, E_1, and 2, E_2, (with number densities of n_1 and n_2), producing a photon with energy $h\nu_{21} = E_2 - E_1$, where ν_{21} is its frequency. These collisional excitations occur at a rate of $n_0 n_1 C_{12}$, while de-excitations have a rate of $n_0 n_2 C_{21}$ (where the number density $n_0 \approx n_e$, the electron density) and the transition coefficients are given by,

$$\frac{C_{21}}{C_{12}} = \frac{g_1}{g_2} e^{-h\nu_{21}/kT}, \tag{1.27}$$

where g is the statistical weight of each energy level.

The emissivity of the emission line is given by $\varepsilon = h\nu_{21} n_2 A_{21}$, where A_{21} is the Einstein coefficient which describes the probability of spontaneous emission. This ultimately means that for a steady state, collisional excitation is balanced by spontaneous and collisional de-excitations, $n_0 n_1 C_{12} = n_2(A_{21} + n_0 C_{21})$. Ultimately this leads to an emissivity

$$\varepsilon = \frac{h\nu_{21} n_0 n_1 C_{12}}{1 + \frac{n_0 C_{21}}{A_{21}}}, \tag{1.28}$$

where below the critical density, $n_{0,\mathrm{crit}} = A_{21}/C_{21}$, the emissivity $\varepsilon \propto n^2$ (given that generally $n_0, n_1, n_2 \propto n$, where n is the total gas density). However, when $n_0 \gg n_{0,\mathrm{crit}}$, $\varepsilon \propto n$ and the line begins to be collisionally de-excited.

The emissivity of these lines are therefore functions of electron temperature and density. Emission line doublets, such as [O II]λ3727,3730Å, have similar excitation energies and hence the ratio of the doublet is governed by their different critical densities (Maiolino & Mannucci, 2019). These are therefore used as electron density diagnostics (e.g. Kewley

et al., 2019). Alternatively, emission lines resulting from high quantum levels, "auroral" lines, are very sensitive on electron temperature. Auroral lines such as [O III]λ4363Å can be used to estimate the electron temperature, which in turn can be used as a probe of the gas-phase metallicity (Sanders et al., 2016; Curti et al., 2017). These in turn can be used to create calibrations between this electron temperature based metallicity and strong emission line ratios and facilitate gas-phase metallicity estimates while auroral line emission are not detected (Pettini & Pagel, 2004). Alternatively, when transitions from multiple ionisation stages of the same element are observed, e.g. [O II] and [O III], this allows one to understand the occupancy of each energy level and hence is a probe of the ionisation parameter (McGaugh, 1991), U, which is generally defined to be proportional to the ratio of the ionising photon density to the hydrogen density.

Notably, promotion of an electron to the first excited levels of hydrogen or helium requires much greater energies than for heavier elements, and as such higher temperatures are required to produce collisionally excited emission lines. Instead, the emission lines seen from hydrogen and helium are recombination lines, produced when hydrogen and helium which were photo-ionised recombine. The relative strengths of all hydrogen recombination emission lines can be computed, assuming Case B recombination and a given electron density and temperature (Osterbrock & Ferland, 2006).

The most frequent hydrogen recombination transition is the $2p \rightarrow 1s$ transition, producing the intrinsically strongest hydrogen recombination line, Lyman-α at $\lambda_{\rm rest} = 1215.6$ Å. The Lyα line however is a resonant line, as such the presence of neutral hydrogen in the ISM, CGM and IGM in the early Universe makes observations of the Lyα line challenging, and observations of all other Lyman series lines impossible, as discussed in Section 1.3. Instead the strongest observed hydrogen recombination lines are the Balmer lines, Hα and Hβ, the ratio of which is the Balmer decrement. Comparing the known intrinsic ratio to that observed is frequently used as a measure of the dust attenuation around young stellar populations (e.g. Berman & Trumpler, 1936).

Hydrogen recombination can also lead to nebular continuum emission. While ordinarily, the continuum spectra of galaxies is dominated by stellar emission, stellar populations with high ionising photon production efficiencies (discussed in Section 1.3.2) may be able to produce sufficient ionising photons to facilitate the observation of nebular continuum emission. This nebular continuum emission has three main components (Osterbrock & Ferland, 2006; Reines et al., 2010; Byler et al., 2017): free-bound, free-free and two-photon emission (as shown in Figure 1.6). Free-bound emission, occurs when free electrons recombine with ionised hydrogen, producing recombination emission lines as it cascades down the energy levels. Recombinations become more likely with decreasing electron

temperature, however, at the ionisation energy, 3645Å for the Balmer series, there is a sharp discontinuity where there is no excess kinetic energy of the electron to produce free-bound emission. The spectral discontinuity associated with the Balmer series, the Balmer jump, has been identified in low-redshift highly star forming galaxies (Peimbert & Costero, 1969; Guseva et al., 2006, 2007). Free-free (bremsstrahlung) emission, is produced by the deceleration of electrons as they are deflected by ions. This is generally sub-dominant to the free-bound emission that occurs at similar wavelengths (see Figure 1.6). Two-photon emission is also generally expected to be sub-dominant to stellar UV emission. Two-photon emission facilitates the decay of an electron between two energy levels with the same angular momentum quantum number. The most notable is the $2s \rightarrow 1s$ transition of neutral hydrogen, which produces a characteristic peak at $\lambda_{\rm rest} \approx 1430$Å, smoothly declining at lower and higher wavelengths, predicted by Schaerer (2002); Raiter et al. (2010); Zackrisson et al. (2011) and potentially observed at $z = 3.357$ in the Lynx arc (Fosbury et al., 2003). However, the wavelength range of this two-photon emission is comparable to that of the stellar UV continuum that has a steep increasing slope with decreasing wavelength. In order to produce observable nebular emission young stellar populations ($\lesssim$ 10Myrs) driving significant production of ionising photons are required, however these in turn have steep UV slopes that outshine the nebular continuum emission, therefore two-photon emission is generally expected to be sub-dominant.

The sensitivity and wavelength range of NIRSpec prism observations has, for the first time, facilitated rest-optical continuum observations of high-redshift galaxies. This has revealed that detections of Balmer jumps at high redshift become typical at $z > 6$ (Roberts-Borsani et al., 2024). Moreover, both spectroscopy and photometry indicate that the majority of galaxies at this epoch show extreme nebular emission line strengths (Boyett et al., 2024). These features represent the extreme nebular emission driven by the young ionising stellar populations present in the early Universe. A sub-sample of such galaxies have been shown to also host a UV downturn that is consistent with two-photon emission (see Figure 1.6; Cameron et al., 2024; Katz et al., 2024).

1.4.4 The Balmer break

The Balmer break is the inverse of the nebular continuum Balmer jump discussed in Section 1.4.3. Occurring at $\lambda_{\rm rest} = 3645$Å, the Balmer break is associated with bound-free absorption of photons in stellar atmospheres, and as such the strength of the break is governed by stellar physics. The strength of the Balmer break has long been known as a measure of stellar age and is strongest in A-type stars, reaching a maximum at 300-1000 Myr (see Figure 5 from Kriek et al., 2006). As such, for high-redshift galaxies, whose stellar

ages cannot exceed the ages that correspond to the maximum age of the Universe, the Balmer break is a key tool in diagnosing stellar ages.

With spectroscopic confirmation and photometric constraints on the UV-slope, optical continuum (e.g. Balmer break or Balmer jump strengths) and nebular emission lines, SED-fitting tools can begin to estimate the star-formation histories of galaxies (e.g. Carnall et al., 2018; Johnson et al., 2021) by matching models of stellar and nebular emission spectra to observations (e.g. Eldridge et al., 2017). The nebular emission and Balmer break strength from a galaxy constrains both its recent (on scales of 10 Myrs) and extended star formation, respectively. When these are measured and combined with the detection or lack-there-of dust in its UV-slope one can infer the galaxy's SFH.

The photometric identification of strong Balmer breaks in high-redshift galaxies ($z \gtrsim 9$) was used to suggest they host old stellar populations with stellar mass weighted ages of ~ 300 Myr (Hashimoto et al., 2018; Roberts-Borsani et al., 2020; Laporte et al., 2021). SED-fitting of spectroscopic observations of galaxies that host Balmer breaks and/or nebular emission lines at high redshifts offers the opportunity to understand the SFH of early galaxies, and to constrain the timescales and frequency of the bursty SFHs that they are predicted to have (Fukuda et al., 2000; Immeli et al., 2004; Bournaud et al., 2007; Elmegreen et al., 2009; Faucher-Giguère, 2018; Tacchella et al., 2020).

1.5. GALAXY PHOTOMETRIC PROPERTIES IN THE EARLY UNIVERSE

All of the features discussed in Section 1.4 give high-redshift galaxies distinctive SEDs. While NIRSpec prism spectroscopy gives a vast improvement in spectroscopic depth relative to all other spectrographs, NIRCam imaging offers a more efficient method for obtaining high-redshift galaxy SEDs. Imaging can provide constraints on properties such as the stellar mass and star formation rate (often with significantly larger uncertainties than for spectroscopy) for large samples of galaxies. However, the high sensitivity and spatial resolution ($0.03'' - 0.06''$) of NIRCam makes it particularly useful in identifying the shapes, sizes and environments of galaxies.

While photometric redshifts are inevitably somewhat uncertain, they can be utilised in a probabilistic manor to assess the likelihood that galaxies that appear local in the plane of the sky are also local in the line of sight. This can be used to estimate the probability a galaxy is found with a close companion galaxy, i.e. the pair fraction. Studies have shown that this increases with redshift, up to $\sim 20\%$ by $z \sim 8$, then remains nearly constant (Puskás et al. in prep.; Duan et al., 2024). These photometric redshifts can also

be used to identify large-scale galaxy overdensities, such as Laporte et al. (2022); Tacchella et al. (2023b), and study the properties of the resident galaxies. Galaxies resident within overdensities appear to have higher specific SFR (the ratio of star-formation rate to stellar mass) as well as redder UV slopes, however show no differences in their effective radii (Li et al., 2024). Galaxy sizes are however seen to exponentially decrease with redshift from a maximum of $R_{\rm eff} \sim 4.5$ kpc at $z = 0$, to a plateau of $R_{\rm eff} \sim 1$ kpc at $z > 5$ (Ormerod et al., 2023; Morishita et al., 2024).

1.6. The James Webb Space Telescope

First proposed in 1996, the then Next Generation Space Telescope was intended to launch in 2007 with a budget of $1 billion. As the successor to the *HST*, the aim was to build a much larger and cooler telescope in order to observe the formation of early galaxies. A plan to build a 6.5-metre class telescope, orbiting at Lagrange point 2 was settled on, but multiple scheduling problems and funding challenges led to a total delay of 14 years and a ballooning of the budget to ~ $10 billion. However, on the 25$^{\rm th}$ December 2021 the Ariane 5 rocket launched from French Guiana carrying onboard the *James Webb* Space Telescope. The orbital location, sunshield and cryocooler onboard significantly reduce the background emission relative to telescopes like *HST* allowing deeper observations at longer wavelengths. This allows *JWST* to host the Mid-Infrared Instrument (MIRI), as well as three other scientific instruments: NIRCam, NIRSpec and the Near-Infrared Imager and Slitless Spectrograph (NIRISS). Moreover, the wider diameter (2.7 times wider) and larger collecting area (6.6 times larger) than *HST* significantly increases the sensitivity and spatial resolution of *JWST*, respectively.

The NIRCam instrument consists of two $2.'2 \times 2.'2$ modules that can each simultaneously image in short ($0.6-2.3\mu$m) and long ($2.4-5\mu$m) wavelengths. NIRCam also employs a pupil wheel and filter wheel, that can be used in various combinations facilitating grism, coronographic and imaging observations with extra-wide-, wide-, medium- and narrow-band filters. These filters span the wavelength range $0.6-5\mu$m. The medium band filters (corresponding to a spectral resolution of $R \sim 10$) are effective at differentiation between nebular and continuum emission (Endsley et al., 2023; Laporte et al., 2023), especially in the rest-optical for $z \sim 5-10$ galaxies and hence are crucial in improving the accuracy of photometric redshift estimates.

The NIRSpec instrument provides spectroscopy over a $0.6-5.3\mu$m wavelength range utilising four different observing modes: multi-object, integral field, fixed slit and wide aperture spectroscopy. Of these, multi-object spectroscopy with the micro-shutter array

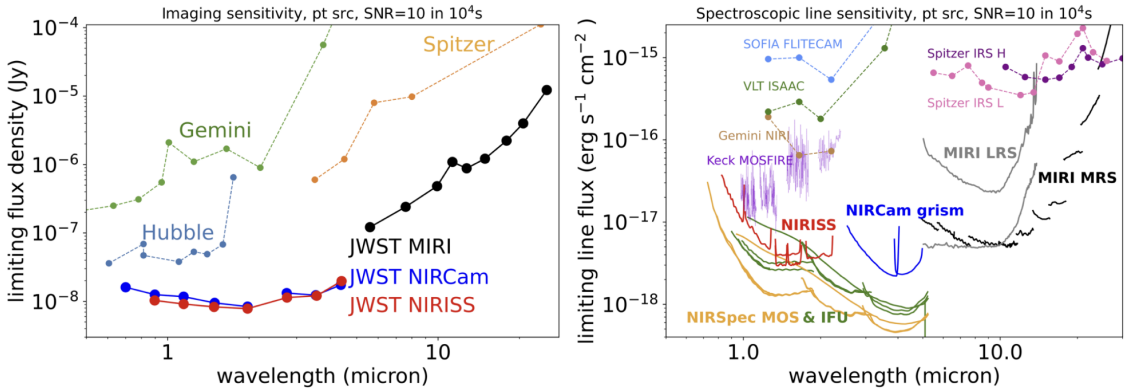


Fig. 1.7 The limiting flux density and line flux detected at 10σ for a point source in 10^4s using (Left:) NIRCam, NIRISS and MIRI imaging and (Right:) NIRSpec, NIRISS, NIRCam grism and MIRI spectroscopy, respectively. The sensitivity of these *JWST* instruments is compared to existing ground- and space-based telescopes. At $0.6\mu\text{m} \lesssim \lambda \lesssim 20\mu\text{m}$ *JWST* instruments offer significant increases in sensitivity to existing telescopes underlining the advancement in imaging and spectroscopy offered by *JWST*. This figure is adapted from Rigby et al. (2023).

(MSA) and spatially resolved spectroscopy with the integral field unit (IFU) are the most utilised for high-redshift science. The NIRSpec MSA configuration facilitates spectroscopic observations of galaxies, with $0.2'' \times 0.46''$ area shutters, covering a total area of $3.'6 \times 3.'4$, while the IFU provides spatially resolved (each pixel is $0.1'' \times 0.1''$) spectroscopy over a $3'' \times 3''$ area. NIRSpec utilises a filter wheel and grating wheel to select the transmission and resolution of the spectrograph, respectively. The transmission filters typically utilised are four long-pass filters and one bandpass filter, while there are seven dispersers, six of which provide high ($R \sim 2,700$) and medium ($R \sim 1,000$) resolution gratings for the long-pass filters and one low resolution ($R \sim 100$), double-pass prism for the bandpass filter. These grating-filter configurations provide a variety of resolutions, wavelength ranges and sensitivities (as seen in Figure 1.7), and hence are utilised for different studies from rest-UV or -optical emission line characterisation to rest-UV-to-optical continuum observations.

While the instruments discussed above are the most utilised for high-redshift science, *JWST* also hosts MIRI and NIRISS. MIRI provides spectroscopy and imaging over a wavelength range that is inaccessible to NIRCam ($\lambda > 5\mu$m), while NIRISS provides slitless spectroscopy over a lower wavelength range than that of NIRCam grism spectroscopy ($\lambda < 2\mu$m). However, the sensitivity of MIRI imaging and spectroscopy is an order of magnitude poorer than that of NIRCam and NIRSpec, respectively (see Figure 1.7), and hence is only of use for the brightest targets at high-redshifts. Similarly, the sensitivity of NIRISS

spectroscopy is poorer than that of NIRCam grism, and moreover, the strongest emission lines (e.g. [O III] and Hβ) at $z \sim 7-9$ fall within NIRCam long-wavelength filters and not NIRISS's lower wavelength range reducing the ability of NIRISS as a redshift confirmation machine. However, it remains useful for detecting the Lyman break and strongest rest-UV emission lines of bright, high-redshift galaxies (e.g. Roberts-Borsani et al., 2022).

The sensitivity offered by *JWST*, in both imaging and spectroscopy, at $0.6\mu\text{m} \lesssim \lambda \lesssim 20\mu\text{m}$ is unrivaled by other telescopes, opening an entirely new regime in high-redshift galaxy studies (see Figure 1.7). The order of magnitude increase in sensitivity of NIRCam relative to the second most sensitive imaging instrument, on *HST*, at $\lambda \sim 1\mu$m, has facilitated the detection of a population of, previously undetectable, faint $z > 5$ galaxies (e.g. Endsley et al., 2023). Moreover, the greater than two orders of magnitude increase in depth at $\lambda \gtrsim 2\mu$m, where the Lyman break of galaxies at $z > 10$ will fall, has opened up a new parameter space in high-redshift galaxy studies. Where the wavelength range of *HST* placed its redshift detection limit at $z \sim 11$ (Oesch et al., 2016; Ellis et al., 2013), *JWST* can now observe galaxies out to $z \gtrsim 14$ (Carniani et al., 2024).

1.7. OPEN QUESTIONS

While ground (e.g. VLT and Keck) and space-based (*HST* and *Spitzer*) telescopes have revolutionised our understanding of galaxies in the first billion years of the Universe, our understanding of galaxies at $z > 7$ remained limited. Largely constrained by the wavelength range and sensitivity of these instruments, just ~ 700 photometric candidate galaxies existed at such redshifts (Bouwens et al., 2015). As such some key questions remain open at beginning of the *JWST* era:

i. Can enough ionising photons escape UV-bright galaxies as to drive reionisation?

ii. What facilitates the escape of Lyα photons from high-redshift galaxies?

iii. What are the environments of high-redshift galaxies?

iv. Do high-redshift galaxies show evidence of bursty star formation histories?

Existing ground-based observations of UV-bright galaxies, when combined with multiple diagnostics of f_{esc}(LyC), are capable of beginning to resolve question (i), as discussed in Chapter 2. The advent of *JWST* and the NIRCam instrument's high sensitivity and resolution has resulted in imaging of high-redshift galaxies. When LAEs are imaged, they are found to have close companion galaxies, and with the help of simulations, the

escape of Lyα emission can be understood through interactions between these galaxies, as described in Chapter 3, thus helping to answer question (ii). Integral field unit (IFU), on *JWST*'s NIRSpec, observations of the current highest redshift LAE, GN-z11 (Bunker et al., 2023b), reveal an abundance of UV-faint line emitters in its proximity in Chapter 4, helping to show that some galaxies in the early Universe appear to exist within extremely overdense environments of galaxies, providing hints towards the answer of question (iii). Chapter 5 utilises the wavelength range and sensitivity of NIRSpec to target spectra that show Balmer breaks and strong nebular emission lines, evidencing their complex bursty SFHs, thus providing an answer to question (iv). Finally, Chapter 6 summarises these findings and the state of the field, while discussing future avenues for expanding our understanding of the galaxies that existed in the very early Universe.

CHAPTER 2

EVIDENCE FOR A LOW LYMAN CONTINUUM ESCAPE FRACTION IN THREE MASSIVE, UV-BRIGHT GALAXIES AT $z > 7$

2.1. INTRODUCTION

Cosmic reionisation is the processes by which the neutral hydrogen in the universe, formed after the recombination phase (380 000 years after the Big-Bang), underwent a phase change - to almost complete ionisation by $z = 5.5$ (Bosman et al. 2022, Kulkarni et al. 2019, Robertson et al. 2015). Although the period over which reionisation occurs is well constrained, the sources of the ionising photons are not. The current debate is regarding whether or not the faint galaxies which represent the bulk of galaxies at $z \geq 6$ (e.g. Bouwens et al. 2015, Finkelstein et al. 2015, Atek et al. 2018) are the main drivers (Ocvirk et al. 2020, Trebitsch et al. 2022) or if the most massive objects are the major contributors (Naidu et al., 2020; Sharma et al., 2017) to reionisation. For example, Naidu et al. (2020) claim that bright galaxies may produce > 80% of the reionisation budget in order to match the observed rapid fall in neutral hydrogen fraction at $z < 8$. Directly measuring the Lyman continuum radiation escaping a source and hence ionising the universe is however not possible due to the neutrality of the IGM during the epoch of reionisation (Inoue et al., 2014, eg.). Instead the Lyman-α profile is heavily affected by neutral hydrogen and hence it can be used as a probe of the amount of ionising photons escaping the galaxy.

Given the potential Lyα emission has as a probe of neutral hydrogen, research into the properties of the line emerging from a source through surrounding neutral hydrogen were

conducted by Harrington (1973); Neufeld (1990). The absorption of Lyα at line centre by neutral hydrogen results in the main escape mechanism for photons being through diffuse in frequency producing a double-peaked spectrum. The wings of the Lyα line, that form the consequent two peaks are defined by the density of the scattering neutral hydrogen. As such the separation between the two peaks ($V_{\rm sep}$) has been proposed as a key diagnostic of the Lyman-continuum escape fraction $f_{\rm esc}$(LyC) due to their tight dependence (Izotov et al., 2017; Verhamme et al., 2015).

While this double-peaked spectrum has now been observed (eg. Yee & De Robertis, 1991; Venemans et al., 2005; Vanzella et al., 2008), especially at high-redshift, we often see reduced or a complete absence of the blue peak (Verhamme et al., 2015). This process is now well understood as a result of inter-galactic medium (IGM) radiative transfer (RT) effects (Gunn & Peterson, 1965) at high redshifts where we do not expect to find strong outflows (Vito et al. 2022, Murray et al. 2005). Photons are redshifted as they travel through the expanding universe, and hence blue peak photons get shifted into resonance. If this occurs in the vicinity of neutral hydrogen, we observe absorption, and hence any intrinsic blue peak is not observed for high neutral hydrogen column densities as predicted by Garel et al. (2021); Laursen et al. (2011).

Emission from high redshift objects typically travels through high neutral hydrogen patches in the IGM and as such, we rarely observe blue Lyα peaks. There have however been instances of double-peaked high redshift galaxies such as that discussed in Meyer et al. (2021) and Matthee et al. (2018). These galaxies typically live within large ionised bubbles resulting in a largely reduced neutral hydrogen column density, allowing the presence of a blue-peak.

In this letter, we aim to measure the mean LyC escape fraction of bright $z \geq 7$ Lyman Alpha Emitters (LAEs) based on their Lyα profile and to give new insights on their contribution to the reionisation. In section 2.2, we describe the sample of LAEs we identified at $z \geq 7$. We then describe the stacking method we use to improve the signal-to-noise ratio on our Lyα line in section 2.3 and we discuss the results and their implications in sections 2.4 and 2.5.

2.2. TARGETS SELECTION

Our primary goal is to infer the mean escape fraction of bright, $z \geq 7$ galaxies by constraining the shape and properties of Lyα, 1215.67Å, redshifted into the near-infrared. We therefore search in the literature for all $z \geq 7$ spectroscopically confirmed galaxies that have been observed by high-resolution NIR spectrographs, to guarantee the ability to

Table 2.1 The properties of the archival observations of our three target galaxies. z_{syst} indicates the systemic redshift of the target galaxy that we use for our analysis and the stellar mass is determined by the BAGPIPES fitting discussed in Section 2.2.

Name	z_{syst}	$\log(M_\star)$ $[M_\odot]$	Exposure time	Lyα luminosity $[10^{43}$ erg/s]	Telescope	Reference
EGSY-8p68	8.671	$10.1^{+0.1}_{-0.2}$	4 hrs. 45 min.	1.95±0.49	MOSFIRE	Zitrin et al. (2015)
EGS-zs8-1	7.721	$10.2^{+0.2}_{-0.1}$	4 hrs.	1.2±0.1	MOSFIRE	Tilvi et al. (2020)
			4 hrs.	1.2±0.2	MOSFIRE	Oesch et al. (2015)
COSY	7.142	$10.2^{+0.1}_{-0.8}$	12 hrs.	1.43±0.19	X-Shooter	Laporte et al. (2017a)

resolve any double-peak in the Lyα profile. Moreover, to obtain a good estimate of the UV beta slope, which can be used to measure the escape fraction at high-redshifts (Zackrisson et al., 2017), we require that our selected galaxies have at least 3 constraints on their SEDs. 11 objects with a redshift ranging from 7.15 to 9.11 were selected with Lyα luminosities ranging from ~0.05 to 2×10^{43}erg/s.

Additionally, Schenker et al. (2012) reports the detection of A1703-zd6, a redshift $z = 7.045$ galaxy that satisfies all of our selection criteria, except one - it is a lensed, low-mass galaxy (Stark et al. 2015 - $\log(M_\star) \sim 8.7$). We therefore do not include this target within our stack, but we note the measured Lyα red-peak separation from line-center of ~ 60 km/s. This is consistent with a high Lyman-continuum escape fraction as expected if faint galaxies are responsible for re-ionization (Gazagnes et al., 2020; Izotov et al., 2018b).

Because the observed shape of the Lyα line profile is a result of propagation through neutral gas, it does not trace the precise redshift of a galaxy. In order to produce a proper stack of this emission line, we also require that the galaxies used in our study have a systemic redshift measurement through the observation of another emission line. Among the 11 galaxies mentioned above, only 3 (described in table 2.1) satisfy this criteria, returning 4 datasets combining 24 hours and 45 minutes on-source exposure time. For MOSFIRE, we make use of the following programs : C228M (PI: A. Zitrin), Y288M (PI: Moncheva) and N190 (PI: Malhotra) ; and for XSHOOTER : 097.A-0043(A) (PI: Ellis). The three aforementioned galaxies have UV luminosities $M_{UV} \sim -22$, placing them on the extreme end of galaxies luminosities at $z = 7-8$ (Bowler et al., 2014, 2020).

The Spectral Energy Distribution (SED) of the 3 galaxies have been extracted from 3DHST catalogues (Brammer et al. 2012, Skelton et al. 2014). The physical properties of each individual galaxy have been estimated by SED-fitting using BAGPIPES (Carnall et al., 2018). One of the main advantages of this code is to allow the user to choose between several Star Formation Histories (SFHs). In this work, we run BAGPIPES with 4 different SFHs, namely a burst, a constant, a delayed and a combination of burst+constant. The best SED-fit is obtained for the SFH that minimised the BIC (see Laporte et al. 2021 for

more details). BAGPIPES being a parametric code, we used the following ranges for the parameters of the SFH :

- ionisation parameter : log U ∈ [-3.0,-1.0]
- dust attenuation (assuming a Calzetti law) : $A_v[mag] \in [0.0, 1.0]$
- age of the stellar population : Age[Gyr] ∈ [0.0,1.0]
- mass formed : $M_\star[M_\odot] \in [10^6, 10^{12}]$
- metallicity : Z $[Z_\odot] \in [0.0, 1.5]$

The redshift was fixed to the systemic redshift of each galaxy, and the IMF used in BAGPIPES is a Kroupa IMF (Kroupa, 2001). As expected by the selection function of our sample, our galaxies are good examples of the most massive galaxies at $z \geq 7$ with stellar masses ranging from 1.36×10^{10} to 1.75×10^{10} $M_\odot$, placing them on the extreme end of the galactic stellar mass function at $z = 7 - 8$ (see Table 2.1). The properties of our stacked spectrum obtained using BAGPIPES are found in Table 2.1.

The detection of Lyα at $z \geq 7$ implies the existence of an ionized bubble around the galaxies. (e.g. Castellano et al., 2022; Roberts-Borsani et al., 2023b). The origin of the ionising photons that produce these ionised bubbles is still highly debated and could be either due to the intrinsic nature of the object (e.g. star formation or an active galactic nucleus) or the over-dense environment near the most massive galaxies formed at high-redshift (Leonova et al., 2022; Laporte et al., 2022). Indeed previous research into these galaxies indicates that they reside within ionised bubbles, large enough that any blue-peak escaping the host galaxy would be redshifted past Lyα line-centre before leaving the ionized bubble and hence should be unaffected by IGM absorption.

Based on a relation between Lyα luminosity and ionized bubble size derived from theoretical models, Tilvi et al. (2020) estimate that EGS-zs8-1 sits in a common bubble with at least 3 neighbouring galaxies, with a size of 1.02 pMpc. More recently, Leonova et al. (2022) present analysis of *Hubble* Space Telescope (*HST*) imaging that supports the conclusion of a bubble surrounding EGS-zs8-1 given it resides in an overdensity. They additionally find that EGSY-z8p68 is found with an overdensity and again comparing to simulations find an expected bubble radius of ∼ 1 pMpc. Laporte et al. (2017b) detect Lyα, HeII and NV, as well as upper bounds on the flux of CIII] and CIV in the spectrum of COSY allowing for the determination of the source of its radiation field. The most likely hypothesis is that the radiation field of COSY is inconsistent with that from star-forming galaxies, instead it likely is produced by an active galactic nucleus (AGN) (Laporte et al., 2017b; Costa et al., 2014) and additionally its high UV luminosity makes it likely to trace

Table 2.2 The values for various parameters used for a range of diagnostics, and the $f_{\rm esc}(\mathrm{Ly}\alpha)$ that these diagnostics predict.

Diagnostic	Value	$f_{\rm esc}(\mathrm{Ly}\alpha)$
Spitzer: flux (Hβ + [OIII])	7.13×10^{-17}erg/s/cm^2/Å	[0.02:0.32]
BAGPIPES: flux (Hβ)	$1.1^{+0.9}_{-0.5}\times10^{-17}$erg/s/cm^2/Å	~ [0.09:0.18]
Spitzer: UV slope β and $\log_{10}$(EW(Hβ + [OIII]))	[-0.6, -2.35], [2.0 Å, 3.1 Å]	~ 0
BAGPIPES: UV slope β and $\log_{10}$(EW(Hβ))	$-1.89^{+0.09}_{-0.11}$, 2.67 ± 0.25	~ 0
Lyα profile: Asymmetry and Peak separation	$A=3.3^{+1.4}_{-0.8}$, $V_{\rm redpeak}=330^{+190}_{-70}$ km/s	< 0.15

an overdense region (eg. Barkana & Loeb, 2004; Furlanetto et al., 2004). Therefore it is likely COSY additionally resides within a large, ionized bubble.

2.3. METHOD

Data reduction of archival Keck MOSFIRE and VLT XSHOOTER data was performed through the standard MOSFIRE data reduction pipeline[1] and EsoReflex[2] which include standard data reduction procedures such as flat-fielding, wavelength calibration and background subtraction. The flux calibration was performed using a bright photometric standard observed during each night.

The systemic redshift is obtained from the additional emission lines present in each galaxy's spectrum (EGSY-8p69: N V, Mainali et al. (2018); EGS-zs8-1: C III], Stark et al. (2017); COSY: [C II], Pentericci et al. (2016)). We then shift each spectrum into the rest frame wavelength based on the systemic redshift reported in Table 2.1. Following this we define a new wavelength basis, spanning the range of wavelengths of interest with a wavelength bin equal to that of our lowest resolution spectra (~ 0.15 Å). The total flux of each object within each wavelength bin is calculated, converted into a luminosity, and the median of these luminosities is then taken as the value for our stacked spectrum. We then return this median spectrum to units of flux by dividing through by the luminosity distance of the median redshift in our sample. We take the standard deviation of the fluxes in each bin to obtain the error in our stacked spectrum.

We then use a Monte Carlo (MC) error propagation method to estimate the uncertainties on the observed red-peak velocity offset and asymmetry. We use the error in our spectrum, described above, to redraw each spectral bin's flux from a Gaussian centred on the median stack value with a standard deviation equal to the aforementioned error. We then repeat this process, measuring the red-peak offset and asymmetry, one hundred thousand times in order to understand the uncertainty in our measurements driven by the

[1]https://keck-datareductionpipelines.github.io/MosfireDRP/
[2]https://www.eso.org/sci/software/esoreflex/

error associated with each flux measurement. We take the red-peak offset and asymmetry associated with our stack and measure the standard deviation of those values above and below the median for our upper and lower bound uncertainties on this measurement respectively.

We have additionally evaluated the results when normalising all of our spectra by dividing through by their peak Lyα flux, as not to weight our stack based on the most luminous galaxy. Given that every galaxy has a similar luminosity, we see no change in our Lyα profile regardless of normalisation. A median stacking method is chosen as it acts to reduce the affect of sky-lines, however, in the case of COSY, significant contamination from a sky-line in a region of the spectrum where we expect to observe the tail of our red-peak profile did act to pollute our stacked spectrum and it was therefore masked. Adopting a mean stacking method and additionally changing the position and widths of bins provides a very similar stacked spectrum thus indicating the robustness of our stack.

2.4. RESULTS

2.4.1 Constraints from SED

The Lyα escape fraction is the ratio of the Lyα flux escaping a galaxy over the intrinsic Lyα flux of the object. In order to ascertain limits on the potential intrinsic flux of Lyα we use recombination lines with known relations to Lyα. Unfortunately no direct observations of recombination lines are made in any of our target spectra. Instead we make use of two independent methods to estimate the equivalent width (EW) of the Hβ emission line: (i) using the flux ratio between the 4.5μm and 3.6μm assuming that the 3.6μm flux is the stellar continuum and the 4.5μm is the sum of the stellar continuum with a contamination of OIII+Hβ and (ii) using BAGPIPES to directly predict the flux of Hβ (see Table 2.2 for the results of both methods). The first method returns the flux of Hβ and OIII in combination and thus by assuming no contribution from OIII supplies us with a lower bound estimate of the flux of the recombination line Hβ. We assume a maximum contribution of $\log_{10}([\text{OIII}]/\text{H}\beta) = 1$ - based on the assumption our stack lies in the extreme AGN region of the Baldwin, Philips and Terlevich (BPT) diagram (Baldwin et al., 1981; Veilleux & Osterbrock, 1987). This conclusion is in itself unlikely as NV emission lines are either weak or not present in our individual galaxy spectra but allows us to obtain an absolute upper bound on the Hβ flux. The second method is based on SED-fitting and depends on the best fit parameters such as the stellar mass, the reddening, the age, the metallicity, etc. To verify that the Hβ flux estimated by BAGPIPES is not strongly dependant

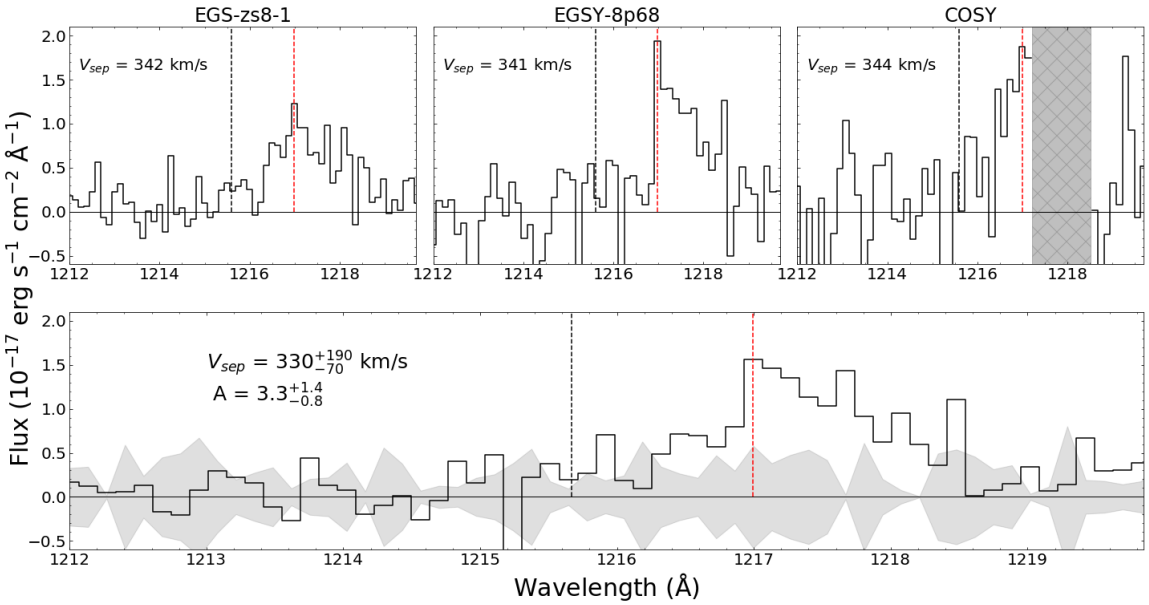


Fig. 2.1 Line profiles of Lyα emission in our target galaxies (top) and the stacked spectrum (bottom). Line centre is denoted by a black dashed line, while the red-peak flux is indicated by a red dashed line. The velocity offset of the red-peak from line centre is additionally found within each panel. (Top:) From left to right: EGS-zs8-1, EGSY-8p68, COSY. The resolution of the spectrum of COSY has been reduced to the resolution of the two MOSFIRE spectra (0.14 Å). The hatched box indicates a region of the spectrum that has been removed due to significant pollution by a sky line. (Bottom:) The grey region indicates the 1-sigma error obtained by taking the standard deviation of the constituent galaxies of the stack. Additionally, red peak velocity offset and the asymmetry of the red-peak as defined in Section 2.4 are found within the panel.

on other parameters, we study the evolution of Hβ flux as a function of the metallicity and stellar mass and as a function of the metallicity and reddening. The variation in the EW is estimated as $\Delta \log EW < 0.5$.

We assume, following Gazagnes et al. (2020), that $L_{Ly\alpha} = 8.7 L_{H_\alpha}$, and in turn that $L_{H_\alpha} = 2.85 L_{H_\beta}$, however Ly$\alpha$ emission can include a large contribution from collisional excitation (Mitchell et al., 2021; Smith et al., 2022b) thus increasing the intrinsic $L_{Ly\alpha}$ further decreasing the measured escape fraction.

However, given we observe no Lyα blue-peak we must assume the Lyα profile has undergone some absorption due to IGM attenuation. Given many Lyα profiles of low redshift LAEs, that have not travelled through a high neutral hydrogen density IGM, exhibit often equal or less than equal blue-to-red peak flux ratios (eg. Izotov et al., 2018a,b) we assume that the upper bound of the $L_{Ly\alpha}$ escaping the galaxy is double the $L_{Ly\alpha}$ that we observe, while the $L_{Ly\alpha}$ that we observe represents a lower bound. Taking the bounds of L_{H_β} that we obtain from Spitzer provides us with a potential range of $0.02 < f_{esc}(Ly\alpha) < 0.32$, while the best-fit L_{H_β} from BAGPIPES returns $f_{esc}(Ly\alpha) = 0.09^{+0.07}_{-0.04}$, while taking the bounds on $L_{Ly\alpha}$ the best-fit value ranges from $0.09 < f_{esc}(LyC) < 0.18$.

The bounds of the Hβ flux as well as the best-fit value and the associated escape fraction for the stacked spectrum can be found in Table 2.2. It has been shown that the escape fraction of the Lyman-continuum (LyC) is almost always lower than the Lyα escape fraction (Maji et al., 2022) and hence we take the bounds on f_{esc}(LyC) to be the same as those on $f_{esc}(Ly\alpha)$ (also reported in Table 2.2).

We additionally consider the relation between the UV slope β and the EW(Hβ) first determined by Zackrisson et al. (2013) at redshifts $z > 6$ and then at $z \approx 7-9$ by Zackrisson et al. (2017) by studying the evolution in synthetic galaxy spectra with changing f_{esc}(LyC). The results of this analysis can be seen in figure 2.2 where galaxies with log(EW(Hβ)) $\gtrsim 2$ exclusively have $f_{esc}(LyC) = 0$. As is clear in figure 2.2 both the best-fit EW(Hβ) and UV slope β (corrected for dust extinction following Meurer et al., 1999) returned from BAGPIPES and the range in EW(Hβ) and UV slope β (determined following Bouwens et al., 2014) that can be estimated from Spitzer data (the values of which are reported in Table 2.2) constrain our stack to a region of the figure that is not compatible with an $f_{esc}(LyC) \gg 0$.

We do note that this diagnostic is potentially limited in its ability to diagnose high escape fractions given examples of low redshift galaxies that have high f_{esc}(LyC) and high EW(Hβ) (eg. Izotov et al., 2018a). The diagnostic requires many assumptions in order to estimate f_{esc}(LyC), notably in the stellar models employed. Changing these stellar models can significantly affect the result of the diagnostic, as seen when binary

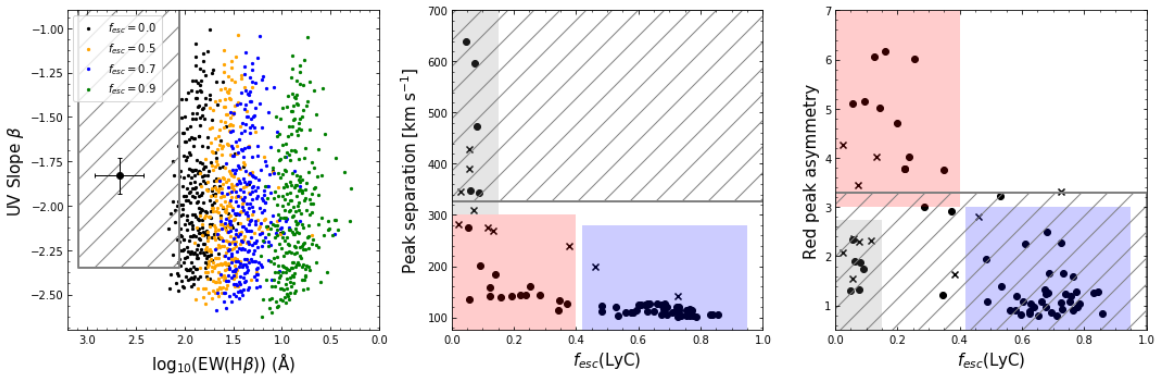


Fig. 2.2 The indirect diagnostics of $f_{\rm esc}$(LyC), with the region which our stack resides indicated with a grey hatched box. (Left) The EW(Hβ) of simulated $z = 7-9$ galaxies against their UV slope β, assuming a Calzetti attenuation law with $E(B-V)_{\rm stars} = E(B-V)_{\rm neb}$ (from Zackrisson et al., 2017). Their Lyman-continuum escape fraction is denoted by their colour, $f_{\rm esc}$(LyC) = 0.0, 0.5, 0.7, 0.9 correspond to red, orange, green and blue respectively. The grey hatched region indicates the location of our stack using Spitzer data, while the black data point indicates the position using the BAGPIPIES best-fit on the SED. (Centre) The peak separation of Lyα profiles against their LyC escape fraction (from Kakiichi & Gronke, 2021). The dots denote simulated results from Kakiichi & Gronke (2021), while crosses indicate the results for $z \sim 0.3$ LyC-detected galaxies from Izotov et al. (2016, 2018a,b). The coloured regions indicate the three regimes of LyC escape - leakage by full break, through holes and small leakage with few or no holes indicated by blue, red and grey respectively. (Right) The red peak asymmetry of Lyα profiles against their LyC escape fraction (from Kakiichi & Gronke, 2021). The markers and shading are the same as the central panel.

evolution is considered (see figure 6 in Zackrisson et al., 2017). However, we are aware of the main outcome of Zackrisson et al. (2017) - that galaxies with $f_{\rm esc}$(LyC)>0.5 should have EW(Hβ) < 30 Å. Therefore, we instead consider a high EW(Hβ) to be a necessity for low $f_{\rm esc}$(LyC), although perhaps not sufficient. Given the inclusion of multiple diagnostics all indicating low $f_{\rm esc}$(LyC) we consider the potential uncertainty surrounding this diagnostic not to be a significant issue.

We additionally note that Zackrisson et al. (2017) provides this diagnostic for a range of different dust attenuation laws. While the panel in figure 2.2 assumes a Calzetti attenuation law with $E(B-V)_{\rm stars} = E(B-V)_{\rm neb}$, we find that the conclusion, that our stack lies within a region of the diagram corresponding to $f_{\rm esc}$(LyC) = 0, is consistent regardless of the dust attenuation law used in Zackrisson et al. (2017).

2.4.2 Lyα profile

Figure 2.1 clearly indicates that for all of our targets, we find the red-peak of the Lyα profile to be offset from line centre by ~ 340 km/s, such a large separation is indicative of a low Lyman-continuum escape fraction (Izotov et al., 2018b; Gazagnes et al., 2020; Kakiichi & Gronke, 2021). The large offset of the red-peak is additionally present in the stacked spectrum, as well as a clear asymmetry. This asymmetry, A, is the ratio of the blue-to-red flux of the red peak (as defined in Kakiichi & Gronke 2021). We find that the two targets for which we are able to observe the shape of the red-peak profile, we observe clear asymmetry.

2.5. Discussion

The limits placed on $f_{\rm esc}$(Lyα) from photometry discussed in Section 2.4.1 are already low enough to rule out the possibility of these three massive, bright galaxies currently being significant contributors to re-ionization. However, we wish to use multiple diagnostics in order to confirm these findings. The results from our Lyα stack are therefore crucial to further constrain the escape fraction.

2.5.1 Interpretation of the velocity offset

While the velocity offset of Lyα from the systemic redshift initially appears as though it may be primarily driven by outflows in these massive galaxies, we believe this to be unlikely. Neufeld (1990) and Michel-Dansac et al. (2020), using a static medium with large neutral hydrogen column densities, find velocity offsets in their simulations that are

comparable to those that we observe indicating these velocity separations are achievable within simulations without modelling for outflows.

Any such shift in the Lyα profile due to the expansion velocity, $v_{\rm exp}$, of neutral gas would still result in the expected double-peaked profile of Lyα centred on the systemic redshift. Results from Verhamme et al. (2015) indicate that increasing $v_{\rm exp}$ has the effect of reducing the peak separation, such that when $v_{\rm exp} > 300$ km/s, they cannot recreate the red peak to line-centre separation that we observe. These results constrain the neutral gas column density in our stack to be greater than $10^{20}{\rm cm}^{-2}$ and any outflow velocity $v_{\rm exp} < 300$ km/s.

The conclusion that any expansion velocity will act to reduce peak separation thus allows us to conclude that our red-peak offset is a minimum separation. Additionally, most diagnostics of escape fraction that use the separation between the peaks of the Lyα profile are based on observations of lower mass galaxies than our targets and therefore using any such diagnostic is challenging (Izotov et al., 2018b; Gazagnes et al., 2020). Instead we choose to use the relation determined by simulations from Kakiichi & Gronke (2021), allowing us to avoid mass biases. The observed red-peak-offset in Figure 2.1 can be used as a lowest bound for the blue-red peak separation given that we know the blue peak lies on the blue side of line-centre. As such, we expect the blue-red peak separation to far exceed the red-peak offset of ~ 300 km/s and comparing this to the relation from Kakiichi & Gronke (2021) we find that $f_{\rm esc}({\rm Ly}\alpha) \lesssim 10\%$ and Lyα photons escape through an optically-thick medium with few or no holes.

2.5.2 Interpretation of the red-peak asymmetry

This asymmetry allows us to quantify the amount Lyα photons have to scatter, in doing so creating a broad wing component of the emission line, in order to escape the galaxy. A high asymmetry (A > 3) is hence indicative of Lyα photons having multiple routes to escape and hence scatter significant amounts in order to find low-density channels to escape the galaxy (leakage through holes), while a low asymmetry (A < 3) is indicative of Lyα photons only having one method of escape possible either through predominantly optically-thin (leakage by full break) or optically thick (small leakage due to few or no holes) media Kakiichi & Gronke (2021). Therefore we can use the asymmetry to attempt to diagnose the properties of the medium through which the Lyα photons have traversed.

The asymmetry that we observe, in Figure 2.1, is an upper bound on the asymmetry of the intrinsic spectrum. This is due to IGM attenuation reducing the flux close to line-centre hence reducing the flux between the red-peak and line-centre relative to the flux on the red side of the red-peak, therefore the observed asymmetry is greater than the intrinsic

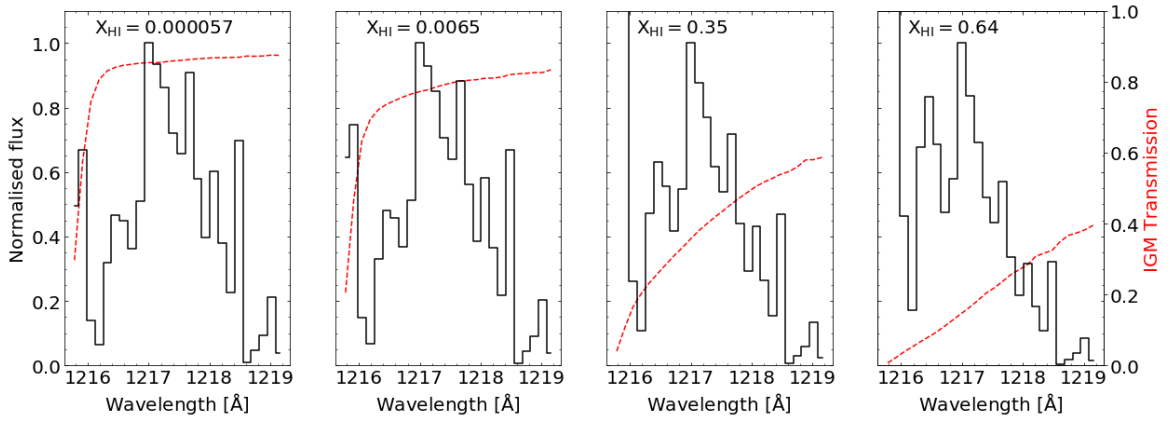


Fig. 2.3 The potential intrinsic stacked spectrum created by dividing the observed spectrum by the IGM transmission for a range of different volumetric neutral fractions indicated in the top left of each panel. The spectrum normalised by the red peak flux is indicated by the solid black line, while the IGM attenuation curve associated with the volumetric neutral fraction, taken from Garel et al. (2021), is indicated by the red dashed line.

asymmetry. As such we find the asymmetry, A < 3, results in the interpretation that Lyα photons have either escaped through a full break environment or by leakage through few or no holes. Given the aforementioned limits on the escape fraction ($f_{\rm esc}$(Lyα)$\lesssim$ 10%) we can constrain ourselves to small leakage without the presence of optically-thin channels (Kakiichi & Gronke, 2021).

2.5.3 Effects of IGM attenuation

Neutral hydrogen in the IGM causes attenuation of Lyα close to line-centre (see Garel et al., 2021), therefore in order to confirm that we do not misidentify the location of the red-peak flux, we divide our observed stacked spectrum through by attenuation curves for varying volumetric neutral fraction, taken from Garel et al. (2021). Given our target galaxies all likely reside within large ionized bubbles we do not expect significant IGM absorption. Figure 2.3 indicates the effect of correcting for the different IGM transmission curves and we see no notable difference in our observed spectra even at the most extreme volumetric neutral fraction, indicating our observed red peak separation likely trace that intrinsic to the galaxy before IGM absorption of Lyα. While we do see a notable decrease in the asymmetry, as predicted in Section 2.5.2, we have already consider the asymmetry to be a lower bound and as such this does not affect the interpretation of the result.

We do note a significant increase in the flux at line-centre, in Figure 2.3, due to the effectively zero transmission through the IGM at that wavelength. This is merely an

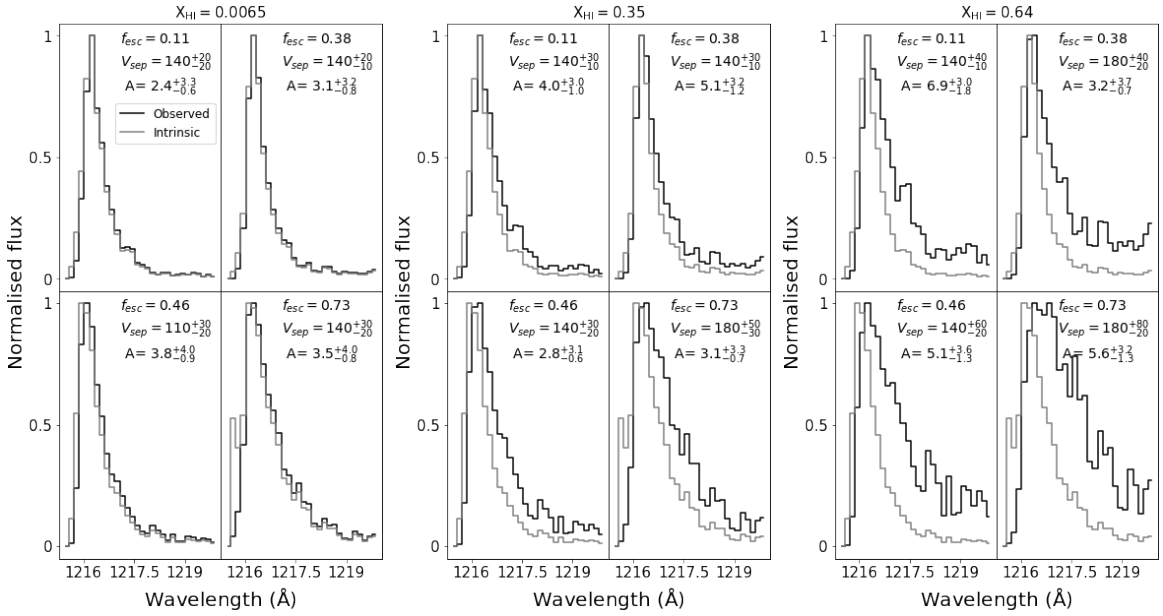


Fig. 2.4 The simulated spectra of high redshift, high $f_{\rm esc}$(LyC) galaxies for varying assumed neutral hydrogen fractions. Each column uses an increasing neutral hydrogen fraction from left to right, that in turn dictates which IGM attenuation curve, taken from Garel et al. (2021), is applied to the original spectrum. Each panel includes the same 4 galaxies from Izotov et al. (2018a,b) (clockwise from top left sub-panel: J1011+1947, J1256+4509, J1243+4646, J1154+2443), whose Lyman continuum escape fraction are indicated at the top of each sub-panel. The simulated observed spectrum normalised by the red peak flux is indicated by the solid black line, while the original spectrum (with resolution degraded) is indicated in grey.

artefact of noise being divided through by a number tending to zero rather than any physical intrinsic property of the galaxy. We know this to be true as Lyα emission will immediately be absorbed at line-centre by any neutral hydrogen within the host galaxy and as such we must observe negligible flux at line-centre escaping the host galaxy.

Finally, in order to confirm that we are not being affected by high escape fraction interlopers that due to IGM transmission, spectral resolution and noise are being interpreted as having a low escape fraction, we attempt to recreate high redshift observations of galaxies with $f_{\rm esc}$(LyC) greater than our sample ($f_{\rm esc}$(LyC)> 0.1). We use the spectra of galaxies from Izotov et al. (2018a,b) as examples of Lyα profiles associated with an $f_{\rm esc}$(LyC) greater than our sample up to a value of 72% at low redshifts ($z \sim 0.3$). We apply the IGM attenuation curves from Garel et al. (2021) to the Lyα profile, we then reduce the resolution of these spectra down to the resolution of our stacked Lyα profile and finally we use the MC error propagation described in Section 2.3 to estimate uncertainties on the $V_{\rm red}$ and asymmetry of each galaxy given a noise level similar to our stacked spectrum (by assuming the peak flux to be at SN = 5).

Figure 2.4 shows our simulated high redshift, high $f_{\rm esc}$(LyC) spectra with various IGM transmission curves applied for differing neutral hydrogen fractions. We find that applying the IGM transmission and degrading the resolution of the spectra result in velocity offsets that are consistent with the original spectra, even for the most extreme neutral hydrogen density, to within 40 km/s. This peak separation when considered in the context of the Kakiichi & Gronke (2021) diagnostic appears to indicate these galaxies are likely high $f_{\rm esc}$(LyC). All of these galaxies exhibit red peak offsets $\ll 300$ km/s thus allowing us to conclude that the observation of a peak offset of ~ 300 km/s is not only indicative of a low $f_{\rm esc}$(LyC) but also that we are likely not being affected by high $f_{\rm esc}$(LyC) interlopers. We do however note that the asymmetry is more challenging to understand as a function of the neutral fraction. It is clear that for a sharp Lyα red-peak, increasing the neutral density will act to increase the flux on the red side of this peak hence increasing the asymmetry. When the red-peak is more broad, increasing the neutral density can push the peak of the Lyα profile red-ward and hence act to increase the amount of flux on the blue side of the peak. This complicated interplay of effects leads to a highly uncertain asymmetry in some of our Lyα profiles. Therefore, the use of asymmetry to diagnose $f_{\rm esc}$(LyC) alone at high redshifts, where neutral hydrogen in the IGM causes large uncertainties on the intrinsic asymmetry, should be avoided. However, the asymmetry of our stacked spectrum appears relatively well defined with a sharp drop in flux blue-ward of the peak flux. Therefore, we conclude that our asymmetry is most likely an upper-bound asymmetry, where the relative boosting of flux to the red side of the Lyα peak due to IGM attenuation, is the most

likely of the two aforementioned effects at play. Given the relatively small uncertainty on the intrinsic Lyα asymmetry of our stacked spectrum we believe that, in combination with multiple other diagnostics, the observed asymmetry can be used to infer the ability of Lyα photons to escape their host galaxy, with the caveat that this diagnostic should not be used on high redshift LAEs without the use of other supplementary diagnostics.

2.6. SUMMARY

In order to probe the potential ionising properties of the most massive ($\log_{10}(M_\star[M_\odot]) > 10$), UV-bright ($M_{UV} \sim -22$), high redshift ($z > 7$) galaxies, we target all archival data on telescopes with resolution $R \gtrsim 3000$, allowing us to obtain a resolved Lyα profile. We find a total of four observations of three satisfactory galaxies with Lyα emission, totalling an exposure of 24 hours and 45 minutes. Using a median stacking method we obtain a deep stacked spectrum representing massive, UV-bright, high redshift Lyα leaking galaxies. Through the analysis of the stacked Lyα profile, using the red-peak velocity offset from line-centre and the red-peak asymmetry, we deduce the Lyman-continuum escape fraction to be less than 10% and that the few Lyman-continuum photons that do escape, escape through an optically-thick medium with few or no holes. Through the use of Spitzer observations of our target galaxies, stacking these and SED-fitting using BAGPIPES we obtain bounds on the recombination line Hβ. Given this we constrain the escape fraction to $9\% < f_{\rm esc}(\text{Ly}\alpha) < 18\%$ in strong agreement with the results of our stacked Lyα profile. We additionally confirm that neither IGM attenuation or a significant outflow velocity could affect our conclusion regarding a low $f_{\rm esc}$(Lyα) for massive, UV-bright, high redshift galaxies. Our study shows that despite the fact the 3 galaxies analysed lie within ionised bubbles, they are not capable themselves of ionising their own bubbles.

However, we emphasize that our result is obtained using only 4 datasets of 3 different galaxies at $z \geq 7$ – the only observations currently available in telescope archives. Increasing the number of Lyα detections at $z \geq 7$ with high-resolution spectrographs is therefore crucial to confirm our conclusions. Furthermore, the high-fraction of neutral gas underlying galaxies within the epoch of reionisation limits the detection of Lyα to galaxies in overdense regions. Spectrographs with a large field-of-view will therefore be ideal instruments to push forward this project. MOONS, a 3^{rd} generation instrument at the Very Large Telescope, will be one of those. It combines high-resolution (R>4000), a large field of view (~500 arcmin^2) and a huge number of fibres (~1000).

CHAPTER 3

DECIPHERING LYMAN-α EMISSION DEEP INTO THE EPOCH OF REIONISATION

3.1. INTRODUCTION

Young, vigorously star-forming galaxies have been identified in the very early Universe (Curtis-Lake et al., 2023; Arrabal Haro et al., 2023; Williams et al., 2023). These galaxies should be excellent sources of Lyman-α emission (Lyα, λ=1215.67Å) – the intrinsically brightest emission line (Dijkstra, 2014) – which stems from the recombination of hydrogen that has been ionised by their young stellar populations. However, deep in the epoch of reionisation galaxies are expected to be exceptionally gas-rich such that their stellar nurseries are enshrouded in copious amounts of neutral hydrogen, which leads to extreme damped absorption of Lyα (Heintz et al., 2024c). Furthermore, the intergalactic medium (IGM) is increasingly neutral as we probe to higher redshift (Robertson et al., 2015; De Barros et al., 2017) and this neutral gas is expected to resonantly scatter Lyα emission. Hence, due to 'local' attenuation by a gas-rich interstellar medium (ISM) and scattering by a neutral IGM, Lyα emission should only be detectable towards the end of the reionisation era, about one billion years after the Big Bang (Dijkstra, 2014; Endsley et al., 2022; Bolan et al., 2022; Hsiao et al., 2023).

While the decreasing observability of Lyα emission with increasing redshift has been repeatedly claimed to be observed (Stark et al., 2017; Jones et al., 2024), this picture has been challenged by the occasional, surprising detection of Lyα emission in several galaxies deep in the reionisation era (Bunker et al., 2023b; Jung et al., 2024). It has been suggested that Lyα can escape through the neutral IGM if the galaxies reside in sufficiently large ionised bubbles embedded within the neutral IGM, driven either by active galactic nuclei

(AGN) (Laporte et al., 2017b; Larson et al., 2023; Maiolino et al., 2024a) or by an enhanced radiation field produced by an overdensity of associated objects (Castellano et al., 2016; Tilvi et al., 2020; Leonova et al., 2022; Morishita et al., 2022; Tang et al., 2023; Tacchella et al., 2023b; Witstok et al., 2024c). However, a solution to the escape of Lyα emission through the ISM and CGM of what are expected to be very gas-rich galaxies remains elusive. Before the advent of *JWST*, the sensitivity and resolution of imaging instruments meant that studies of Lyα emitters (LAEs) at high redshift were spatially unresolved. Hence, it was not possible to probe the physical processes that could explain the escape of Lyα emission from the ISM.

To address this crucial issue we study a sample of nine galaxies that have been spectroscopically confirmed with the detection of Lyα emission at $z > 7$ with high-resolution spectrographs and which have been observed with publicly-available *JWST* Near Infrared Camera (NIRCam) (Rieke et al., 2023a) imaging (the properties of which are reported in Table 3.2). These galaxies fall within the GOODS-North, GOODS-South, EGS and COSMOS fields, and have been observed as part of five different programs: PRIMER (PI: Dunlop), FRESCO (PI: Oesch)Oesch et al. (2023), CEERS (PI: Finkelstein)Bagley et al. (2023), JADES (PI: Eisenstein)Rieke et al. (2023b) and DDT program 4426 (providing NIRSpec IFU observations of GN-z11, PI: Maiolino)Maiolino et al. (2024b). Six of these nine Lyα-emitting galaxies are known to lie within overdensities (Tang et al., 2023; Jung et al., 2022; Leonova et al., 2022; Tacchella et al., 2023b; Witstok et al., 2024c), three of which also likely host accreting black holes (Laporte et al., 2017b; Larson et al., 2023; Maiolino et al., 2024a). Moreover, it has been shown that the remaining galaxies are incapable of alone blowing a large enough ionised bubble to facilitate the escape of Lyα through the neutral IGM, suggesting that an underlying population of faint galaxies must be surrounding these LAEs (Witstok et al., 2024c).

However, the presence of spectroscopically confirmed galaxies within known ionised bubbles that do not show Lyα emission (Tang et al., 2023; Jones et al., 2024; Witstok et al., 2024c) indicates that there *must* be further *local* processes at play driving Lyα emission deep into the epoch of reionisation. With this in mind, we use the unparalleled sensitivity and resolution of *JWST*/NIRCam in order to accurately determine the properties of these nine LAEs, with equivalent width (EW) greater than 10Å, at $z > 7$.

We utilise *JWST* imaging and spectroscopy to study the environments of these high-redshift LAEs, this data and its reduction is described in Section 3.2. The analysis of these environments and the detection, characterisation and confirmation of companion galaxies is discussed in Section 3.3. A comparison to the *Azahar* simulations provides

further understanding of the physical processes driving the observed Lyα emission in Section 3.4. Finally, these results are summarised in Section 3.5.

3.2. DATA

The emergence of *JWST* with its significant technical advancements over both ground- and space-based telescopes, has the potential to reveal new hints as to the driving forces behind the observed Lyα emission discussed in this paper. We therefore decided to analyse all spectroscopically-confirmed LAEs at $z > 7$, with *JWST*/NIRCam imaging. This search originally revealed eleven such LAEs, however, further constraints on the SED of these galaxies from NIRCam imaging reveals the Lyα emission line detected in two of these LAEs is in fact not Lyα (discussed in more detail below). Therefore, the sample becomes nine LAEs (seen in Table 3.2). These galaxies are found in the GOODS-North and GOODS-South, EGS and COSMOS fields which have been observed with NIRCam by FRESCO (PI: Oesch, ID: 1895) and JADES (PI: Eisenstein, ID: 1180), CEERS (PI: Finkelstein, ID: 1345) and PRIMER (PI: Dunlop, ID: 1837), respectively.

The PRIMER survey covers $\sim 150\text{arcmin}^2$ of the COSMOS field with the F090W, F115W, F150W, F200W, F277W, F356W, F410M and F444W filters all at a 5σ limiting depth of $m_{\rm AB} > 28$. The CEERS survey took imaging of 100arcmin^2 of the EGS field in the F115W, F150W, F200W, F277W, F356W, F410M and F444W filters, reaching a 5σ limiting depth of $m_{\rm AB} > 29$ in the short-wavelength or $m_{\rm AB} > 28.4$ in the long-wavelength filters (Finkelstein et al., 2023). The June 2023 JADES data release surveys a ~ 36 arcmin^2 area of the GOODS-S field using the F090W, F115W, F150W, F200W, F277W, F335M, F356W, F410M and F444W, reaching 5σ limiting depths of $m_{\rm AB} \sim 30$ (Eisenstein et al., 2023).

The FRESCO survey's primary aim is to obtain NIRCam Long-Wavelength (LW) Grism spectra over the F444W filter for galaxies across GOODS-S and GOODS-N. However, the simultaneous imaging in the short-wavelength channel provides imaging in both F182M and F210M filters, and direct imaging also obtained in the F444W filter additionally provides us with a total of 3 filters for our galaxies in GOODS-S/GOODS-N. FRESCO's LW WFSS brings significant advantages as the F444W filter covers the wavelength range of $3.9 - 5\mu$m, therefore observing both [O III]and Hβ at $z \sim 7 - 9$. These emission lines ordinarily contaminate imaging in the F444W filter and hence constraints on the fluxes of these emission lines allow us to reduce the uncertainty in galaxy properties such as stellar mass, star-formation rate and age. We use these constraints on the emission lines present in F444W to create a mock F430M filter. Furthermore, when we fit the SED (discussed further below) we check the fluxes of the [O III]and Hβ emission lines in the best-fit SED

model and confirm that these are consistent with the measured fluxes from the WFSS spectrum.

3.2.1 Data Reduction

Throughout our analysis we use the final reduced data products created by the PRIMER team for the COSMOS field, the publicly available reduced data products from the CEERS team for the EGS field, the publicly available reduced data products from the JADES team for the deep *HST* region of GOODS-S and our own reduction of the GOODS-S FRESCO imaging for regions outside of the publicly available JADES imaging. In order to produce the GOODS-S images, we take all available NIRCam imaging data of GOODS-S, available from the FRESCO survey, and reduce them using version 1.8.5 of the *JWST* pipeline. We first download the uncalibrated files from the MAST archive and follow the steps of the *JWST* pipeline: (Stage 1) detector-level corrections and ramp fitting, (Stage 2) instrument-level and observing-mode corrections producing fully calibrated exposures, (Stage 3) producing the final mosaic. We take additional care after Stage 1 to ensure the removal of horizontal and vertical striping that can be present in the rate files.

We extract the photometry of our objects using SExtractor (Bertin & Arnouts, 1996) on PSF-matched images extracting all objects with more than 4 pixels (DETECT_MINAREA 4) detected at more than 1.5σ (DETECT_THRESH and ANALYSIS_THRESH 1.5). We used large deblending parameters to allow efficient separation of each component (DEBLEND_NTHRESH 64 and DEBLEND_MINCOUNT 0.000001), however, the small separation between the components of the EGSY8p68 and JADES-GS-z7-LA systems makes the extraction of each component's photometry impossible with SExtractor. For JADES-GS-z7-LA we employ the photometry of the LAE and its companion reported in Saxena et al. (2023) while for EGSY8p68 we model the shape of each component using Galfit (Peng et al., 2002) assuming a Sersic profile and NIRCam PSFs, simulated using WebbPSF. We first run Galfit on one of the NIRCam images where all the components are clearly resolved (F115W), and we then fix the resultant position, Sersic index, axis ratio and angle for each component when extracting the photometry for the remaining images. Figure 3.1 shows the modelling of the components for this system.

The NIRSpec Mutli-Shutter Array, used to observe three of the systems (EGSY8p68, CEERS-1027 and JADES-GS+53.15682-27.76716), has a small shutter size and hence only the LAE and not its environment is coincident with the shutter position. We do however make use of all existing ground-based observations (which typically have larger slit sizes) of our sample of LAEs, of which four of the sample have been observed by MOSFIRE (z7-13433 - ID: N199, PI: Jung; EGSY8p68 - ID: C228M, PI: Zitrin; GSDY - ID: U069, PI: Treu;

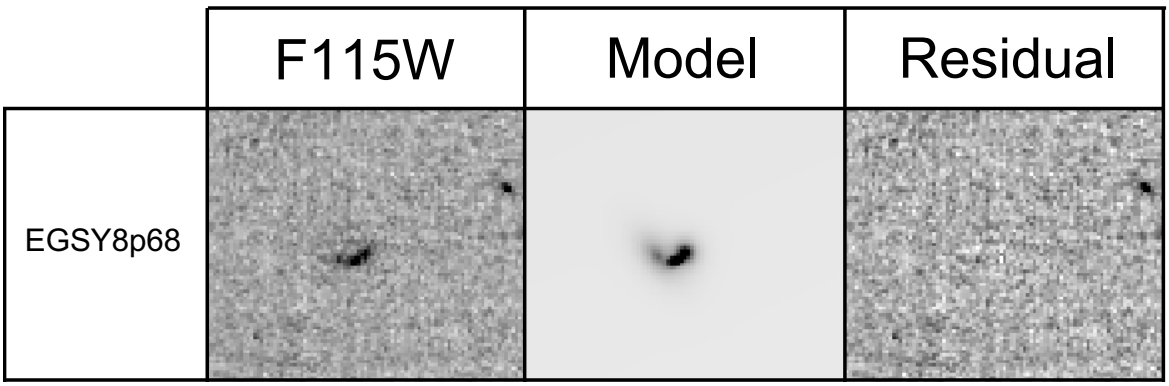


Fig. 3.1 Example of GALFIT modelling for the EGSY8p68 system for which the separation between each component is too small for the photometry to be extracted by SEXTRACTOR.

z7-GSD-3811 - ID: C182M, PI: Scoville) and one with X-shooter (COSY - ID: 097.A-0043, PI: Ellis).

We reduce the spectroscopic data using the standard MOSFIRE data reduction pipeline and EsoReflex. These pipelines perform wavelength callibration, sky subtraction and flat fielding as well as further standard data reduction steps in order to produce two-dimensional (2D) spectra. The resultant 2D spectra span the length of the original slit used for the observations and have spectral and spatial resolutions of $R \sim 3380$, 0.1798 $''$/px and $R \sim 8900$, 0.158 $''$/px for MOSFIRE and X-shooter respectively, but are limited by the seeing of each observation which can be as poor as $\sim 0.9''$.

Two of our sample are located within the FRESCO LWSS footprint in the GOODS-S field (z7-GSD-3811 and GSDY) . We reduced this Grism data following the steps described in Sun et al. (2023a) including flat fielding, background subtraction, 1/f noise subtraction and WCS assignment. We also perform a careful astrometric analysis to avoid any offsets between the LW direct imaging in F444W and the LW channel Grism spectra in F444W. This allows us to extract the 2D Grism spectra associated with the astrometric position of any source, these spectra are then stacked and the 1D spectrum at the position of the source with an aperture of 3 pixels is extracted. We also perform a continuum subtraction from the 1D spectrum to remove any contaminant continuum and we focus on measuring [O III]and Hβ emission line fluxes.

3.3. RESULTS

Remarkably, the NIRCam images of our sample reveal the presence of multiple components to all of these LAEs, as shown in Figure 3.2. It is first crucial to constrain the photometric redshift in order to ascertain whether these are indeed high-redshift galaxies

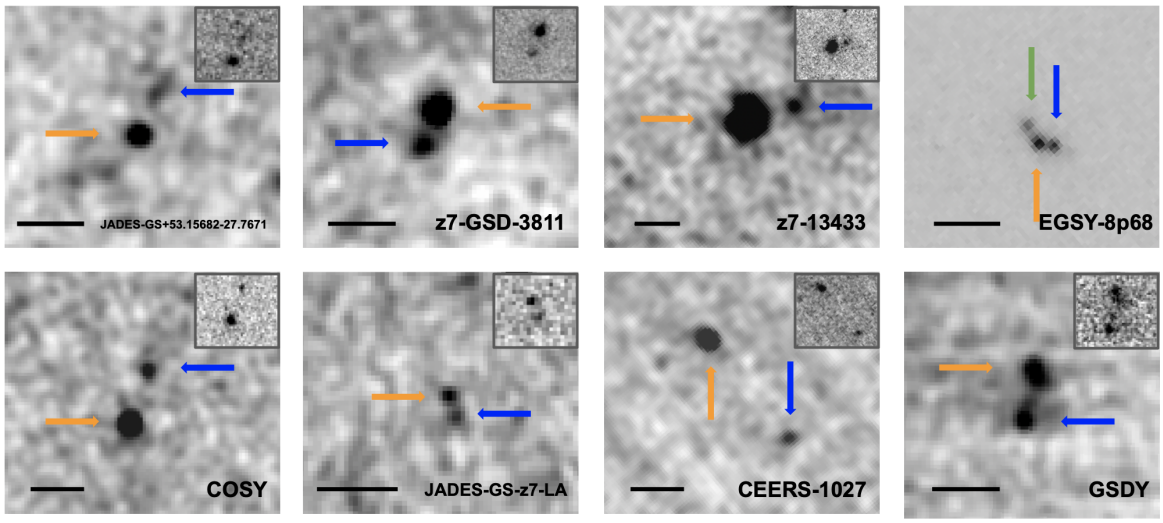


Fig. 3.2 Cutouts of *JWST* images (as different surveys make use of different filters, the filter within which the companion is most clearly present is shown: F182M for GSDY and z7-GSD-3811, F200W for EGSY8p68, F115W for CEERS-1027, JADES-GS-z7-LA and COSY and F150W for JADES-GS+53.15682-27.7671 and z7-13433) showing the multiple components of each system. All images are smoothed with a Gaussian of FWHM ~ 0.7 kpc to enhance the visibility of the companion, except for EGSY8p68 where the three components become unresolved if we smooth the images. We include the unsmoothed images as sub-panels in the top right of each panel in order to emphasise that the companion is always visible before smoothing. We use different colorbar ranges for each system in order to make the companion galaxy well defined, hence some unsmoothed images can have what appears like a nosier background than others. The position of the main target (*A in the text) is displayed by an orange arrow, the position of the first companion (*B) is indicated by a blue arrow and the third companion, if any, by a green arrow. The name of each candidate is indicated in the bottom right corner, and in the bottom left of each panel is a black scale showing 0.5″.

or contaminants. We estimate the redshift of each candidate companion by carefully extracting the Spectral Energy Distribution (SED) of both the main component and its companions. We then fit each SED by assuming several parametric star formation histories (SFH - burst, constant, delayed, exponential), allowing a large range for all parameters (Carnall et al., 2018). In all cases, we deduce the photometric redshifts of the companions to be consistent with that of their spectroscopically-confirmed main component. One key diagnostics of redshift is strong [O III]emission falling into a medium band filter (seen in Figure 3.3 and 3.4), however, it then becomes crucial to understand the probability of these being unrelated, high-redshift objects.

3.3.1 Redshift Determination

We initially take care to ensure that, for galaxies that only have Lyα emission detected in their spectra, the emission line detected is indeed Lyα. There are initially five LAEs at $z > 7$ in the literature, imaged by NIRCam, for which Lyα is the only observed emission line (z7-GSD-3811, z8-19326, z7-13433, z7-20237 and GSDY). In order to confirm that this emission line is Lyα, we use the SED-fitting code BAGPIPES (Carnall et al., 2018) on the available NIRCam data to produce an improved photometric redshift of each LAE. One of the key diagnostics in this fitting becomes the bright [O III]emission lines, expected from high redshift objects, that fall into either the medium-band filter F410M (for $z \sim 6.8-7.6$)/F430M (for $z \sim 7.3-7.8$) or a lower flux consistent with continuum emission in this filter but a boost in the F444W filter (for $z \sim 6.8-9$). This photometric-redshift fitting results in a redshift that is consistent with the redshift of the emission lines observed being Lyα in all but two of these galaxies (z8-19326 and z7-20237). The photometric redshift, constrained by a significant boost in F410M ($z \sim 7.1$) for z8-19326, becomes inconsistent with the emission line detected being Lyα and therefore this galaxy is not included in our sample. While boosts in both F356W and F444W strongly indicate that z7-20237 is in fact at $z \sim 6$. The original *HST* photometry reported in Jung et al. (2022) also slightly favoured a $z \sim 6$ solution making the observed emission line likely to be C IV. The removal of these two galaxies as contaminants underlines the challenges in determining galaxy redshifts with just one emission line and wide-band *HST* filters pre-*JWST*. The *JWST* offers significant advancement in this area given NIRSpec's sensitivity allowing the detection of multiple lines, but also given the frequency of programs that exploit the NIRCam medium-band filters around $\sim 4\mu$m. All of the galaxies in our final sample show a clear [O III]excess in NIRCam imaging. We can therefore be confident that the emission line detected is indeed Lyα.

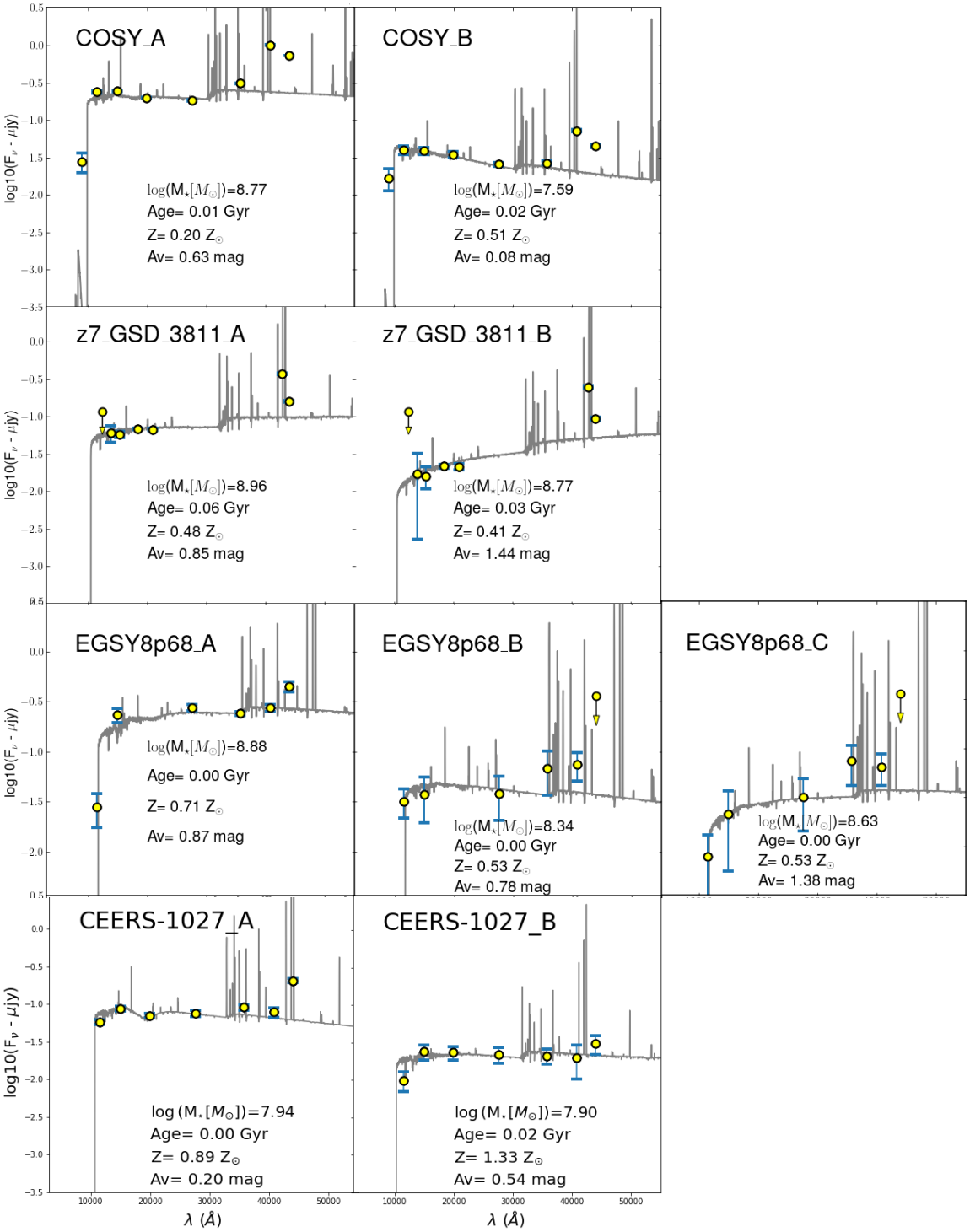


Fig. 3.3 Best SED-fits for all systems with strong Lyα. The yellow dots are the observed photometry, the grey line shows the best fit. The physical parameters of the best fit are also indicated (see text for details).

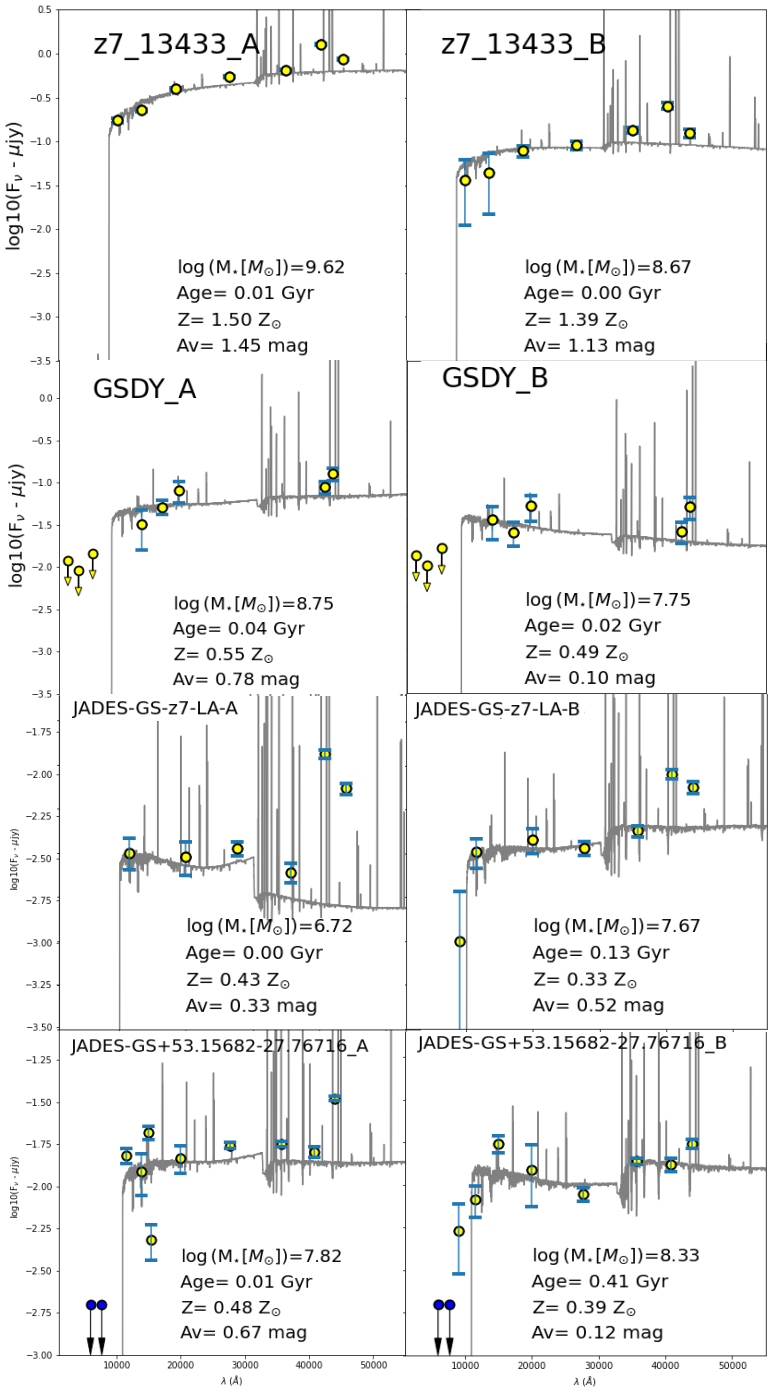


Fig. 3.4 Same as Figure 3.3.

Following this previous step, in order to confirm that all components of each system are at a similar redshift, we again employ BAGPIPES to estimate their photometric redshifts. We then compare these with the spectroscopic redshift of the main component. This fitting confirms that all the companion galaxies preferentially have high photometric redshifts, all of which are consistent with the spectroscopic redshift of the massive galaxy. In combination with the prior of the companion being closely located to a confirmed high-redshift galaxy, these become strong-photometric candidate companion galaxies. We also exploit all available ground and space based spectroscopic observations that cover the companions and use these to spectroscopically confirm the companion galaxy where possible. Below we discuss the available redshift constraints on our sample.

EGSY8p68, CEERS-1027 and JADES-GS+53.15682-27.76716 constitute the sub-sample of LAEs that have photometric candidate companions with no spectroscopic confirmation. This lack of confirmation is due to a combination of factors. The observations of CEERS-1027 and JADES-GS+53.15682-27.76716 were made with the NIRSpec MSA and the positioning of the shutter misses the companion, and as such there is no ability to place any spectroscopic constraints on these companions. EGSY8p68 has been observed by both the NIRSpec MSA and the ground-based telescope MOSFIRE/Keck. The MOSFIRE observations of the system cover the companion galaxies, however due to the seeing of the observations ($\sim 0.7''$) and the small separation between the components of the system ($\sim 0.1''$) resolving the components was not possible. Follow-up observations with the NIRSpec MSA reveal EGSY8p68 is likely an AGN given the presence of high-ionisation emission lines and a significant broadening of the Hβ emission line (Larson et al., 2023). While one of the companion galaxies lies in the shutter of this observation, the available spatial resolution and strength of the emission lines from EGSY8p68 makes it challenging to disentangle the emission from each component. However, the reported NIRSpec MSA flux is seven times fainter than that observed in the MOSFIRE spectrum (Zitrin et al., 2015). Given the MSA shutter size is significantly smaller than the MOSFIRE slit, one solution to this discrepancy in flux is that some of the Lyα emission is extended outside of the MSA slit. We note that the component EGSY8p68-C sits outside of the MSA slit position and as such this component may additionally host Lyα emission at the same redshift as the main component. Moreover, Larson et al. (2023) claim that the system is consistent with being involved in a major merger and recent NIRSpec IFU observations of EGSY8p68 spectroscopically confirm they are physical companions [Carniani et al. in prep.]. However, given these results are not yet published and the data is not yet public we continue our analysis of the redshift constraints assuming we do not have this spectroscopic confirmation.

Despite the lack of spectroscopic constraints on this sub-sample, the similarities in the SEDs of the main and companion galaxies are distinct. The galaxies in each system show similar colours, Lyman-breaks that are consistent and an emission-line boost, driven by [O III]and Hβ emission, in the same filters, as can be seen in Figure 3.3 and 3.4. This results in a photometric redshift for each of the companions that clearly favours a $z > 7$ solution, and given the emission line feature in the F444W filter, the photometric redshift is further constrained to $z = 7.6 - 9$. While the redshift of these components are very confidently constrained to within this $\Delta z = 1.4$, further constraining their redshift is challenging without spectroscopy. However, we make use of the UV luminosity function from Bouwens et al. (2015) in order to estimate the expected number of galaxies within the volume that the [O III]emission redshift constraint constrains the companion galaxies to. We note that the photometric redshift from BAGPIPES (reported in Table 3.2) is much more tightly constrained than that provided by the [O III]boost alone, however we use this wider redshift constraint ($\Delta z = 1.4$) to provide an upper limit on the expected number of galaxies. Given that we have searched for companions in a 3″x3″ region, we use this as the area for this estimate and the [O III]redshift constraints to produce a volume. We then integrate the $z = 7$ or $z = 8$ (dependent on the spectroscopic redshift of the main component) UV luminosity function down to the 5σ magnitude limit of the NIRCam images. We estimate that the expected number of galaxies in this field-of-view that are consistent with Hβ and [O III]producing the excess F444W flux and with observed UV magnitude is $0.09^{+0.16}_{-0.04}$ for JADES-GS+53.15682-27.76716, $0.007^{+0.009}_{-0.002}$ for CEERS-1027 and $0.003^{+0.006}_{-0.001}$ for EGSY8p68. It is therefore very unlikely that another galaxy that is not associated with the spectroscopically confirmed LAE would be present in this volume.

JADES-GS-z7-LA is an extremely high EW LAE (400 Å) discovered in JADES NIRSpec observations of galaxies in the GOODS-S field (Saxena et al., 2023). The presence of a photometric candidate companion galaxy to this LAE has already been reported (Saxena et al., 2023), but the companion lies outside of the NIRSpec MSA shutter, and hence is not spectroscopically confirmed. We perform the same analysis as above, noting that the [O III]emission is in the F410M filter, limiting the redshift to between $z = 6.8 - 7.6$, and find the expected number of galaxies to be $0.13^{+0.14}_{-0.05}$. Therefore, once again, it is unlikely that the observed companion is at a redshift that makes it unrelated to the spectroscopically confirmed LAE. Moreover, we note the observed excess in the F115W filter that may be driven by Lyα emission in both the main and companion galaxy (Saxena et al., 2023), supporting the conclusion that these two galaxies are at the same redshift.

COSY and z7-13433 have companions that show excess fluxes in the F410M filter driven by [O III]emission. SED-fitting of their companions strongly prefers a $z > 7$ solution,

and the presence of [O III]emission confines the redshift to $z = 6.8 - 7.6$. We again perform the same analysis as above, estimating the expected number of galaxies in the FoV. The expected number of galaxies is $0.03^{+0.03}_{-0.01}$ for COSY and $0.04^{+0.03}_{-0.01}$ for z7-13433 and thus we conclude it is very likely that the central and companion galaxies are associated with each other.

Moreover, the companions of both COSY and z7-13433 are within the slits used for the observations with X-shooter and MOSFIRE, respectively. The two components of each system are shown in Figure 3.9, where we can identify both the positive (white) and negative (black) traces of the objects, caused by the ABBA nodding pattern of the observations. We find that, for z7-13433, two emission lines are offset by ~ 3 pixels, which is consistent with the $\sim 0.5''$ offset between the two components in the slit. The Lyα emission that is coincident with the position of the companion is detected at 2.8σ and corresponds to $z = 7.479$, while the Lyα flux associated with the main galaxy is at 4σ corresponding to $z = 7.482$. Therefore, these two components are offset by ~ 125 km/s. In the case of COSY, we expect a similar spatial offset between the two components, but the $\sim 0.7''$ seeing during the observations makes distinguishing between the two components very challenging. However, there is a second, previously unidentified, component to the Lyα emission of COSY at $\sim 0.9898\mu$m. We attribute this 2.7σ emission line to COSY-B and note a velocity offset of ~ 280 km/s between COSY-A and COSY-B. While these Lyα detections are clearly tentative, given they are coincident with the expected position in the 2D spectra of strong-photometric candidate galaxies and given their proximity to confirmed LAEs this boosts their likelihood of being a true detection.

Finally, we note that recent NIRSpec IFU observations of COSY spectroscopically confirm its companion galaxy (Übler et al., 2024). However, as with EGSY8p68, given these results are not yet published and the data is not yet public we continue our analysis without this spectroscopic information.

z7-GSD-3811 and GSDY are both in the FoV of the FRESCO NIRCam WFSS survey. Their final 2D WFSS spectra, shown in Figure 3.6 show [O III]emission lines for both z7-GSD-3811, GSDY and their companions. For z7-GSD-3811 both components of the [O III]emission are present and there is a $\sim 3\sigma$, faint detection of Hβ. Moreover, [O III]emission is also detected in the companion galaxy. This emission is offset from the central galaxy by 170 ± 70 km/s. The 2D spectrum of GSDY shows a clear detection of $[\mathrm{OIII}]_{5007}$ in both the central and companion galaxies, and there is a very faint indication of $[\mathrm{OIII}]_{4959}$ and Hβ in the companion galaxy, but at a less than 3σ confidence for both. The emission detected in the central and companion galaxy are offset by just ~ 26 km/s,

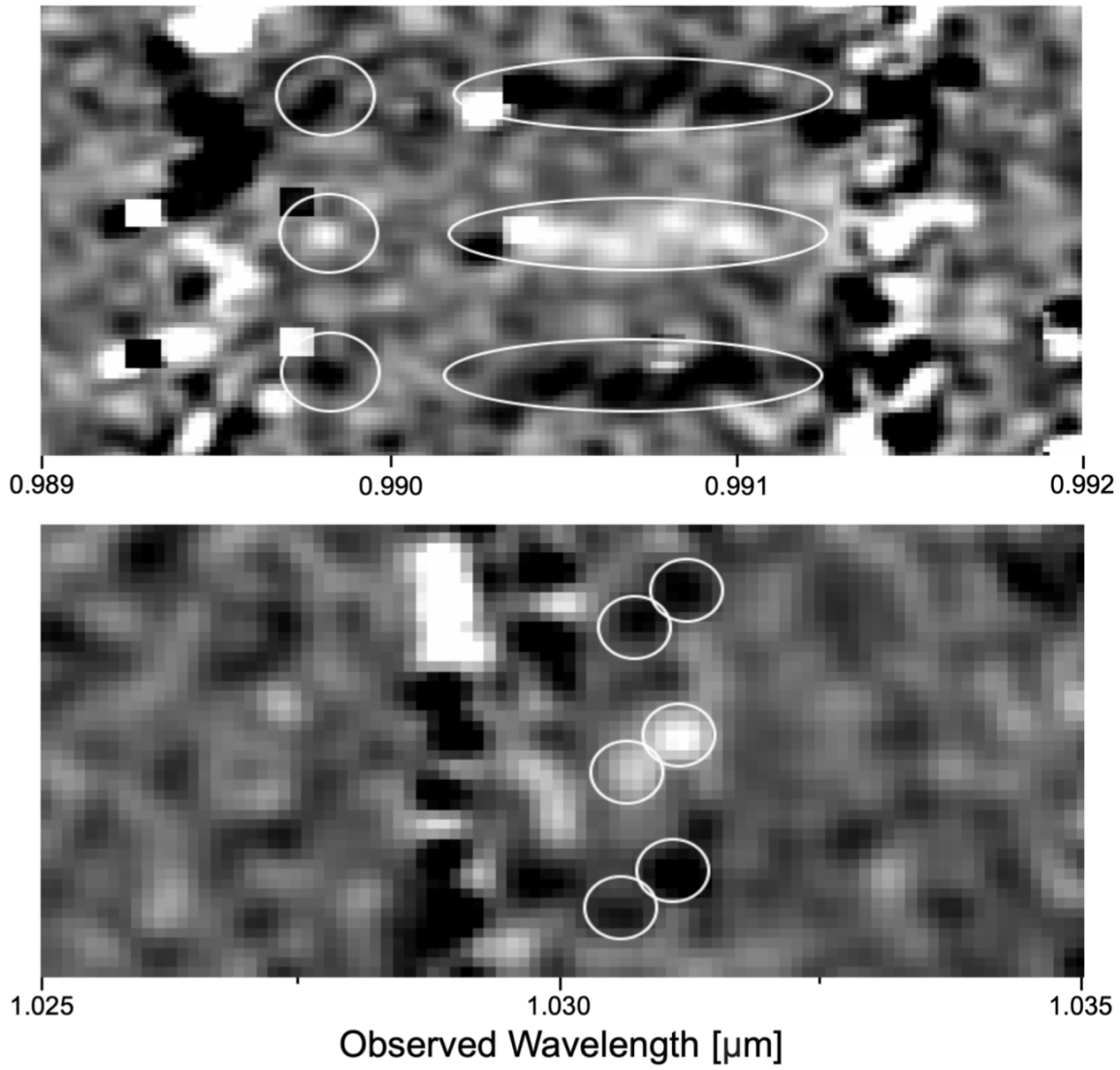


Fig. 3.5 The X-shooter (above) and MOSFIRE (below) Gaussian-smoothed 2D spectra of COSY and z7-13433, respectively. These cover the wavelength range of the previously spectroscopically confirmed Lyα emission of the system. White circular apertures indicate the proposed two components of the Lyα emission that are coincident with the expected positions of the main and companion galaxies in the slit. The positive trace of the emission is indicated by white pixels, the negative trace by black pixels, this effect of positive and negative traces is caused due to the ABBA nodding procedure employed by both instruments.

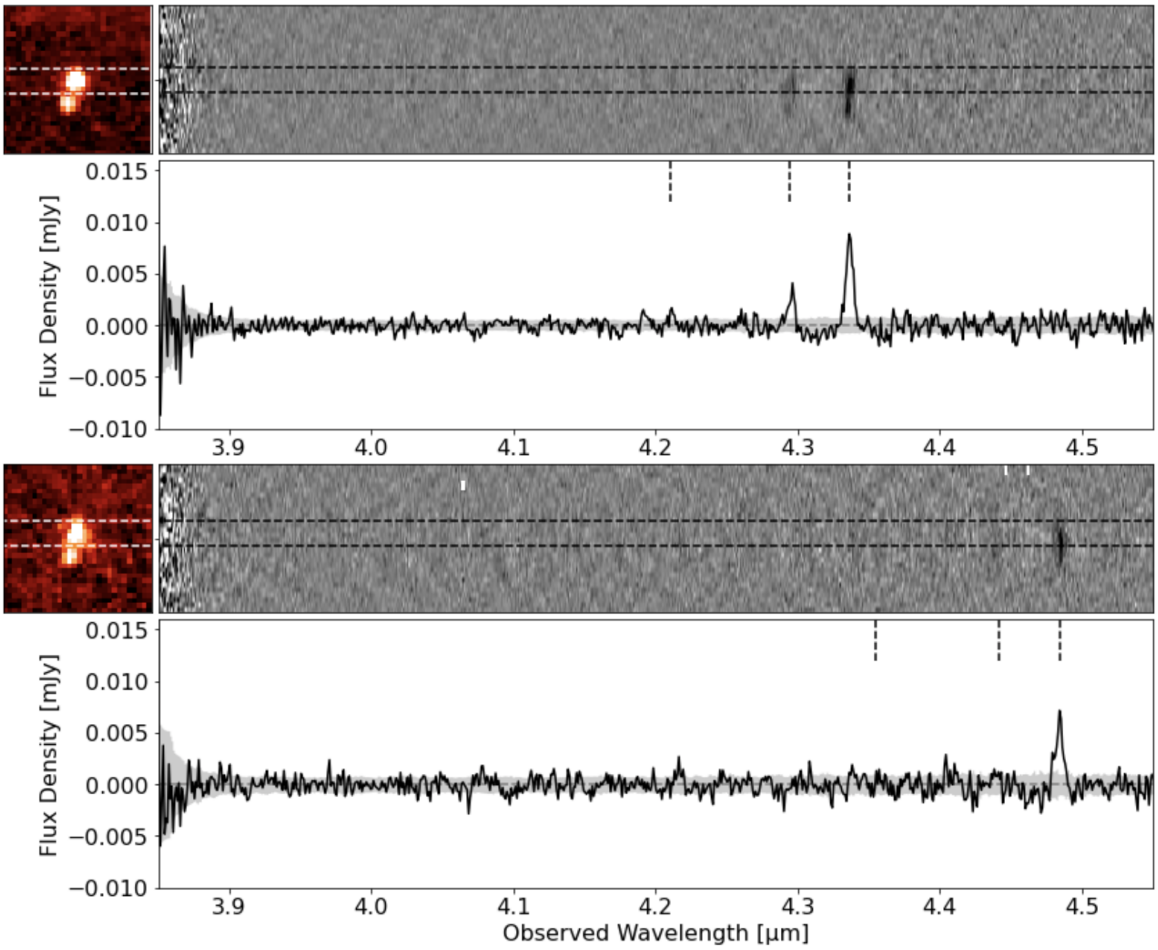


Fig. 3.6 NIRCam WFSS Grism spectra of z7-GSD-3811 (above) and GSDY (below). Within each panel, the continuum subtracted 2D spectrum (above) and 1D spectrum (below) of the central galaxy are shown. The upper left sub-panel shows the F444W image of the system where two components can be identified. The horizontal dashed lines indicate the position of the central galaxy in both the image and 2D spectrum. The three dashed lines in the 1D spectrum denote the expected positions of the [O III]doublet (central and right dashed lines) and Hβ (left dashed line) emission lines given the object's spectroscopic redshift, while the grey shaded region denotes the one-sigma noise level.

but given the low spectral resolution of the NIRCam WFSS, this is significantly smaller than the ~ 70 km/s uncertainty on this measurement.

Hβ emission can be used to estimate the intrinsic Lyα flux and hence the escape fraction of Lyα photons, however, it can also be used as a probe of the instantaneous SFR. We assume a Hα to Hβ ratio of 2.85 and Case B recombination at $T_e = 10,000$ K and $N_e = 10^4$ cm^{-3} and following Kennicutt (1998) we estimate the instantaneous SFR which is reported in Table 3.1. Given the lack of constraints on the reddening of these galaxies, we do not dust-correct the Hβ flux and as such these measured instantaneous SFRs are lower bounds.

The observation of multiple, spatially resolved components to each emission line is taken as spectroscopic confirmation of the companion galaxies. Moreover, the velocity offset observed in all of these galaxies is indicative of either a line-of-sight separation, or more likely, a difference in the line-of-sight velocities of the two objects as they interact with each other. We therefore consider this as further evidence that the systems are indeed interacting, as opposed to being two components of the same stable system.

ID	z	[OIII]$_{5007}$	[OIII]$_{4959}$	Hβ	SFR$_{\rm inst}$[M$_\odot$/yr]
z7-GSD-3811-A	7.663	8.3 ± 0.4	2.7 ± 0.4	1.0 ± 0.3	16.5 ± 4.9
z7-GSD-3811-B	7.658	4.8 ± 0.4	1.5 ± 0.3	< 0.5	< 8.2
GSDY-A	7.956	5.8 ± 0.7	1.0 ± 0.4	< 0.5	< 9.0
GSDY-B	7.957	3.1 ± 0.6	1.5 ± 0.6	0.8 ± 0.3	14.4 ± 5.4

Table 3.1 The observed redshift and fluxes of emission lines (in units of 10^{-18} erg/s/cm^2) detected in the NIRCam/WFSS spectra. The instantaneous star-formation rate (SFR) is not dust-corrected given the lack of constraints on dust in these galaxies.

GN-z11 was recently observed with the NIRSpec IFU which has revealed the presence of three regions emitting either HeII (GN-z11-C) (Maiolino et al., 2024b) or Lyα emission (GN-z11-B and a more tentative candidate) (Scholtz et al., 2024). While none of these regions show evidence of a UV continuum in deep NIRCam imaging, a tentative indication of CIV]$\lambda\lambda$1548,1551 emission in GN-z11-B and analysis of the emission in GN-z11-C (Maiolino et al., 2024b) indicate they are indeed being driven by a UV-faint stellar population in these regions. As such, we conclude that GN-z11 does have spectroscopically-confirmed companion galaxies. Moreover, the presence of multiple LAEs in the local vicinity of GN-z11 suggests that peculiar velocity is not playing a vital role in the escape of Lyα through the neutral IGM (see Scholtz et al. 2024 for a detailed analysis of the Lyα emission present in this system). Moreover, the density of objects in this field makes GN-z11 a likely protocluster core (Scholtz et al., 2024).

3.3.2 Galaxy properties

We then use BAGPIPES, with the redshift fixed to that of the spectroscopic redshift of the system, using different SFHs (constant, delayed, burst and burst+constant) to fit the SED of all central and companion galaxies, as seen in Figure 3.3 and 3.3. The best fit model is then obtained by taking the SFH that leads to the smallest Bayesian Information Criterion (BIC) as described in Laporte et al. (2021). We then report the galaxy properties returned from this method in Table 3.2.

ID	z	$\log(M_*\,[M_\odot])$	SFR $[M_\odot/yr]$	Separation [kpc]	$L_{Ly\alpha}$ $[10^{43}$ erg/s]	$f_{esc}(Ly\alpha)$	Reference
COSY-A	*7.142*	$8.77^{+0.02}_{-0.02}$	$5.91^{+0.27}_{-0.25}$	–	1.37	< 0.1	Roberts-Borsani et al. (2016),Witten et al. (2023)
COSY-B†	*7.142*	$7.59^{+0.13}_{-0.11}$	$0.39^{+0.14}_{-0.08}$	2.8	–	–	
JADES-GS-z7-LA-A	*7.278*	$6.72^{+0.14}_{-0.12}$	$1.13^{+0.05}_{-0.03}$	–	0.15	0.96 ± 0.22	Saxena et al. (2023)
JADES-GS-z7-LA-B	$7.27^{+0.07}_{-0.06}$	$7.68^{+0.18}_{-0.19}$	$2.44^{+0.82}_{-0.47}$	1.2	–	–	
z7-13433-A	*7.482*	$9.62^{+0.01}_{-0.01}$	$42.19^{+1.09}_{-1.00}$	–	1.00	–	Jung et al. (2022)
z7-13433-B	*7.478*	$8.67^{+0.03}_{-0.04}$	$4.65^{+0.33}_{-0.36}$	2.6	–	–	
z7-GSD-3811-A	*7.661*	$8.96^{+0.12}_{-0.13}$	$9.25^{+2.79}_{-2.39}$	–	0.386	0.22 ± 0.08	Song et al. (2016), this work
z7-GSD-3811-B	*7.658*	$8.77^{+0.21}_{-0.21}$	$5.88^{+3.73}_{-2.28}$	1.0	–	–	
CEERS-1027-A	*7.819*	$7.94^{+0.05}_{-0.04}$	$0.87^{+0.10}_{-0.08}$	–	0.432	0.085 ± 0.018	Tang et al. (2023)
CEERS-1027-B	$8.07^{+0.12}_{-0.14}$	$7.90^{+0.28}_{-0.21}$	$0.80^{+0.74}_{-0.30}$	8.6	–	–	
GSDY-A	*7.957*	$8.75^{+0.21}_{-0.27}$	$17.80^{+4.56}_{-5.02}$	–	0.19	> 0.11	Roberts-Borsani et al. (2023b), this work
GSDY-B	*7.958*	$7.75^{+0.21}_{-0.28}$	$2.73^{+1.02}_{-0.73}$	2.2	–	–	
JADES-GS-A‡	*7.982*	$7.82^{+0.11}_{-0.09}$	$3.97^{+1.78}_{-0.88}$	–	0.056	0.09 ± 0.01	Jones et al. (2024)
JADES-GS-B‡	$8.11^{+0.10}_{-0.08}$	$8.33^{+0.14}_{-0.16}$	$10.47^{+5.16}_{-3.46}$	1.5	–	–	
EGSY8p68-A	*8.683*	$8.88^{+0.10}_{-0.18}$	$7.62^{+1.87}_{-2.54}$	–	1.59	< 0.1	Zitrin et al. (2015),Witten et al. (2023)
EGSY8p68-B†	$8.40^{+0.37}_{-0.27}$	$8.34^{+0.28}_{-0.36}$	$2.22^{+1.99}_{-1.25}$	0.5	–	–	
EGSY8p68-C†	$8.74^{+0.50}_{-0.46}$	$8.63^{+0.26}_{-0.40}$	$4.26^{+3.49}_{-2.55}$	0.6	–	–	
GN-z11-A	*10.603*	$9.1^{+0.3}_{-0.4}$	21^{+22}_{-10}	–	0.324	$0.03^{+0.05}_{-0.02}$	Bunker et al. (2023b)
GN-z11-B	*10.62*	*	*	1.6	–	–	Scholtz et al. (2024)
GN-z11-C	*10.603*	*	*	2.5	–	–	Maiolino et al. (2024b)

Table 3.2 Physical properties of all the systems studied in this paper. The columns are : (1) spectroscopic (italic) or photometric redshift of the candidate, (2) stellar mass, (3) Star formation rate averaged over at most the last 100 Myr, (4) separation between the main component (A) and each companion, (5) luminosity of observed Lyα emission (6) escape fraction of Lyα photons and (7) reference for the original spectroscopic confirmation and the Lyα escape fraction measurement. * The companions of GN-z11 are spectroscopically confirmed (Scholtz et al., 2024; Maiolino et al., 2024b) however given that they are very UV-faint we cannot estimate their properties. † The companions of both COSY and EGSY8p68 have recently been spectroscopically confirmed by NIRSpec IFU observations (Übler et al., 2024);[Carniani et al. in prep.]. ‡ JADES-GS+53.15682-27.76716 has been abbreviated to JADES-GS.

3.3.3 Close companion fraction of non-LAEs

To confirm that the presence of a close companion is the primary factor governing the visibility of Lyα, we determine the fraction of companions seen in a mass-matched sample of $z > 7$ galaxies, we take 30 galaxies with stellar masses from $7.4 < \log(M_*[M_\odot]) < 9.3$ from Laporte et al. (2022); Tang et al. (2023); Harikane et al. (2022); Bouwens et al. (2022). We find that 47% of these galaxies have photometric-candidate companions within $5''$ of the central galaxy. Given some of these photometric candidates will be low-redshift contaminants, this companion fraction represents an upper bound. Moveover, most of the companion galaxies have not been spectroscopically followed-up and as such, we do not have constraints on their Lyα emission, thus further making this fraction an upper limit.

Unfortunately, examples of spectroscopically confirmed $z > 7$ non-Lyα emitting galaxies that have been observed by high-resolution spectrographs are rare in the literature. As such it is challenging to probe the companion fraction for known non-LAE galaxies. Moreover, the challenges of slit/aperture spectroscopy mean that it is incredibly rare to have spectroscopic information on the local environment of such galaxies and hence, while a galaxy may be confirmed as non-LAE, there is no information regarding the Lyα emission of any companion that may or may not be present.

The recent JADES data release of NIRSpec and NIRCam observations in GOODS-S includes seven $z > 7$ galaxies, observed using the NIRSpec high-resolution G140M grating, that show no Lyα emission (Jones et al., 2024). This does not provide any Lyα information on any would-be companion galaxies, however, the companion fraction for this sample, $\sim 43\%$, is once again consistent with both our larger sample, and that of Puskás et al. (in prep.). We therefore conclude that the 100% companion fraction of our LAE sample is clearly atypical compared to the companion fractions of samples not selected for Lyα emission.

3.3.4 Close companions are necessary but not sufficient for Lyα emission

We note the recently reported $z = 9.3$ merging system (Boyett et al., 2024) with an absence of detected Lyα emission. The reported specific-star-formation rate of this galaxy (sSFR ~ -8) is consistent with the high sSFR seen in our sample of merging LAEs, and therefore, one may naively posit that this is a contradiction to our results. We argue that non-detection of Lyα emission from a merging system is entirely reasonable and in fact our simulation modelling predicts this. While we see significant boosts in the SFR and hence

intrinsic Lyα emission from merging systems, the escape of Lyα emission is only possible when the escape fraction of Lyα emission out of both the host galaxy and the surrounding IGM is non-zero. We explore the effect of viewing angle on the observed Lyα emission escaping our simulated merging systems and find that while the majority of viewing angles facilitate the escape of Lyα, several viewing angles exist for which the only Lyα emission escaping the galaxy is diffuse and is therefore unobservable. Moreover, without the presence of a large ionised bubble, the neutral IGM will not facilitate the escape and observation of Lyα emission. Ultimately we conclude that close companions are necessary but not sufficient for the observation of Lyα from galaxies deep in the epoch of reionisation. A combination of a large-scale ionised bubble and a preferable line-of-sight, stripped of neutral hydrogen by the interaction with the companion, are also required.

3.4. *Azahar* SIMULATION

To reinforce the observational evidence supporting the idea that ongoing interactions drive an increase in the detectability of Lyα emission during the epoch of reionisation, we explore comparable galaxy mergers using the novel *Azahar* simulation suite (Martin-Alvarez et al., in prep.). The new *Azahar* simulations are a suite of multiple models spanning various combinations of canonical hydrodynamics, magnetic fields, radiative transfer, and cosmic rays physics with the aim of understanding their effects on galaxy formation as well as their complex interplay. The simulations are generated using the magneto-hydrodynamical code RAMSES (Teyssier, 2002), which employs a constrained transport method for divergence-less evolution of the magnetic field (Teyssier et al., 2006). The main target of study for *Azahar* is a massive spiral galaxy with $M_{\rm halo}\,(z=1) \sim 2{\cdot}10^{12}\,M_{\odot}$ in a relatively large zoom-in region, approximately 8 cMpc across along its largest axis. *Azahar* has a maximum spatial resolution in cells with a full cell-width of $\Delta x \sim 20\,$pc (or equivalently, an approximate cell half-size or radius of 10 pc), and refinement is triggered whenever its size is larger than a quarter of its Jeans length, or it contains a total mass larger than $8\,m_{\rm DM}$, where $m_{\rm DM} \sim 4.5 \cdot 10^5 M_{\odot}$ is the mass of dark matter particles in the zoom region. The stellar mass particle is $m_* \sim 4 \cdot 10^4 M_{\odot}$. *Azahar* follows the set of physics presented by the pathfinder Pandora (Martin-Alvarez et al., 2023). We provide here a brief summary, and refer the reader to that reference and Martin-Alvarez et al. (in prep.) for further details.

In this work, we investigate one of the most complete simulations in the *Azahar* suite, the RTnsCRiMHD model. In this manuscript, we refer to the RTnsCRiMHD simulation simply as *Azahar*. Radiative cooling is modelled both above and below 10^4K (Ferland

et al., 1998; Rosen & Bregman, 1995). The model adopts a magneto-thermo-turbulent prescription for star formation (Kimm et al., 2017; Martin-Alvarez et al., 2020), mechanical supernova (SN) feedback (Kimm & Cen, 2014) and astrophysically-seeded magnetic fields through magnetised SN feedback (Martin-Alvarez et al., 2021). This feedback injects 1% of the SN energy as magnetic energy, and intends to reproduce the approximate magnetisation of SN remnants. This particular simulation of *Azahar* also includes cosmic rays modelled as an energy density allowed to anisotropically diffuse, evolved with an implicit solver (Dubois & Commerçon, 2016; Dubois et al., 2019). In *Azahar*, cosmic rays are exclusively sourced by SN feedback, with each event injecting 10% of their total energy as cosmic rays. We assume a constant diffusion coefficient for cosmic rays of $\kappa_{\rm CR} = 3 \cdot 10^{28}\,{\rm cm}^2\,{\rm s}^{-1}$, and do not account for cosmic ray streaming in this model. This model also includes on-the-fly radiative transfer (Rosdahl & Teyssier, 2015), with a configuration similar to that of the SPHINX simulations (Rosdahl et al., 2018), featuring three energy bins for radiation spanning the ionisation energy intervals from 13.6 eV to 24.59 eV (HI ionisation), 24.59 eV to 54.42 eV (HeI ionisation), and above 54.42 eV (HeII ionisation). Finally, recent work has shown that when expanding simulation models to account for cosmic ray feedback physics at comparable resolutions, the escape fractions of LyC photons from galaxies is lower than in their absence (Farcy et al., 2022).

Therefore, *Azahar* features a realistic (yet computationally taxing) ISM model that considerably approaches the resolution required to converge in the propagation and escape of ionising photons from galaxies (Kimm & Cen, 2014). We note that, while close, our resolution still falls short from that regime and that even more sophisticated ISM models (Katz et al., 2022) may be important for a complete understanding of photon propagation in the ISM. Opportunely, a considerable part of our results relies on gas being ejected from galaxies during merger events, which implies that our estimate of LAE detectability during these events will be more resilient to these caveats.

This simulation suite was performed before we obtained the NIRCam data, but remarkably, we find many simulated galaxies that match the photometry and the spatial geometry of our objects. We post-process our simulations with the publicly available RASCAS code (Michel-Dansac et al., 2020) to account for both the production and resonant scattering of Lyα photons as well as scattering or absorption of Lyα photons by dust (Laursen et al., 2009).

3.4.1 RASCAS post-processing

We post-process the simulation with the publicly available, massively parallel code RASCAS (Michel-Dansac et al., 2020), for modelling the Lyα emission and resonant scattering

in our simulations. RASCAS accounts for two sources for the Lyα emission, recombination and collisional excitation. For recombination, RASCAS adopts the case B recombination coefficient from Hui & Gnedin (1997), and the fraction of recombinations producing Lyα photons from Cantalupo et al. (2008). For collisional excitation, RASCAS uses the fitting function for the collisional excitation rate from level 1s to 2p from Goerdt et al. (2010). We cast $N_{\rm MC} = 10^6$ Monte Carlo photon packets to sample the real Lyα photon distributions, from recombination and collisional excitation, respectively. We sample $N_{\rm MC} = 10^7$ photon packets for the images along the LOS in Figure 3.5.

After each scattering event, the Lyα photon changes its frequency and direction according to a phase function as implemented in RASCAS. RASCAS adopts the phase function in Hamilton (1940); Dijkstra & Loeb (2008) for the scattering of Lyα photons around line centre and Rayleigh scattering for Lyα photons in the line wing. At high H I column density, Lyα photons will scatter many times locally until they shift in frequency large enough so that they can have a long mean free path. To reduce the associated computation cost, RASCAS adopts a core-skipping mechanism (Smith et al., 2015) to transit the photons to line wings while avoiding local scattering in space. RASCAS also implements the recoil effect and the transition due to deuterium with an abundance of D/H = 3×10^{-5}. The dust distribution is modelled according to the dust model described in Laursen et al. (2009). Dust can either scatter or absorb Lyα photons. The probability of scattering is given by the dust albedo $a_{\rm dust} = 0.32$ following Li & Draine (2001). The dust scattering with the Henyey-Greenstein phase function (Henyey & Greenstein, 1941) and the asymmetry parameter is set to $g = 0.73$ following Li & Draine (2001). We generate synthetic images and spectra with a peeling algorithm described in Yusef-Zadeh et al. (1984); Wood & Reynolds (1999); Costa et al. (2022), along 12 lines-of-sight uniformly sampled using the healpix algorithm (Górski et al., 2005).

Finally, we note that the dust distribution has a large effect on the Lyα radiative transfer as well as the [O III]and Hβ emission that is used to infer the intrinsic Lyα emission from observations. In Figure 3.7, we show how reducing the amount of dust changes the observed Lyα images. We see the dust is able to completely mask the Lyα emission of some individual galaxies as resonant scattering of H I increases the probability of Lyα photons encountering dust grains. The dust can suppress the Hα emission by up to 20% around the center of the three merging galaxies, if we assume intrinsic Lyα can be converted directly to intrinsic Hα emission. We therefore conclude that inclusion of dust modelling, as done here, is fundamental for understanding observed Lyα emission. However, our results regarding the boosted intrinsic Lyα emission in star formation bursts

and enhanced Lyα escape fraction in mergers, due to the channels cleared of local neutral hydrogen are robust to the assumed dust model.

We further consider the Lyα profiles escaping along the line-of-sight shown on the upper panel of Figure 3.5 at $z = 7.3$ after applying a reasonable IGM damping wing from Keating et al. (2024). We note that the shape of the Lyα profile is very complex and sensitive to both the LOS (Blaizot et al., 2023) and IGM attenuation (Witten et al., 2023). As such, further analysis of this is needed to constrain the expected Lyα profiles from simulations requiring a larger simulated galaxy sample with careful IGM attenuation modelling as well as a larger observational sample to compare to, which is beyond the scope of this work. However, as an immediate consistency test, we take the Lyα profile from our simulated system at $z = 7.3$ after IGM attenuation (Keating et al., 2024) and compare the velocity offset and FWHM to that reported in our observational comparison, EGSY8p68. We find that the observed Lyα profile (Larson et al., 2023) is well matched with the simulated profile's velocity offset of ~ 200 km/s and FWHM of ~ 200 km/s, and shows a clear asymmetric line shape as observed.

3.4.2 Galaxy selection and measurements in the simulations

In order to select galaxies in our simulation, we employ the HALOMAKER software (Tweed et al., 2009) to detect and characterise dark matter halos. We identify the three progenitor galaxies of interest (as well as the galaxy merging with the main progenitor at $z \sim 8.1$) and follow them through time by tracking their innermost stellar particles. Their centres are determined using a shrinking spheres algorithm applied to their stellar component (Power et al., 2003), and they are assigned their corresponding halo as obtained by HALOMAKER. In order to select the system for study, we broadly reviewed the evolution of all galaxies with stellar masses $M_* > 10^7 M_\odot$ in the redshift interval $z \in [9.0, 6.0]$. The most promising candidate, the main progenitor of the main *Azahar* galaxy was selected for further investigation.

In order to assign measurements to individual galaxies, we measure values within their galactic region, defined by the radius $r_{\rm gal} < 0.2\, r_{\rm rvir}$. During the merging stages, we employ $r = \min[r_{\rm gal}, 0.45 D_{ij}]$, where D_{ij} is the distance between the i and j progenitors, down to a minimum distance of 1.5 kpc.

We have briefly reviewed the observability of lower mass pairs within the simulated domain. While these present a similar behaviour during mergers to that of our main studied system, their observability and that of other galaxies with similar masses remains low. This is due to their low stellar masses generating low luminosities. Their low stellar masses also make them more vulnerable to disruption during powerful star formation

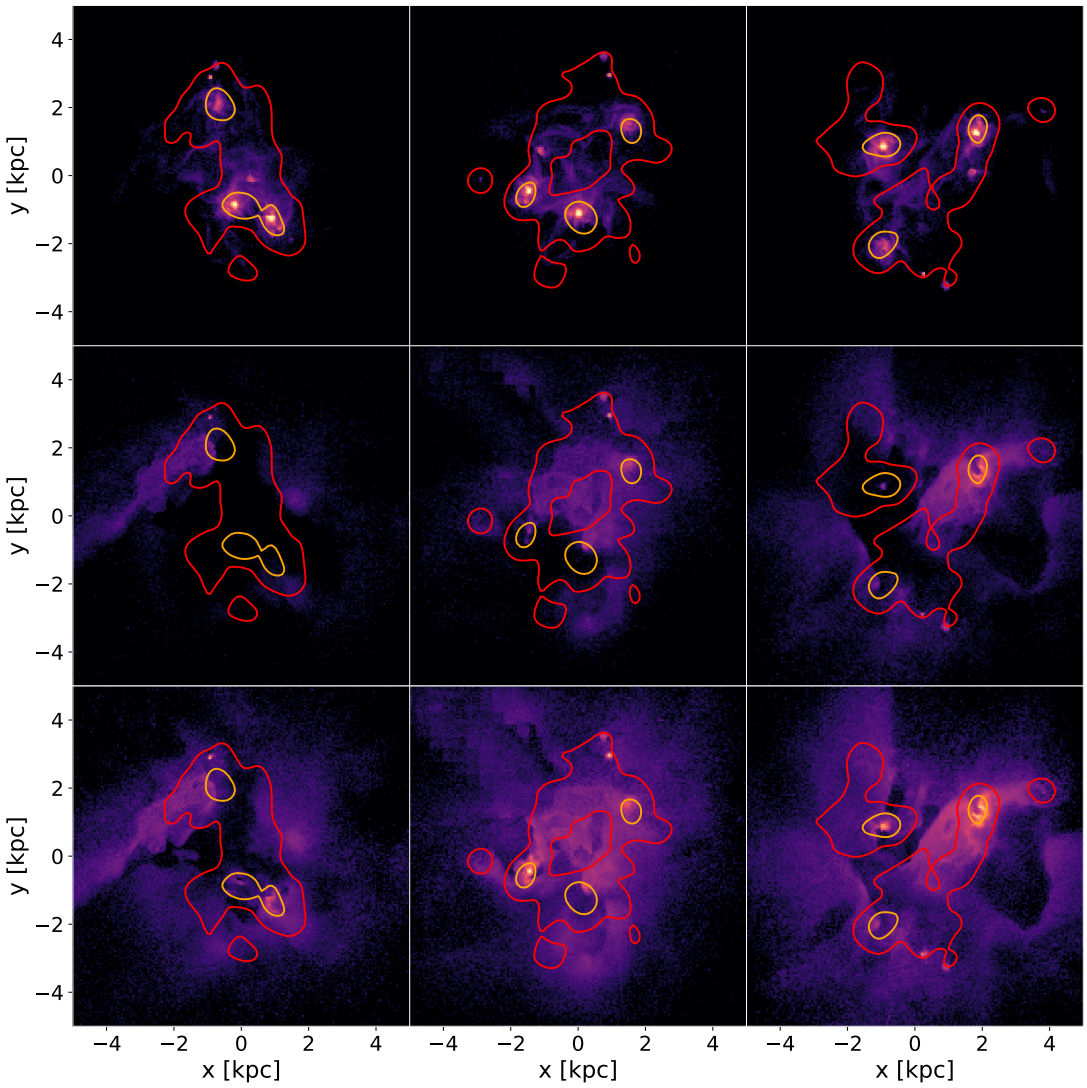


Fig. 3.7 Dust opacities superimposed on synthetic Lyα images at $z = 7.3$. The top row shows the intrinsic Lyα emission, the middle row shows the scattered Lyα emission generated with a default dust model, and the bottom row shows the scattered Lyα emission with practically no dust (dust opacity reduced by a factor of 10^6). The three different columns show the galaxy from three different directions. Red and orange iso-contours correspond to a dust optical depth of 0.2 at the Lyα wavelength and Hα wavelength, correspondingly. Note that the dust does not act as a screen for the Lyα emission because of the resonant nature of Lyα scattering.

bursts, which increases their escape fractions (Rosdahl et al., 2022). Despite this, most isolated systems present quiet star formation histories and in the absence of mergers, have low observabilities throughout their evolution.

3.4.3 Mergers in the *Azahar* simulations

Due to a substantial overdensity of galaxies in our simulated volume, the main progenitor undergoes repeated mergers with multiple other galaxies brought in by the cosmic filaments, often involving several companions or mergers in rapid succession as has been observed in Asada et al. (2023). To highlight one such merger occurring at $z \sim 7.3$, the four rightmost panels of Figure 3.8 show different projections of *Azahar*. This particular interaction features three merging galaxies that will constitute the main progenitor of the spiral galaxy formed by $z \sim 1$. These are very good analogues to the EGSY8p68 observations shown on the left (modelling of the observations with GALFIT in Figure 3.1 confirms the presence of three components), with very similar individual galaxy sizes, mutual distances and merger configuration. Specifically, the three simulated galaxies have stellar masses of $M_{*,1} = 2 \cdot 10^8\, M_\odot$, $M_{*,2} = 3 \cdot 10^8\, M_\odot$, and $M_{*,3} = 8 \cdot 10^7\, M_\odot$ at this redshift, which are comparable to the stellar masses of the EGSY8p68 galaxies (see Table 3.2). Furthermore, the simulated SFRs vary between $1 - 10\,\mathrm{M_\odot yr^{-1}}$ (shown in Figure 3.9) and are consistent with the SFRs of the EGSY8p68 galaxies ($7.62\,\mathrm{M_\odot yr^{-1}}, 2.22\,\mathrm{M_\odot yr^{-1}}, 4.26\,\mathrm{M_\odot yr^{-1}}$). We also consider the total Lyα luminosity of the simulated system and find that, while it varies, its average value is $\sim 10^{43}$ erg/s (shown in Figure 3.9), in agreement with the value for EGSY8p68, reported in Table 3.2. Finally we note that simulated volume hosts a substantial overdensity of galaxies, which is consistent with the local environments of the observed LAEs in our sample (as discussed earlier), including EGSY8p68 (Leonova et al., 2022). We thus conclude that the simulated system matches all key observed photometric and spectrosopic properties of EGSY8p68 (see also Methods).

The observations shown on the left clearly indicate the superior resolution and sensitivity of *JWST* (top panel) over *HST* (bottom panel). The tri-component nature of EGSY8p68 is entirely unresolved in existing *HST* observations, but clearly identifiable in the F150W NIRCam imaging. Panel (*a*) shows the large-scale view (150 kpc across) of the three filaments encompassing the merging system, where large amounts of HI gas (blue) along the filaments are feeding the star formation in these systems, resulting in significant NIRCam F150W emission (yellow-white). This active star formation is driving ionised hydrogen bubbles (grey) away from the filaments and into the low density regions. Panel (*b*) shows the simulated *JWST*-like 150W observation (with dust extinction modelled as an absorption screen along the line-of-sight), where we select the line-of-sight to illustrate

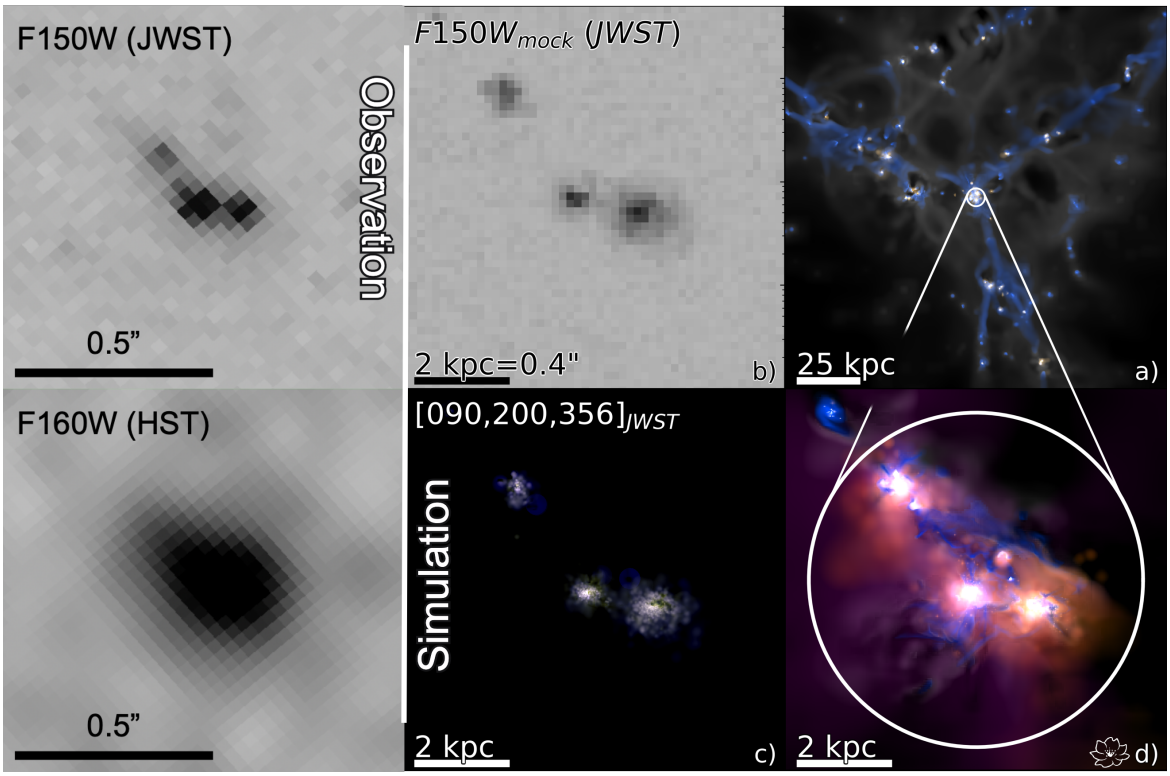


Fig. 3.8 (Left panels:) NIRCam F150W (above) and *HST* F160W (below) imaging of the LAE EGSY8p68. The NIRCam imaging reveals three components to the system that were previously unresolved by *HST*. (Right panels:) Analogue galaxy merger from the *Azahar* simulation: *a)* large-scale view of the filaments encompassing the galaxy merger (blue: HI density, grey-black: HII density, yellow-white: F150W intensity); *b)* simulated NIRCam F150W observation of the *Azahar* merger; *c)* fully resolved simulated observation in the NIRCam filters; and *d)* hydrogen gas and ionising radiation properties (blue: HI density, grey-black: HII density, yellow-white: F150W intensity, purple: LyC radiation energy density, orange: HeII ionising radiation energy density). The appearance of the observed system is very well reproduced by an ongoing merger of three galaxies, with an extended escaping LyC radiation field.

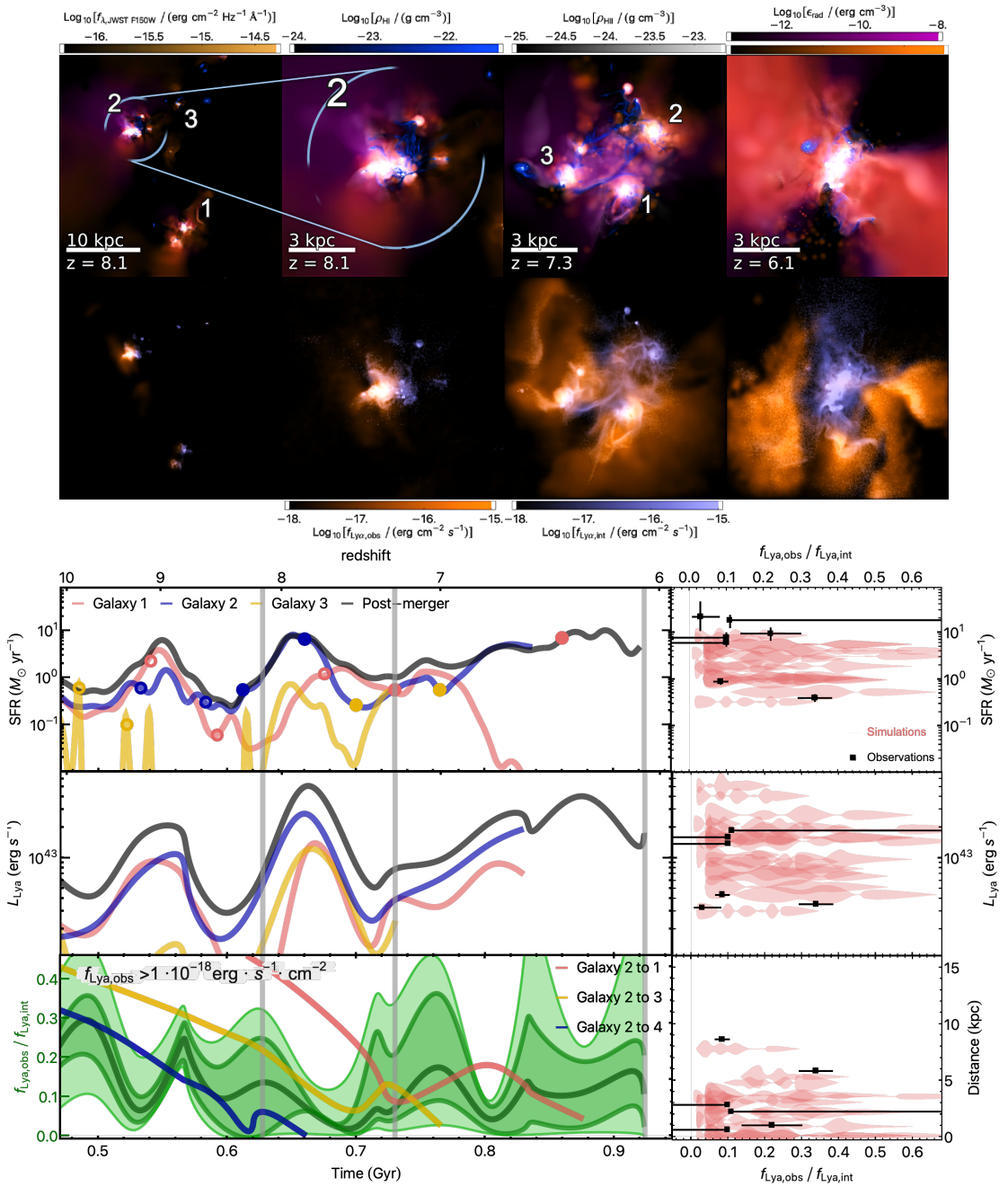


Fig. 3.9 (Rows 1-2) The time evolution of our simulated merger system at $z = 8.1$ (close up and large scale view), 7.3 and 6.1. First row: maps of HI density (blue), HII density (grey-black), NIRCam F150W intensity (yellow-white), LyC radiation energy density (purple) and HeII ionising radiation energy density (orange). Second row: the RASCAS intrinsic Lyα flux (blue) and scattered Lyα flux (orange). (Rows 3-5) The SFR for each of the three galaxies and the final system (black line). Dot symbols show mergers with companions (filled markers) and with other secondary systems (open markers); the intrinsic Lyα luminosity of the three galaxies and the entire system; the fraction of escaped Lyα luminosity filtered for pixels with fluxes higher than a given limit (median (dark green curve), quartiles (darker band), and min-max of the distribution across 12 lines-of-sight (lighter band)). Distance between the main progenitor and its companions is shown as well. The grey vertical lines in Rows 3-5 correspond to the redshifts shown in Rows 1-2. The rightmost panels show key physical relations for the observations (black) and the simulations (pink).

the resemblance with EGSY8p68. Note that the simulated merging system is displayed at a slightly earlier stage of the merger with respect to EGSY8p68, with our simulated galaxies approaching the likely physical separation of the observed galaxies at $z \sim 7$. Interestingly, the mock F150W emission reveals compact galactic cores analogous to these *JWST* observations, with diffuse emission (blending with the background) emerging from the extended, low density stellar discs. Panel (*c*) shows a synthetic colour-composite observation using the *JWST* filters at the full resolution of our simulation. While the two companion galaxies are more diffuse, with stars actively forming in their discs, the main galaxy has a much more compact core due to the preceding rapid growth through repeated mergers. Panel (*d*) is a colour composite of various simulated properties relevant for Lyα detectability. The emission from LyC ionising photons (purple) and HeII ionising photons (orange) is distributed on a scale of a few (physical) kpc. Importantly, there is a clear separation of the stellar emission and the HI gas during the merger event, with the ionising radiation escaping from star forming regions.

To provide a quantitative confirmation and in-depth physical understanding of this effect, we explore in Figure 3.9 the time evolution of the three merging galaxies in detail. In the top row the colours encode the same quantities as in panel (*d*) of Figure 3.8, while in the second row we show the intrinsic Lyα emission (blue) and resonantly scattered Lyα emission (orange), produced by post-processing our simulations with RASCAS. The first column shows a larger scale view of the main progenitor (galaxy 2) and its approaching companions at $z \sim 8.1$. The second column shows a close up views of the main progenitor at $z \sim 8.1$ undergoing a minor merger, with a complex topology of neutral gas and channels through which Lyα photons are escaping. The third column shows the ongoing merger of our galaxies 1, 2 and 3 at $z \sim 7.3$ (with a different orientation with respect to Figure 3.8), where bursty star formation drives ionised channels through the ISM, with a significant amount of gas also tidally stripped from the galaxies. The fourth column shows a post-merger view of the newly formed disc-dominated galaxy at $z \sim 6.1$. Here, the HI is confined to the merger remnant galaxy, and the radiation can only escape through the channels opened by collimated, cosmic ray-driven bi-conical outflows. Focusing now on the Lyα radiation, all three galaxies have significant intrinsic Lyα emission due to their ongoing star formation. However, at this particular viewing angle, due to absorption by dust, Lyα scattered emission is only present around merging galaxy 2 at $z \sim 8.1$ and around galaxies 1 and 3 at $z \sim 7.3$. At $z \sim 7.3$, the remaining emission, not absorbed by dust, is very diffuse on large scales and hence not observable.

To understand what drives the observable Lyα emission, in the third row we show the SFR evolution for our three merging galaxies, together with the SFR history of the merger

remnant at $z \sim 6.1$. SFRs of all three galaxies undergo repeated cycles of bursts, driven by numerous mergers ('major' mergers [filled dots] and 'minor' mergers [empty dots]) and a high SFR 'plateau' during the final merger episode from $z \sim 7$ to 6. This is mirrored in the evolution of the intrinsic Lyα emission of each galaxy as well as the total emission integrated across our three simulated galaxies, as shown in the fourth row, where there is a clear correspondence between the SFR peaks and peaks of intrinsic Lyα. *Our simulations hence demonstrate that high SFR triggered by gas-rich mergers results in brighter intrinsic Lyα emission, which enhances the probability of these systems being detected.*

To constrain this more robustly, we select an observable Lyα flux threshold of 10^{-18} erg s^{-1} cm^{-2} consistent with our observations, and compute the ratio of scattered to intrinsic Lyα emission over the spatial extent of our mock observations of the galaxies. This is shown in the fifth row, where the dark green curve denotes the median of the ratio of scattered to intrinsic Lyα flux and the shaded bands (quartiles and minimum-maximum values) show its distribution computed over 12 different lines-of-sight. As the fraction of escaping Lyα emission has a number of high peaks we conclude that the Lyα emission should be detectable for a number of the redshifts studied here. Interestingly, during close galactic encounters (as shown by the distance trajectories of three main galaxies shown in the same panel by coloured lines) and mergers, *significant gas tidal stripping and bursty star formation feedback leads to the opening of more low-column density sight-lines, hence facilitating the escape of Lyα radiation.* This effect can be clearly seen when comparing the most important merger instances to the peaks (high-tails) in the distribution of Lyα escape fraction. Our findings regarding higher intrinsic Lyα emission and enhanced Lyα escape fraction in mergers are robust to the assumed dust model (see Methods and Figure 3.7). It is worth noting that inclusion of AGN feedback in our simulations would only reinforce our findings, as expected black hole fuelling during mergers would power AGN (Costa et al., 2022).

The rightmost panels of Figure 3.9 show various physical relations for the observations (black points) and the simulations (pink 'violin' symbols, denoting the distribution of Lyα escape fraction over 12 lines-of-sight), using the same flux filter as above, demonstrating a good correspondence between our simulations and our observed LAEs. Interestingly, both for observations and simulations we do not find any clear correlations between the Lyα escape fraction and any other galaxy properties (distance, SFR, stellar mass), further corroborating our finding that 'favourable' lines-of-sights with evacuated neutral hydrogen channels lead to directional Lyα photon leakage.

3.4.4 Simulated ionised bubble evolution in the neutral IGM

In the high resolution patch of the Universe probed by our simulation ($\lesssim 8$ Mpc on a side), centred on the main progenitor of our $z \sim 1$ galaxy, the first ionised bubbles emerge at $z \sim 20$. By $z \sim 15$ several ionised bubbles exist that are a few (physical) kpc in radius, and rapidly begin merging as redshift approaches $z \sim 10$. This drastically increases the mean free path of ionising photons, with the coalesced bubble reaching scales of 100 kpc. The resulting main bubble continues to grow, reaching a radius ~ 0.5 Mpc at $z \sim 7.3$.

For intrinsically bright LAEs, as studied here, a $\lesssim 0.5$ Mpc local ionised bubble may be sufficient for these sources to be detectable (Hayes & Scarlata, 2023). We further note that due to ongoing mergers our simulated galaxies have considerable relative peculiar velocities of up to 230 km s^{-1}. Sufficiently large velocities relative to the intervening neutral IGM gas may reduce damping wing suppression for particular lines-of-sight, hence requiring smaller local ionised bubbles to facilitate the detection of Lyα. Our simulation also does not consider star formation outside of the high-resolution zoom-in region. Hence, our simulated ionised bubbles could be even larger in size, e.g. due to the presence of a neighbouring, large-scale overdensity or large-scale clustering (Weinberger et al., 2019) on spatial scales beyond our zoom-in region.

Nevertheless, knowing from our observations that the LAEs sit in sufficiently large ionised bubbles, we primarily focus on the probability of Lyα photons escaping the host galaxy and its intermediate surroundings. We find that a large amount of gas remains neutral within the local ionised bubble, actively feeding the galaxies along the cosmic web.

3.5. Conclusion

In this paper, we have utilised a sample of LAEs observed with *JWST* imaging. This imaging has revealed that 100% of our sample have a photometrically identified companion galaxy within 5 pKpc. We exploit existing ground- (XSHOOTER/VLT and MOSFIRE/Keck) and space-based (NIRspec and NIRCam grism) spectroscopy to confirm two-thirds of our companion galaxies. For the remaining one-third of systems we show that it is statistically highly likely that they are physically associated. We therefore obtain a 100% companion fraction for LAEs in our sample and compare this to the $\lesssim 40\%$ fraction found for mass-matched samples at the same redshift. Combining these constraints from high-resolution and high-sensitivity *JWST* data with state-of-the-art radiative transfer on-the-fly magneto-hydrodynamics simulations, we demonstrate that *three ingredients* are key to making Lyα emission detectable from our sample of galaxies deep in the epoch of reionisation: *galactic*

mergers driving high intrinsic Lyα emission in the host galaxy; *a 'favourable' line-of-sight* cleared of local neutral hydrogen in the host galaxy by tidal interactions with companions and by star-formation feedback; a sufficiently large ionised bubble facilitating the escape of Lyα emission through the IGM. This paper therefore introduces a new interpretation to explain the unexpected detection of Lyα in $z \geq 7$ galaxies in an epoch where the IGM is mostly neutral and gas-rich, early galaxies are often heavily 'locally' Lyα damped.

CHAPTER 4

THE ENVIRONMENT OF AN AGN AT $z = 10.603$: NEW INSIGHTS ON THE MOST DISTANT LYα DETECTION

4.1. INTRODUCTION

One of the most intriguing challenges of modern extragalactic astronomy is to understand how the first luminous objects emerged from a dark and neutral Universe less than 300 million years after the Big Bang. This quest for *Cosmic Dawn* has been a scientific objective of many telescopes and instruments since the 1950s (for a review see Ellis 2022). Deep and large photometric near infra-red surveys have been undertaken to identify galaxies during the Epoch of Reionisation (EoR), at $z \geq 6$, using the Lyman Break technique (Steidel et al., 1996), a well-known and tested method based on the fact that UV photons emitted at energies higher than Lyα ($\lambda \leq 1216$Å) are almost completely absorbed by the neutral gas surrounding these galaxies. However, this technique can lead to the selection of interlopers, such as dusty galaxies (e.g., Hayes et al. 2012, Arrabal Haro et al. 2023), brown dwarfs (e.g., Wilkins et al. 2014) or supernovae (e.g., Yan et al. 2023). Therefore, spectroscopic follow-up is crucial to confirm the nature of high-redshift galaxy candidates.

Because of the abundance of ionising photons from primeval galaxies and AGN in the early Universe, Lyα is expected to be very strong at high-redshift (EW$\geq$80Å, e.g., Partridge & Peebles 1967, Charlot & Fall 1993) and has been used to spectroscopically confirm some of the most distant galaxies (e.g., Zitrin et al. 2015, Roberts-Borsani et al. 2016, Jung et al. 2022). Several mechanisms are involved in the production of Lyα photons. In the

interstellar medium (ISM), young massive stars ionise the neutral hydrogen leading to the emission of Lyα photons by the recombination of hydrogen atoms. This effect is enhanced by the presence of active galactic nuclei (AGN). Further away, in the circum-galactic medium (CGM), Lyα can potentially be excited in shock-heated gas associated with accreting flows. Similarly, outflows can result in Lyα emission in shock-heated gas. Resonant scattering is also an important mechanism for producing Lyα photons since neutral hydrogen is optically thick to these photons (for a review see Ouchi et al. 2020). In addition to the mechanisms discussed above, the formation of Lyα halos could be reinforced by the presence of nearby companions (Bacon et al., 2021; Leonova et al., 2022). Constraining the spatial extent of this strong emission line provides constraints on the ISM and CGM properties, on the nature of the source (star-forming galaxy or AGN) and on the environment of galaxies, especially those found in the early Universe.

The large fraction of neutral hydrogen surrounding galaxies in the epoch of reionisation reduces the detectability of Lyα emission originating at the earliest epochs (Robertson et al., 2015; De Barros et al., 2017). However, within the last decade, several surprising Lyα detections have been obtained from galaxies deep in the epoch of reionisation (Zitrin et al., 2015; Hashimoto et al., 2018; Bunker et al., 2023b). A recent comparison between *JWST* observations and simulations proposes an explanation for the detection of Lyα at $z \geq 7$ through processes driven by interacting galaxies (Witten et al., 2024b) located in overdense regions (eg. Leonova et al., 2022; Jung et al., 2022; Tang et al., 2023; Witstok et al., 2024c). The most distant Lyα emission line, detected with NIRSpec/*JWST* data in a *HST*-selected galaxy (Bouwens et al., 2010), is from GN-z11, at $z = 10.6$ (Bunker et al., 2023b) with an integrated flux of 2.3×10^{-18} erg/s/cm^2 and an equivalent width (EW) of 18.0±2.0 Å. Interestingly, the presence of extremely dense gas ($> 10^{10}$ cm^{-3}, typical of the Broad Line Region of AGN), the detection of [NeIV]$\lambda\lambda$2422,2424 and of CII*1335 emission, indicate that GN-z11 also hosts an AGN (Maiolino et al., 2024a). Spectroscopic follow-up with IFU data has recently provided indications of the presence of PopIII stars in the vicinity of GN-z11 (Maiolino et al., 2024b), making this galaxy one of the most interesting and intriguing objects in the early Universe.

In this work, we investigate the spatially resolved Lyα emission in the vicinity of GN-z11, the most distant Lyα emitter currently known, by exploiting the recent NIRSpec IFU observations presented in Maiolino et al. (2024b). In section 4.2 we briefly describe the observations and the data reduction steps taken to obtain our final reduced datacubes. In section 4.3 we characterise the properties of the Lyα halo and in section 4.4 we search for companion galaxies in the local vicinity of GN-z11. Throughout this work, we use the AB

magnitude system (Oke & Gunn, 1983) and assume the Planck Collaboration et al. (2020) flat ΛCDM cosmology.

4.2. OBSERVATIONS AND DATA REDUCTION

In this work, we use data obtained by the NIRSpec IFU (Jakobsen et al., 2022; Böker et al., 2022) which is presented in full detail in Maiolino et al. (2024b). This data was obtained on 22nd – 23rd of May 2023 under the DDT program 4426 (PI: Roberto Maiolino). The observations used a medium cycling pattern with 10 dithers using a total of 3.3 h on source with the medium-resolution grating/filter pair G140M/F100LP and 10.6 h with the medium-resolution grating/filter pair G235M/F170LP. This configuration covers 0.97–1.89 μm and 1.66–3.17 μm respectively, with a nominal spectral resolution of $R \sim 1000$. The detector was set with the improved reference sampling and subtraction pattern (IRS^2), which reduces significantly the readout noise (Rauscher et al., 2012).

Despite *JWST*'s excellent pointing accuracy (i.e. 1 $\sigma = 0.1''$, [1] Rigby et al. 2023) and accurate target coordinates using GAIA-aligned astrometry, the observations were not centred properly as the guide star selected by the observatory was a binary system, introducing a large offset ($1.4''$ towards SE; Maiolino et al. 2024b). This offset caused GN-z11 to be close to the edge of the FoV of the NIRSpec/IFU. As a result, for three dither positions, GN-z11 was located at the edge of FoV, while for the remaining seven dithers varying parts of the region west of GN-z11 were outside the FoV of the instrument. This results in steeply decreasing sensitivity of the observations to the West of GN-z11 (and also beyond $0.7''$ North of GN-z11). We re-aligned the astrometry of the NIRSpec/IFU observations to the GAIA-aligned NIRCam. We collapsed the IFU cube in the wavelength range of the NIRCam filter and realigned it with the UV continuum of GN-z11 in the NIRCam observations (Tacchella et al., 2023b).

The full data reduction is described in Maiolino et al. (2024b), here we briefly describe the process. The raw data were processed with the *JWST* Science Calibration pipeline[2] version 1.8.2 under CRDS context jwst_1068.pmap. We made a number of modifications to the reduction steps to improve the data reduction (see Perna et al. 2023 for more details). The final cubes were created using 'drizzle' method setting the spatial pixel scale to $0.06''$ to match the pixel scale of the NIRCam images. For this step, we used an official patch to correct a known bug[3].

[1] https://jwst-docs.stsci.edu/jwst-observatory-characteristics/jwst-pointing-performance

[2] https://jwst-pipeline.readthedocs.io/en/stable/jwst/introduction.html

[3] https://github.com/spacetelescope/jwst/pull/7306

4.3. PROPERTIES OF THE MOST DISTANT LYα HALO

In order to map the individual emission lines we first subtracted the background spectrum. We created the master background by producing a median spectrum of object-free spaxels with a running median filter of 25 channels to avoid any noise spikes. This master background was subtracted from each spaxel in the cube. To create the Lyα map, we collapsed the channels within $\pm$400 kms^{-1} centred on the peak of the Lyα emission line from Bunker et al. 2023b (channels 695-703, $\sim 1.4115 - 1.4152$ μm) and subtracted the continuum image (created by taking the channels immediately red-ward of the Lyα emission line and taking the median flux). We smoothed the Lyα image with a Gaussian kernel with a sigma of 1.5 pixels (0.09$''$). We show this Lyα image in the left column of Figure 4.1 as red contours (2, 3, 4 and 5 σ levels; with the noise level estimated using the object free central region of the cube) plotted over the NIRCam F277W image. We indicate the region of the data cube with no more than 70% of dither positions (because of the observatory guiding problem) by a green dotted line and we also highlight the region with <30% of the dither positions by green shaded region.

The Lyα is extended in the NE direction, with a secondary spatial peak 0.4$''$ (i.e. 1.6 kpc) away from the location of the GN-z11 continuum peak, see Figure 4.1. We also detect an extension of the Lyα to the SW of the continuum centre of GN-z11 on a scale of 0.2 arcseconds, similar to the previous reports of this extension in the MSA observations in Bunker et al. (2023b). The finding of extended CIII]λ1909 emission in this direction, with a funnel-shaped geometry (Maiolino et al., 2024a), indicates that the SW Lyα extension is tracing the AGN ionisation cone. However, the Lyα extension in the latter region should be considered with great care because it is only partially covered by the dithers of the observations and is also potentially subject to field edge issues.

We extracted the spectrum from two separate regions: one encompassing the NE extension (i.e. companion) and the other one centered on the GN-z11 continuum location and also including the SW extension (these regions are shown by blue dashed lines on the NIRCam images in Figure 4.1). Specifically, we selected these regions to coincide with the structure and peaks of the Lyα map in order to extract the total emission line flux. In the combined 1D spectra, we detect Lyα emission at 4.5σ and 6.0σ for the central and NE regions, respectively. There is a data artefact around 1.42 μm, which is only seen in the region where we have 50% fewer exposures (the SW region of GN-z11), possibly associated with field edge issues (furthermore the second peak hinted by this data artifact is not seen in the much deeper MSA observations, Bunker et al., 2023b).

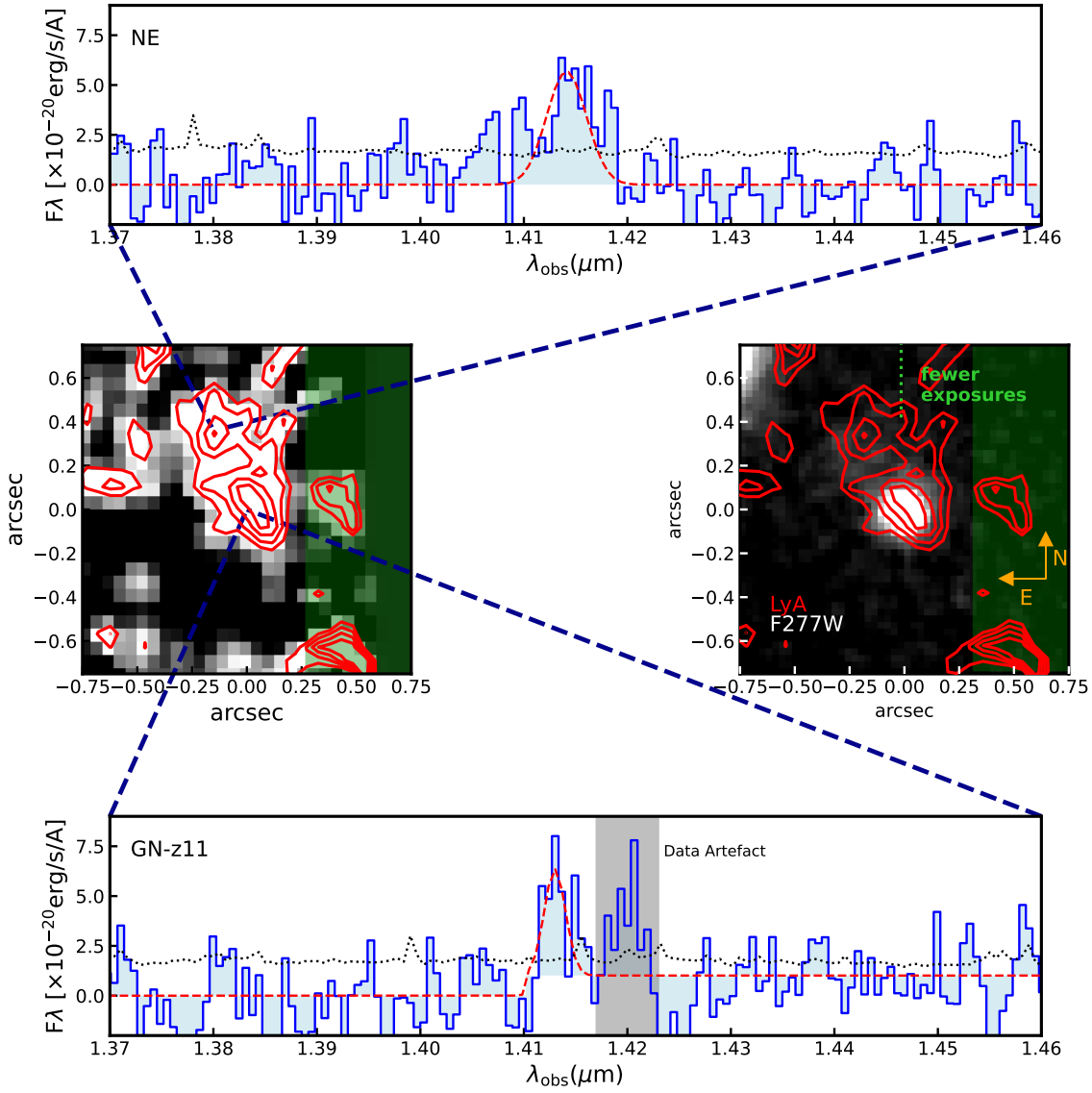


Fig. 4.1 Middle row Left: Map of the Lyα emission with red dashed contours (2,3,4 and 5 σ levels). Right: *JWST*/NIRCAM F277W continuum image with Lyα emission map overlaid as red dashed contours (2,3,4 and 5 σ levels). The green dotted lines indicate where less than 70% dither positions cover the area and the green shaded area is covered by 30% of the dithers. We show extracted spectra from the two Lyα blobs to confirm the Lyα map morphology. The best fit to the emission line is shown as a red dashed line, while the noise is shown as a black dotted line. We highlight the spectral region affected by the data artefact at the edge of the FoV as a grey-shaded region.

We fitted the extracted spectrum with both a Gaussian profile describing the emission line and a power law with a break at 1.41 microns to describe the galaxy continuum. Although Lyα emission often has a complex line profile, given the low signal-to-noise ratio (SNR) of these observations we fit a simple Gaussian profile. This model was fitted to the data using the Markov chain Monte Carlo (MCMC) algorithm with Python's `emcee` package (Foreman-Mackey et al., 2013). The final quoted parameter values and their uncertainties are the median value and 68% confidence intervals of the posterior distribution.

The combined measured flux from the two separate regions (3.9×10^{-18} erg s^{-1} cm^{-2}) is 70 % higher than those measured in the MSA observations (Bunker et al. 2023b). Since flux calibrations of the MSA and IFU observations are accurate to < 5% for both instrument modes (Böker et al. 2023), two factors are likely affecting the MSA measurements: 1) potential self-subtraction of the extended emission line in the MSA observations and 2) the Lyα emission being more extended than the size of the MSA shutter. We summarise Lyα emission line results for each region in Table 4.1.

The Lyα emission line profile varies significantly across the two separate regions. The FWHM of the Lyα emission line varies between 440 and 1000 km s^{-1} across the regions. Furthermore, the velocity offset of the Lyα emission with respect to the rest-frame optical emission lines from Bunker et al. (2023b) (z = 10.6034) is 520 $^{+66}_{-55}$ and 915$^{+86}_{-85}$ km s^{-1} for the central and NE blobs, respectively. We report the results of fitting the Lyα emission from these two regions in Table 4.1. We recover a velocity offset of the Lyα emission line in the local vicinity of GN-z11 that is consistent with that measured by Bunker et al. (2023b) and observe an increasing velocity offset in the NE direction, as shown in Figure 4.1. We see a tentative detection of the unresolved CIV$\lambda\lambda$1548,1551 emission line doublet in the location of the NE blob with a 420 km s^{-1} velocity offset compared to the systemic redshift of GN-z11 (see Appendix for more information). This implies that both GN-z11 and the NE blob have a similar velocity offset of their Lyα emission from their respective systemic redshifts of $\sim$ 500 km s^{-1}. This large velocity offset is indicative of a very low escape fraction of Lyman-continuum photons ($\sim$ 0%) (Izotov et al., 2018b; Gazagnes et al., 2020; Kakiichi & Gronke, 2021), which is consistent with the escape fraction derived by Bunker et al. (2023b) (3.8%).

We also note that the Lyα emission lines of both regions do not show significant asymmetry. While this asymmetry is somewhat typical of high-redshift Lyα emitters (Finkelstein et al., 2013; Oesch et al., 2015; Jung et al., 2022), Kakiichi & Gronke (2021) show that with no or few holes in the neutral hydrogen surrounding LyC leaking regions, symmetrical Lyα profiles can be observed. Moreover, Witten et al. (2023) show that the IGM transmission curve can significantly disturb the asymmetry of the Lyα line as it processes through

Table 4.1 Measured SNRs, fluxes, FWHM and velocity offsets to the emission lines in GN-z11 for the Lyα emission line extracted from the MSA (Bunker et al., 2023b) and the regions defined in Figure 4.1 and 4.3. Notes:* Values taken from Bunker et al. (2023b) for the full shutter extraction spectrum.

Region	SNR	Flux ($\times 10^{-19}$) erg s^{-1} cm^{-2}	FWHM km s^{-1}	δv km s^{-1}
MSA*	12.1	23.0±1.9	530 ±65	555^{+32}_{-32}
Central	4.5	$12.5^{+3.3}_{-2.8}$	440^{+204}_{-95}	520^{+66}_{-55}
NE	6.0	$26.8^{+4.4}_{-4.5}$	1000^{+145}_{-142}	915^{+86}_{-85}
E	4.5	$3.34^{+0.81}_{-0.75}$	520^{+120}_{-90}	810^{+66}_{-55}

the neutral IGM and hence, the symmetrical Lyα emission line observed here does not necessarily trace the intrinsic profile of the emission line. Furthermore, Bunker et al. (2023b) do detect some asymmetry in the Lyα line originating from the position of the UV continuum of GN-z11 and therefore it is possible that the IFU observations do not have the required depth to detect this asymmetry.

The morphology of Lyα emission around GN-z11 is complex with two extended connected peaks. As such it is difficult to assess whether this is a single Lyα halo or two overlapping halos around two separate sources. As a result, we define the Lyα halo size in two different ways: 1) we assume that these are two independent Lyα halos and here we report the size of the halo centred on UV continuum of GN-z11 (similar to the definition from Bunker et al., 2023b) which would result in a radius of 0.8 kpc; 2) we assume that the entire structure is single asymmetric Lyα halo encompassing the two peaks and we define the size from the UV continuum centre to the furthermost 2σ contours of 0.78±0.1″ (3.2±0.4 kpc).

Before the arrival of the *JWST*, one of the best instruments to study the profile of Lyα halos at high redshift was the Multi Unit Spectroscopic Explorer (MUSE) at the Very Large Telescope (VLT) (Bacon et al., 2010). However, its wavelength coverage is not sufficient to detect Lyα halos at $z \geq 6.7$. Leclercq et al. (2017) studied the halos of 145 galaxies at $3 < z < 6$ and concluded that 80% of their sources have Lyα emission more extended than the UV continuum. Farina et al. (2019) presented VLT/MUSE observations of 31 quasars (z=5.7–6.6) with 39% of the sources revealing Lyα halos. The UV luminosity of GN-z11 (M_{UV} = -21.5) would be considered a bright star-forming galaxy in Leclercq et al. (2017) and Wisotzki et al. (2018) and a faint quasar in Farina et al. (2019), hence, it is necessary to compare the Lyα halo properties to both star-forming galaxies and quasars.

We compare the surface brightness of GN-z11 to the profiles of star-forming galaxies from Wisotzki et al. (2018) and quasars from Farina et al. (2019) in Figure 4.2. To account

for cosmological dimming across different redshifts, we multiply the surface brightness profiles by $(1+z)^4$ factor to make an easier comparison. We show two separate measurements for GN-z11 based on the definitions outlined above; 1) a single smaller halo centre on GN-z11 UV continuum (dark blue circle); 2) the extent of the full larger halo encompassing both peaks (light blue square). Remarkably, the size and surface brightness of the GN-z11 halo is consistent with those around quasars rather than star-forming galaxies. We also note that the asymmetric Lyα halo currently observed around GN-z11, although potentially as a result of insufficient dithers due to the guiding problem, does mimic the structure of Lyα halos around quasars in overdense regions at lower redshifts (e.g. Cantalupo et al., 2014; Hennawi et al., 2015; Bacon et al., 2021).

We also compare our measurements with what is expected from zoom-in simulations of a large halo of $M_{\rm h} = 6.9\times10^{12}$ M$_\odot$ at $z = 6$ (Bennett et al., 2024; Sijacki et al., 2009), with variations in the feedback model from the FABLE suite of simulations (Henden et al., 2018). These simulations are a higher resolution version of those described in Bennett et al. (2024), and use the same physical models. Bunker et al. (2023b) estimate a metallicity for GN-z11 of 0.12 Z$_\odot$, which corresponds to a dust-to-metal ratio (DMR) of ~0.01 (Li et al., 2019). Our Lyα modelling uses the COLT code (Smith et al., 2015), which includes Lyα emission from unresolved HII regions around stars and the effect of ionisation from the AGN on the surrounding gas, and we use two different DMRs of 0.01 and 0.1. Our simulations conclude that (i) the size of the halo depends on the DMR (the halo is 25-33% larger with a DMR=0.01 than with a DMR=0.1) and (ii) the expected size of the GN-z11 halo ranges between 2.5 to 3.0 kpc.

4.4. SEARCH FOR COMPANIONS IN THE VICINITY OF GN-z11

The presence of nine galaxies within an area of ~ 100 Mpc^2 with photometric redshifts within Δz=1 of the redshift of GN-z11, places GN-z11 at the centre of a clustered region, reported by Tacchella et al. (2023b). This could provide an explanation for the escape of Lyα radiation through the intervening neutral IGM, via the production of a large ionised bubble (Tilvi et al., 2020; Leonova et al., 2022; Witstok et al., 2024c). Moreover, the presence of a HeII clump, reported by Maiolino et al. (2024b), that may host PopIII stars, provides the first evidence of a potential companion galaxy just 0.6″ from GN-z11. The presence of such companions helps to explain the production and escape of Lyα radiation through the local environment of high-redshift galaxies (Witten et al., 2024b).

With this in mind, we search for emission lines in the NIRSpec IFU FoV that can be attributed to companion galaxies at a redshift consistent with that of GN-z11.

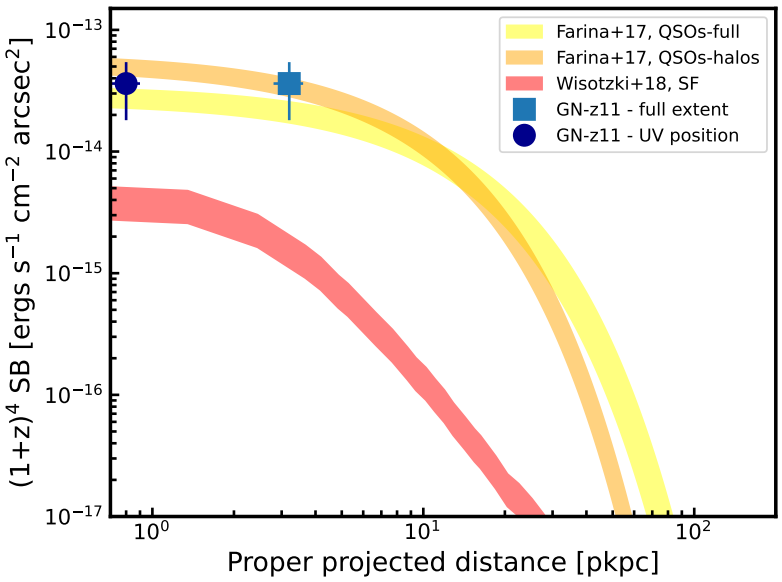


Fig. 4.2 Comparison of the Lyα halo surface brightness of GN-z11 to the profiles of Lyα halos around star-forming galaxies at z=5.0-6.0 (red shaded region Wisotzki et al., 2018) and quasars z=5.7-6.6 (yellow and orange shaded region Farina et al., 2019). For the quasars from Farina et al. (2019) we show the radial surface brightness profiles for the full sample and for only those with detected Lyα halos as yellow and orange regions, respectively. We show two separate definitions of halo size for GN-z11: a smaller halo in the vicinity of UV continuum (dark blue circle) and the full extent of the asymmetric halo (light blue square). The surface brightness and the extent of GN-z11's Lyα halo are in agreement with quasars rather than star-forming galaxies at z~6 .

We note that no photometric candidate companions exist within the limited FoV of these observations. Any such companions identified with these observations will therefore have no associated UV continuum. However, Kerutt et al. (2022) found that one-third of LAEs in MUSE have no associated UV continuum counterpart and hence this would not be unexpected.

4.4.1 Lyα emitting blobs

Figure 4.1 reveals a Lyα "blob" $\sim 0.4''$ to the north-east of GN-z11, originating just above the "haze" reported by Tacchella et al. (2023b). This 6 σ detection of Lyα spectroscopically confirms the presence of a companion galaxy. Indeed, we also have a tentative CIVλ1550 detection in the NE Lyα blob at 3.2σ (see Appendix 4.6.1) indicating a potentially highly ionising source in that region, such as a very young stellar population or a faint AGN (see Maiolino et al., 2024b, for more details).

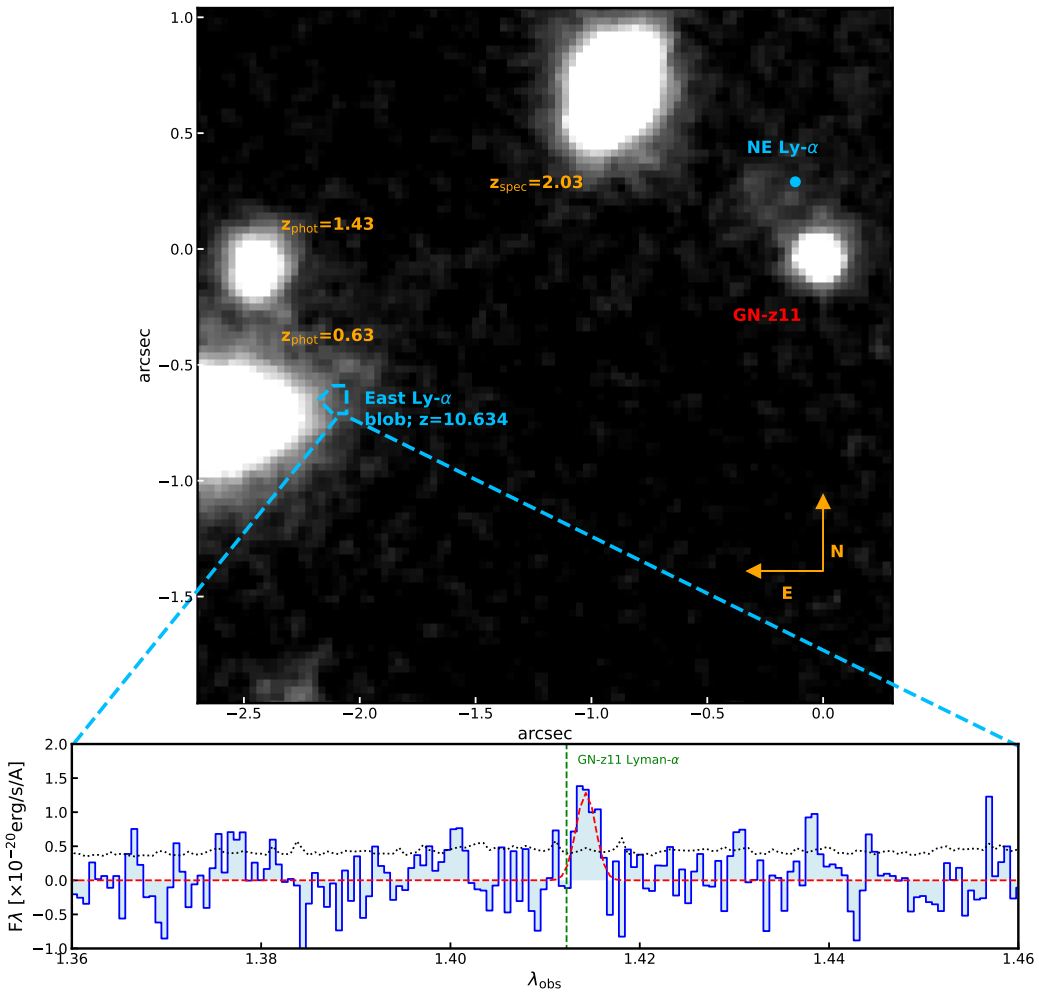


Fig. 4.3 Detection of an additional Lyα emitter, 2.1 arcseconds to the East of GN-z11. Top: *JWST*/NIRCam F277W image corresponding to the FoV of our NIRSpec/IFU observations. We indicate the region from which we extracted the spectrum on the right with blue dashed contours. We also note three foreground galaxies visible in the NIRCam image with their corresponding redshifts. Bottom: Extracted Lyα spectrum from the region indicated in the NIRCam image. The red dashed line shows the best fit to the spectrum while the black dotted line shows the error spectrum. The green dashed line shows the wavelength of Lyα emission seen in the main component of GN-z11. The foreground galaxies in the vicinity show no visible emission lines that would contaminate this spectrum.

Moreover, the detection of HeII emission from a clump ∼ 0.6″ to the north-east of GN-z11, originating from a distinct region to the Lyα halo (Maiolino et al., 2024b) indicates the presence of pristine gas, possibly photoionised by a cluster of PopIII stars.

This detection of emission lines originating from clumps that are distinct from GN-z11 offers a deeper insight into the processes driving Lyα emission at these epochs. It acts as further support for the idea that galaxy interactions are driving Lyα emission in the epoch of reionisation (Witten et al., 2024b), and confirms that a significant amount of neutral hydrogen around these sources must have been displaced in order to facilitate the observation of Lyα photons from both sources, despite their different properties. This result also suggests that peculiar velocities may not be playing a significant role in the escape of Lyα photons given that both GN-z11 and the NE region have relatively large velocity offsets from each other (∼ 400 km/s). If a large velocity offset driven by gravitational interactions between GN-z11 and these clumps was driving the escape of Lyα photons through the neutral IGM, it could be expected that only one source would show Lyα emission. That source would emit Lyα photons that are heavily redshifted relative to the neutral IGM and hence, are more likely to escape without attenuation, while the other source would emit Lyα photons that are blueshifted and hence would be absorbed by the neutral hydrogen in the IGM unless a significant ionised bubble is present (eg. Tilvi et al., 2020; Leonova et al., 2022; Witstok et al., 2024c; Witten et al., 2024b). While the interaction between GN-z11 and its companions is more complex than a simple two-body interaction, this result offers an early indication that peculiar velocities may not be the most important factor in facilitating Lyα escape.

We also searched for more distant Lyα-emitting sources in the full FoV of the NIRSpec IFU observations (specifically the area fully covered by the 10 dithers) within ±2000 km s^{-1} of the systematic redshift of GN-z11. We find a single candidate detected 2.15 arcseconds East of GN-z11 and we show the location of the candidate and its spectrum in Figure 4.3. The emission line is detected at 4.5σ with FWHM of 520^{+120}_{-90} km s^{-1} at z=10.6346, 290 km s^{-1} offset redwards of the Lyα emission line of GN-z11.

Since the Lyα emitter is close to the foreground galaxy, we verified that the emission line is indeed from a z ∼10.63 galaxy and not the foreground source ($z_{\rm phot}$=0.63; Hainline et al., 2024). We extracted a spectrum with a radius of 0.3″ centred on the source East of the Lyα blob in both band 1 and band 2 IFU observations. Although the continuum is well detected in both bands, we do not detect any emission lines in this foreground source. Assuming that the Lyα emission from the candidate is in fact an emission line coming from the foreground source at $z \sim 0.63$, the only possible emission line would be

CaIIλ8664, which we do not expect to see without detection of the [SIII]λ9532, HeIλ10830 and Paδ emission lines.

Given that these emission lines are undetected, we conclude that this emission is not coming from the foreground galaxy, but is indeed an additional Lyα emitter at the redshift similar to GN-z11.

4.4.2 A candidate $z = 10.6$ protocluster core

Tacchella et al. (2023b) previously identified a photometric candidate overdensity surrounding GN-z11 within a ~ 10 cMpc × 10 cMpc region, which corresponds to ~430 pkpc at z=10.6. The detection of the HeII blob (Maiolino et al., 2024b), the Lyα blob ~ 0.4″ NE of GN-z11 and the detection of a further Lyα blob 2.1″ to the East brings the number of spectroscopically confirmed $z \sim 10.6$ objects in the NIRSpec IFU FoV to four. Previous theoretical works (eg. Chiang et al., 2017) have demonstrated that the expected size of a protocluster core is already on scales of just hundreds of ckpc by $z \sim 7$. Observations of protocluster cores at a similar redshift (eg. Capak et al., 2011; Arribas et al., 2024; Hashimoto et al., 2023; Morishita et al., 2023) have also found a central core of order 150 ckpc. This finding motivates our conclusion that the objects seen within the IFU FoV (140 ckpc × 140 ckpc) are representative of the core of a protocluster. Moreover, theoretical work by Chiang et al. (2017) has shown that the overall size of the protocluster at high redshifts like these is typically of order 10 cMpc, consistent with the larger scale overdensity of galaxies identified by Tacchella et al. (2023b).

Comparing the number of objects identified in the IFU FoV at $z = 10.6$ we find an overdensity parameter of $\delta \geq 27$. We estimate the number of galaxies that are expected in the FoV by integrating the luminosity function of Harikane et al. (2024), at $z \sim 10$, down to a UV magnitude that is determined by assuming that the Lyα blobs have a maximum Lyα rest-frame equivalent width of ~ 1000Å. We make this assumption as these Lyα blobs have no associated UV continuum. We give the overdensity parameter as a lower bound as the FoV of the NIRSpec IFU is too small to accurately constrain the overdensity parameter, and thus we use the upper-bound expected number of galaxies to give us an absolute lower bound on our overdensity parameter. This value is significantly larger than the overdensity parameter for the $z = 7.66$ protocluster core presented in Laporte et al. (2022) and, following the same method as above, we estimate an overdensity parameter of $\delta \geq 11$ for the $z = 7.88$ protocluster core presented in Hashimoto et al. (2023). Given that the region surrounding GN-z11 is more overdense than previously observed protocluster cores at $z \sim 8$, that it resides in a large-scale photometric redshift overdensity (Tacchella et al., 2023b) and that protoclusters are expected to fuel AGN in their central galaxy (eventually

forming the observed population of massive $z \sim 6$ quasars, see Bennett et al., 2024), as is observed in the case of GN-z11 (Maiolino et al., 2024a), the region surrounding GN-z11 appears to be a strong candidate for a protocluster core. Moreover, GNz11 is by far the brightest galaxy in this structure, potentially making it a precursor of the brightest cluster galaxies seen at lower redshifts.

As a preliminary analysis, we estimate the dark matter halo mass of this structure following Behroozi et al. (2013). We assume that the total stellar mass of this protocluster core is dominated by GN-z11 and obtain a dark matter halo mass of $M_{\rm h}$=$2.96^{+0.44}_{-0.39}\times10^{10}$ $M_\odot$. This is a reasonable assumption given that all objects within the IFU field-of-view do not have any UV counterpart in the NIRCam images. Moreover, candidates identified in Tacchella et al. (2023b) are 3 magnitudes fainter than GN-z11 and will therefore have much smaller stellar masses than GN-z11. We compare this dark matter halo mass with previous findings at lower redshift (Polletta et al. 2021, Casey 2016, Champagne et al. 2021, Wang et al. 2016, McConachie et al. 2022, Toshikawa et al. 2018, Long et al. 2020, Calvi et al. 2021, Chanchaiworawit et al. 2019, Harikane et al. 2019, Laporte et al. 2022) and with the expected evolution of a Coma-like cluster at $z = 0$ (Chiang et al., 2013) assuming a constant slope in the evolution at $z \geq 4$ (see Figure 4.4). The estimated dark matter halo mass of the GN-z11 protocluster is fully consistent with what is expected for a Coma-like cluster progenitor at z =10.6.

4.5. Conclusions

We have presented an analysis of *JWST*/NIRSpec IFU observations of GN-z11, an extremely luminous galaxy hosting an AGN at z =10.603. Due to a telescope guiding problem, GN-z11 is near the edge of the FoV and hence we are only able to map the Lyα halo extension on and to the North and East of GN-z11.

We detect extended Lyα emission at the location of GN-z11 and towards the NE. Extracting the spectrum from the two separate regions (GN-z11 and the NE extension) we have measured a velocity offset of the Lyα emission with respect to GN-z11 of $520\ ^{+66}_{-55}$ and 915^{+86}_{-85} km s^{-1} for the central and NE extension, respectively. We have measured the extent of the Lyα halo to be 0.8 – 3.2kpc; based on two separate definitions of the size (i.e. edge to edge of the whole structure or the size of the smaller-scale halo around the UV continuum of GN-z11). Comparing the radial surface brightness of GN-z11 to previous studies at z=5-6.5 (Wisotzki et al., 2018; Farina et al., 2019) shows that the Lyα halo is consistent with those of quasars rather than star-forming galaxies.

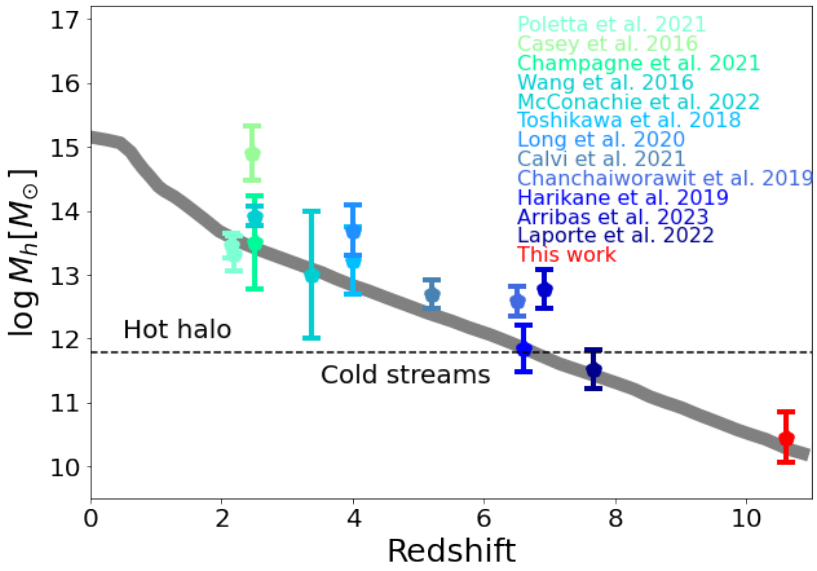


Fig. 4.4 Evolution of the protocluster dark matter halo mass as a function of redshift. The grey line shows the evolution of a protocluster leading to a Coma-like cluster at $z = 0$ (Chiang et al., 2013). We extrapolate this evolution to $z \sim 11$ assuming no change in the slope from $z \sim 4$. The dark matter halo mass of GN-z11 is fully consistent with this evolution.

We also detect a distinct Lyα blob to the NE of GN-z11. Together with the previously identified HeII clump (Maiolino et al., 2024b), this indicates the presence of multiple companions surrounding GN-z11. This appears to support the theory that companion galaxies are playing a key role in the detectability of Lyα emitting galaxies deep into the epoch of reionisation (Witten et al., 2024b).

The further identification of a Lyα blob 2.1″ East of GN-z11 brings the number of objects within the NIRSpec IFU FoV (3″x3″) to four, making this region more overdense than previously discovered protocluster cores at $z \sim 8$. This significant overdensity, combined with the larger scale overdensity reported in Tacchella et al. (2023b) and the presence of an AGN in GN-z11 (Maiolino et al., 2024a), at the centre of this system, makes the region studied in this work a strong candidate for a protocluster core. Our preliminary analysis of this protocluster demonstrates that it has a dark matter halo mass consistent with what is expected for a Coma-like cluster progenitor at $z \sim 10.6$. Our study confirms that the most massive galaxies identified with *HST*, like GN-z11, are indeed surrounded by much fainter, previously undetected, galaxies at similar redshifts which likely form the first large-scale structure in the Universe.

4.6. Appendix

4.6.1 Tentative detection of CIV in NE blob

As we discussed in the §4.4 we search for any other emission line at the same locations as the Lyα blobs detected with NIRSpec/IFU observations. We see a tentative detection of CIVλ1551 emission within the same aperture as the bottom panel of Figure 4.1. We show the extracted spectrum and the best fit in Figure 4.5. The line is detected at 3.2σ in the Band 2 observations at z= 10.620 (i.e. +420 km s^{-1} of GN-z11) with a flux of $7.54^{+3.12}_{-2.35} \times 10^{-19}$ ergs s^{-1} cm^{-2} and a FWHM of 730^{+290}_{-240} km s^{-1}.

The detection of CIV requires very hard ionisation field mostly likely coming from young high-mass stars. However, this tentative detection needs to be confirmed by the future re-observations of this programme.

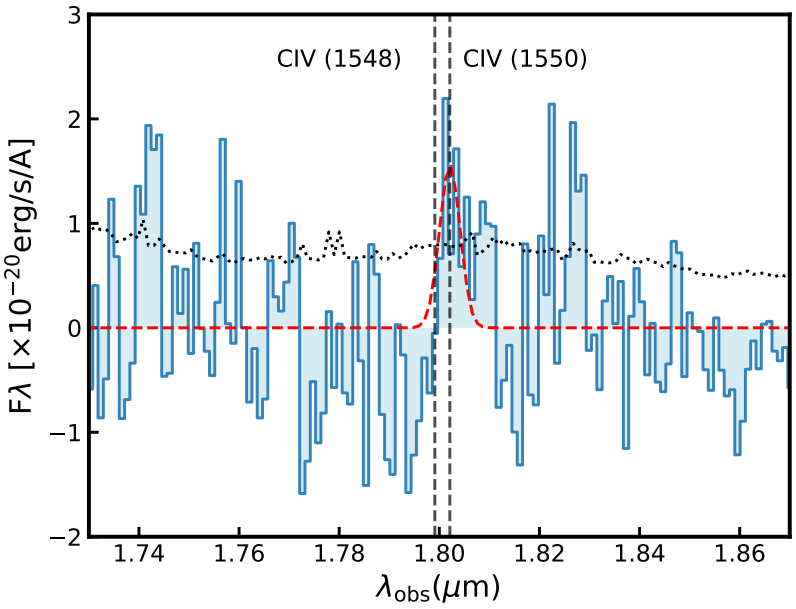


Fig. 4.5 Tentative detection of CIV$\lambda\lambda$1548,1550 in the location of NW Lyα blob. The red dashed line indicates the best fit to the CIVλ1551 line and the black dotted lines show the noise spectrum.

CHAPTER 5

RISING FROM THE ASHES: EVIDENCE OF OLD STELLAR POPULATIONS AND REJUVENATION EVENTS IN THE VERY EARLY UNIVERSE

5.1. INTRODUCTION

The advent of the *James Webb Space Telescope (JWST)* has facilitated the robust detection of extremely high-redshift galaxies out to $z > 10$ (e.g. Curtis-Lake et al., 2023; D'Eugenio et al., 2024; Finkelstein et al., 2024; Arrabal Haro et al., 2023; Carniani et al., 2024; Castellano et al., 2024). The detection of these galaxies pushes back the onset of Cosmic Dawn to at least $z > 14$. The identification of galaxies at such redshifts was previously predicted by various studies (Hashimoto et al., 2018; Roberts-Borsani et al., 2020; Laporte et al., 2021; Tacchella et al., 2022) due to observed excesses in their optical-continuum detected by *Spitzer* relative to the UV-continuum detected with the *Hubble Space Telescope (HST)*. This excess was purported to be evidence of an older stellar population and hence an earlier formation redshift that would push Cosmic Dawn to at least $z \gtrsim 14$.

This excess, is primarily associated with the Balmer break, occurring at a rest-frame wavelength of ~ 3645Å. It is strongest at specific temperatures and densities, namely it is most prevalent in A-type stars. However, due to their high temperature, the break is nearly non-existent in O and B stars. In order for a break to be present, the stellar population must have evolved off the main sequence such that A-type stars can dominate the optical continuum. Moreover, nebular emission from ionised gas that is produced thanks to the ionising radiation from extremely young stars can produce an "inverse Balmer break" (e.g.

Cameron et al., 2024), which is seen as a decrease in flux between the UV and optical continua, and has been observed to become the dominant Balmer break feature in high-redshift galaxies (Roberts-Borsani et al., 2024). Thus, the strength of the Balmer break is frequently used as a probe of the timescale of stellar mass assembly and has been well-studied in galaxy formation simulations (Katz et al., 2019; Wilkins et al., 2024).

Therefore, with the detection of $z \sim 14$ galaxies by *JWST*, one would expect to observe galaxies presenting large Balmer breaks by $z \sim 8$, where some of these galaxies could already be ~ 300 Myrs old (e.g. Wilkins et al., 2024). Extremely high-redshift, UV-bright galaxies seem to require very bursty star-formation histories (e.g. Carniani et al., 2024), however, the frequency, timescales and rapidity of burst and quenching events is currently poorly understood.

Recent works have leveraged the significant gains that *JWST* has facilitated in imaging the rest-optical wavelengths of high-redshift galaxies to identify robust Balmer break candidates (e.g. Laporte et al., 2023; Endsley et al., 2023; Trussler et al., 2024a,b). Most notably, the use of medium-band filters around the Balmer break wavelength has allowed for the degeneracy between strong nebular emission lines and the Balmer break to be broken (Stiavelli et al., 2023; Bradač et al., 2024). This has additionally resulted in vast improvements in stellar mass estimates, finding that most of the Balmer break galaxies sit between $\log(\mathrm{M}_\star[\mathrm{M}_\odot]) = 8-9.5$ (Endsley et al., 2023; Trussler et al., 2024b,a).

Some examples of galaxies with potential Balmer breaks, emission lines and extremely red photometry (e.g. Labbé et al., 2023) are in a different mass regime ($M_\star > 10^{10} M_\odot$) than galaxies with bursty SFHs. Moreover, such galaxies have recently been spectroscopically observed by NIRSpec and have been found to host broad-line Active Galactic Nuclei (AGN) (Wang et al., 2024b). The contribution of AGN to their spectral energy distribution (SED) makes constraining their SFH challenging. While modelling does suggest the presence of an old stellar population, dynamical mass constraints require that the optical light is dominated by a strong AGN contribution, making the fraction and age of old stars hard to constrain accurately (Wang et al., 2024b).

While *spectroscopic* examples of Balmer breaks in galaxies have been seen above $z = 7$, these exclusively show no evidence of strong nebular emission lines, indicative of very little recent star formation (e.g. Looser et al., 2024; Vikaeus et al., 2024; Trussler et al., 2024a). These $M_\star \sim 10^{8.5} M_\odot$ (mini-)quenched galaxies are superb cases for hosting a previous burst in their SFHs, however, they only place constraints on the time since the most recent burst and hence only provide a lower limit on the timescales between bursts.

While many simulations predicted the bursty SFH of early galaxies (Fukuda et al., 2000; Immeli et al., 2004; Bournaud et al., 2007; Elmegreen et al., 2009; Faucher-Giguère, 2018;

Tacchella et al., 2020), observations currently fail to constrain the properties of these bursts, i.e. their duration, frequency and quenching mechanisms. Moreover, direct *spectroscopic* evidence of galaxies rejuvenating on short timescales and hence of mini-quenched periods has yet to be seen. Observational constraints on the burstiness of high-redshift galaxies largely comes from suspected (mini-)quenched galaxies (Looser et al., 2024, 2023; Dressler et al., 2023, 2024; Strait et al., 2023; Trussler et al., 2024a) and galaxies suspected to be in the process of rapidly quenching (McClymont et al., 2024b), however these are not yet strong enough to constrain models of bursty SFHs. These works suggest that mini-quenched periods lasts at least a few tens of Myrs. This timescale between bursts is of significant interest given that it encodes information about the mechanisms that drive quenching and star-formation events (e.g. Tacchella et al., 2020; Dome et al., 2024; Kravtsov & Belokurov, 2024; Sun et al., 2023b; Mason et al., 2023), which are likely a combination of supernovae and AGN feedback, stellar winds and the effects of mergers and environment. Identifying galaxies that are in a rejuvenating phase, following a previous burst, not only offers the first confirmation that we are observing mini-quenched galaxies in the early Universe, but also offers the perfect laboratory for constraining the burst timescales of early galaxies.

In this context, we aim to utilise spectroscopic observations to identify and analyse rejuvenating galaxies in the early Universe. One of the prime locations for identifying old stellar populations is in the core of protoclusters in the early Universe. Protoclusters undergo a phase of inside-out growth between $z \sim 10-5$, resulting in a stellar mass distribution that is dominant in the central region of the protocluster (Chiang et al., 2017). Existing cores have been seen to host some of the most massive, dust enriched galaxies at early cosmic times (Laporte et al., 2017a, 2022; Hashimoto et al., 2023; Morishita et al., 2023). Just three spectroscopically confirmed protoclusters currently exist at $z > 7$: A2744-z7p9OD (Zheng et al., 2014; Hashimoto et al., 2023; Morishita et al., 2023), SMACS0723-PC (Laporte et al., 2022, Witten et al. in prep.) and GN-z11-OD (Scholtz et al., 2024; Tacchella et al., 2023b). Only one of these has ALMA dust detections confirming the most massive, dusty galaxy – A2744-z7p9OD and its constituent A2744-YD4 (hereafter YD4) (Laporte et al., 2017a). Therefore, we first introduce the relevant *JWST* NIRSpec data in Section 5.2. We then focus on YD4 in Section 5.3, which has recently been subject to follow up prism observations as part of the Ultra-deep NIRCam and NIRSpec Observations Before the Epoch of Reionization (UNCOVER) program. With this additional exposure time we have sufficient depth to observe the continuum at rest-frame wavelengths greater than 3600 Å and as such, we are able to analyse the Balmer break of YD4. Using the emission line properties of YD4 we identify a population of similar galaxies and stack their spectra. We discuss the results of this stack in Section 5.4 and use a variety of different models to

interpret these results in Section 5.5. Finally, in Section 5.6 we discuss the implications of these results and our conclusions.

5.2. DATA

A2744-YD4 was originally identified in *HST* Frontier Fields imaging (Lotz et al., 2017) of the Abell-2744 cluster by Zheng et al. (2014) as part of a $z \sim 8$ overdensity, A2744-z7p9OD. Follow up *JWST* NIRSpec spectroscopy confirmed YD4 and nine other galaxies in this extreme overdensity (Hashimoto et al., 2023; Morishita et al., 2023; Cameron et al., 2024; Chen et al., 2024) at $z = 7.87 - 7.88$. With a comparable overdensity parameter to other high-redshift protoclusters (e.g. Scholtz et al., 2024; Laporte et al., 2022), A2744-z7p9OD presents as a spectroscopically confirmed protocluster existing in the first 650 millions years of the Universe (Morishita et al., 2023). YD4 resides in the central 11×11 pkpc core of the protocluster. This region has been recently observed with the *JWST* NIRSpec (IFS) leading to the spectroscopic confirmation of 4 galaxies Hashimoto et al. (2023). Moreover, Atacama Large Millimeter/ submillimeter Array (ALMA) Band 6 (ID: 2018.1.01332.S, PI: N. Laporte; Laporte et al., 2019) and 7 (Laporte et al., 2017a) observations identify dust continuum in three of these galaxies (Hashimoto et al., 2023), including in the vicinity of YD4 as previously reported by Laporte et al. (2017a). Given the dust continuum detections in the core of the protocluster and the large stellar masses of the constituent galaxies reported by Hashimoto et al. (2023), the core of this protocluster appears to be evolved. As such the galaxies resident in the core are prime candidates for observing the Balmer break. Previously, Roberts-Borsani et al. (2020) utilised these ALMA detections and *Spitzer* observations to identify the presence of an old stellar population within YD4. However, the presence of emission lines reported by Hashimoto et al. (2023); Morishita et al. (2023) imply these galaxies are not quiescent and the depth of the original NIRSpec Prism spectroscopy of YD4 and YD7 in the core is insufficient to observe continuum emission at the longer wavelengths where the Balmer break may be found.

The spectroscopic and imaging data exploited, in the case of YD4, were observed as part of the UNCOVER program (ID 2561, PI Labbe; Bezanson et al., 2022) and a *JWST* Director's Discretionary Time (DDT) program (ID 2756, PI Chen; Roberts-Borsani et al., 2023a; Morishita et al., 2023). These two programs observed YD4 with the NIRSpec micro-shutter assembly (MSA), using the PRISM/CLEAR configuration. The exposure time of the DDT and UNCOVER observations were 1.2 and 2.3 hours respectively, resulting in a nearly doubling of the signal-to-noise ratio over the DDT data used in Morishita et al. (2023).

The spectra and NIRCam photometry from each program derive from the reduction and catalogs of Roberts-Borsani et al. (2024), and we refer the reader to that paper for full details. As a summary, the spectra were reduced using the `msaexp`[1] code, which makes use of a combination of the official STScI *JWST* pipelines and custom routines for e.g., snowball masking and $1/f$ removal at the uncalibrated (**_uncal.fits**) and count-rate (**_rate.fits**) exposure stages. 2D spectra for the object were extracted from the exposures and background subtraction was done using the adjacent MSA shutters forming part of the larger slitlet. The 1D spectrum and associated uncertainties were extracted and combined from the 2D spectra using an optimal extraction procedure, which makes use of a Gaussian kernel fitted to the spatial profile of the spectrum from the (wavelength-)collapsed 2D image. The final 1D spectra from both programs were then combined via an inverse-variance-weighted mean.

The photometric fluxes instead were extracted from 0".5 diameter, circular aperture photometry on the reduced and publicly-available images of the `Grizli` v6 Image Release, based on a NIRCam F115W+F277W+F444W detection image. Using these, we account for slit loss effects that may be present: first we scale the combined NIRSpec spectrum to the rest-frame UV continuum as seen from the NIRCam photometry, and secondly we multiply the spectrum (blueward of the Lyman-break) by a 2nd order polynomial fit to the residual fluxes between the NIRCam and NIRSpec (pseudo-)photometry. Finally, we correct both the spectrum and imaging for lensing magnification, assuming the magnification factor of $\mu = 1.96$ listed in Table 3 of Roberts-Borsani et al. (2024), which use the magnification maps of Bergamini et al. (2023).

One immediate concern regarding the spectrum of YD4 is that it may be an agglomeration of different sources, especially given its location within an extremely overdense region (see Figure 5.1). Moreover, in such crowded regions continuum subtraction can become challenging and slit-loss corrections can affect the observed emission line ratios. However, recent *JWST* NIRSpec integral field spectrograph (IFS) prism observations do not show a radical evolution in the [O III]λ5007/[O II] or [O III]/Hβ ratios across the spatial extent of YD4 (Venturi et al., 2024), suggesting that the NIRSpec MSA spectrum that we exploit, is unlikely to be an agglomeration of two radically different sources, like has been suggested by Faisst & Morishita (2024), for the suspected (mini-)quenched galaxy reported by Looser et al. (2024). Moreover, these emission line ratios are consistent with those that we report for YD4 in Table 5.1, therefore indicating that continuum subtraction and slit-loss corrections to our MSA spectrum appear not to be playing a role in driving the observed emission line ratios.

[1] https://github.com/gbrammer/msaexp

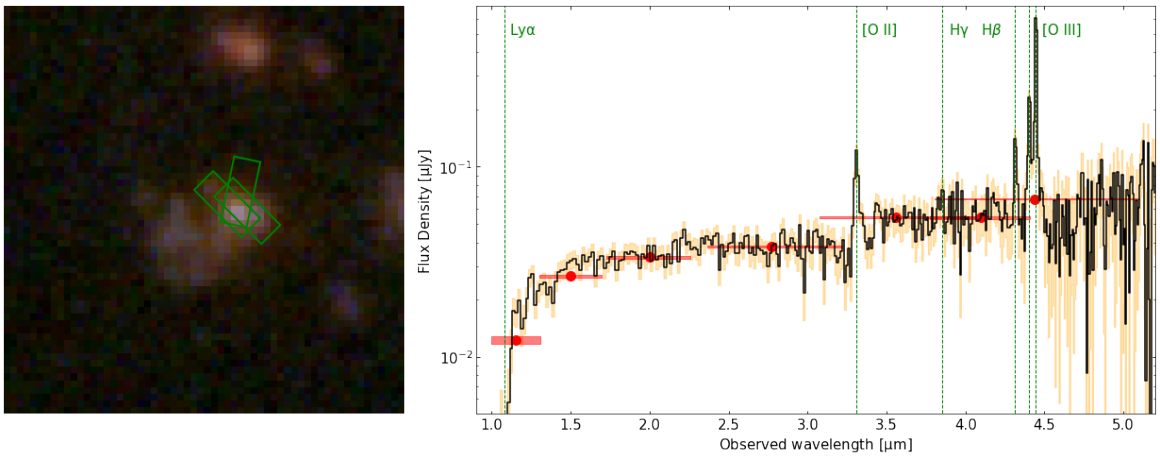


Fig. 5.1 (Left) Positions of the three NIRSpec/*JWST* slit positions used to observe A2744-YD4, the central slit of each observation is indicated in green. For scale, the NIRSpec slit is approximately $0.2'' \times 0.46''$. The NIRSpec observations cover the galaxy YD4 while avoiding significant contamination from YD6 (a seperate $z = 7.88$ galaxy to the lower left of YD4; Venturi et al., 2024) as discussed further in Section 5.2. The background image is a color image obtained by combining F150W (blue), F277W(green) and F444W(red), retrieved from the Dawn *JWST* Archive. (Right) NIRSpec prism spectrum of A2744-YD4 (black) and the 1σ uncertainties (yellow). The red points show the NIRCam/*JWST* photometry and the shaded regions show its uncertainty in the y-axis and the width of the filter in the x-axis. The dashed green lines show the expected positions of notable nebular emission lines. Two key features are seen in the spectrum: a clear Balmer-Break occurring at $\sim 3.3\mu$m and four emission lines (the blended [O II] doublet, Hβ and the [O III] doublet).

The remaining spectroscopic data used in this paper are retrieved in their reduced form from the *JWST* Advanced Deep Survey (JADES) data release (for galaxies in GOODS-S Bunker et al., 2023a) and the Dawn *JWST* Archive (DJA[2]; for galaxies in EGS). The DJA data was reduced using the custom-made pipeline MsaExp v.0.6.12 [3]. Further details off the reduction pipeline are given in Heintz et al. (2024b).

5.3. A2744-YD4

In Figure 5.1 we present the spectrum of YD4. The spectrum shows two notable features: a Balmer break in the spectrum occurring at $\sim 3.3\mu$m and four strong nebular emission line features associated with the [O II]λ3726,3729 Å (hereafter [O II]) doublet, Hβ and both components of the [O III]λ4959,5008 Å (hereafter[O III]) doublet.

[2]https://dawn-cph.github.io/dja/index.html
[3]https://zenodo.org/record/8319596

Table 5.1

Parameter	Measurement *
Fluxes [10^{-19} ergs/s/cm^2]	
[O II]	7.3 ± 0.1
Hβ	2.3 ± 0.2
[O III]$_{4959}$	5.8 ± 1.1
[O III]$_{5007}$	15.1 ± 0.2
EW$_0$ [Å]	
Hβ	25 ± 3
[O III]$_{5007}$	170 ± 18
Line ratios	
$\log_{10}$(O32)	0.12 ± 0.01
$\log_{10}$(R23)	1.18 ± 0.03

Notes:* All measurements have been de-magnified assuming $\mu = 1.96$ from Bergamini et al. (2023). The line ratios are dust corrected using the dust attenuation law obtained from our PROSPECTOR modelling in Section 5.5.2.

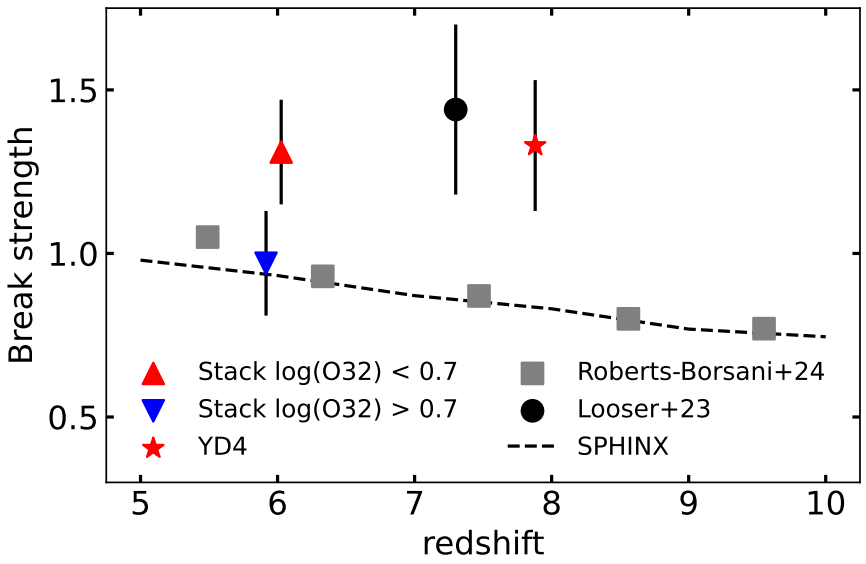


Fig. 5.2 The Balmer break strength, $F_{\nu,4225}/F_{\nu,3565}$, using the definition of Roberts-Borsani et al. (2024), showing the break strength in YD4 (star) and our two stacks (triangles). These are compared to the median break strength found in the SPHINX[20] simulations (dashed line; Rosdahl et al., 2018, 2022; Katz et al., 2023a) and stacked NIRSpec Prism spectra from Roberts-Borsani et al. (2024) (squares). We also show the break strength of the galaxy from Looser et al. (2024) (circle).

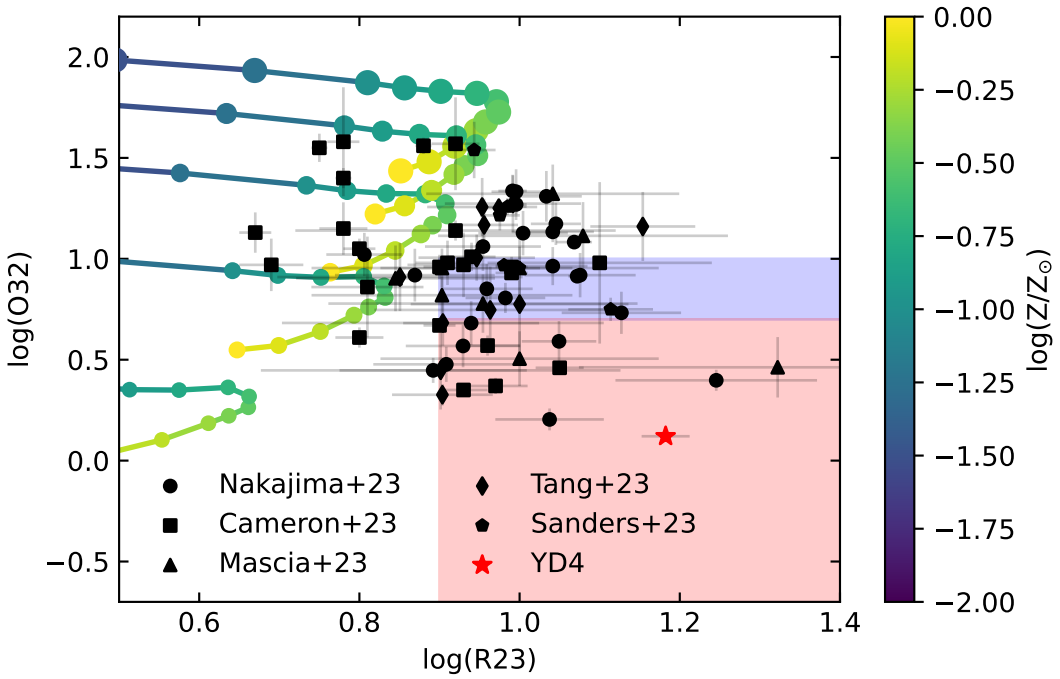


Fig. 5.3 The dust-corrected excitation-ionisation plot, showing NIRSpec data from Nakajima et al. (2023); Cameron et al. (2023b); Mascia et al. (2023); Tang et al. (2023); Sanders et al. (2023). We additionally show the evolution of the O32 and R23 ratios with metallicity (colourbar) and ionisation parameter (increasing marker size, from -3 to -1, in steps of 0.5) using our 5 Myr CLOUDY models (Ferland et al., 2013), discussed further in Section 5.5.1. The values for YD4 are denoted by a red star, while the two stack criteria are indicated by blue and red boxes. Clearly YD4 resides in a poorly populated region of the excitation-ionisation plane.

While observations of Balmer breaks in galaxies at $z > 7$ exist (e.g. Looser et al., 2024), they are certainly rare. Roberts-Borsani et al. (2024) have shown instead that inverse Balmer breaks, so-called Balmer jumps, appear to be typical in galaxies above $z \sim 6$, as shown in Figure 5.2. We choose to exploit the Balmer break definition of Roberts-Borsani et al. (2024), which is more restrictive than typical break definitions (e.g. Binggeli et al., 2019). This definition, $B = F_\nu(4225\text{Å})/F_\nu(3565\text{Å})$, uses windows covering 3500 – 3630 Å and 4160 – 4290 Å thus avoiding any potential strong emission lines, such as the [Ne III]λ3869,3967 Å doublet and He I (λ3889 Å) emission lines. This results in a measured break strength of $B = 1.3 \pm 0.2$ for YD4, where $B > 1$ is a traditional Balmer break, while $B < 1$ indicates significant nebular emission producing an inverse Balmer break. When we compare this break strength to those presented in Roberts-Borsani et al. (2024), in Figure 5.2, it becomes clear that the break observed in YD4 is exceptional relative to galaxies at a similar redshift and is comparable to that of the suspected (mini)-quenched galaxy from Looser et al. (2024) ($B = 1.4 \pm 0.3$).

As we discussed earlier, UV-optical breaks have been observed in AGN-host galaxies. These breaks can be a combination of emission from the accretion disk of the AGN and also a Balmer break (Wang et al., 2024b). As such, given that protocluster cores appear ideal locations to form and fuel massive black holes (Bennett et al., 2024) and that YD4 appears to be evolved thanks to its ALMA dust detections (Laporte et al., 2017a), one could postulate that the break observed in YD4 may be driven by an AGN. However, it is crucial to note that, in galaxies that have optical continua seemingly dominated by AGN accretion disk emission (e.g. the so-called little red dots), the slope of their optical continua is much more red (e.g. Labbé et al., 2023; Furtak et al., 2024; Wang et al., 2024b; Matthee et al., 2024; Kocevski et al., 2024) than the flat optical slope seen in Figure 5.1. Therefore, given YD4 does not exhibit a steep, red optical continuum, evidence of emission lines that require AGN photo-ionisation (e.g. Scholtz et al., 2023) or of broad Balmer emission lines, we conclude that there appears to be no evidence that YD4 is in an AGN phase.

We fit the four emission lines seen in the spectrum of YD4 with Gaussian profiles. This fitting occurs in f_ν space where the optical continuum is flat, we therefore estimate the continuum level of the Gaussian by averaging over the local optical continuum while masking strong nebular emission lines. The resulting flux of these lines is reported in Table 5.1.

Figure 5.3 shows the dust-corrected emission lines ratios from multiple *JWST* programs: JADES (Bunker et al., 2023a), the Cosmic Evolution Early Release Science (CEERS; Arrabal Haro et al., 2023), Early Release Observations (ERO) and GLASS (Treu et al., 2022). In order to compare our observed ratios to these, we dust-correct the emission line fluxes

reported in Table 5.1 using the dust properties obtained by PROSPECTOR in Section 5.5.2. One of the particularly noteworthy features about the emission lines seen in YD4 is the relatively low dust-corrected [O III]λ5007/[O II] ratio (hereafter O32) of $\log_{10}(\mathrm{O32}) = 0.12 \pm 0.01$. In combination with the high dust-corrected ([O III]+ [O II])/Hβ ratio (hereafter R23) of $\log_{10}(\mathrm{R23}) = 1.18 \pm 0.03$, these ratios indicate that YD4 resides in a poorly populated region of the excitation-ionisation plane, shown in Figure 5.3.

The R23 ratio is often taken as a tracer of the gas-phase metallicity (Curti et al., 2022), while O32 is a well-known tracer of ionisation parameter (Kewley et al., 2019). Therefore, these ratios seen from YD4 are indicative of a relatively low ionisation parameter yet a relatively high metallicity. The low O32 ratio appears consistent with a stack of $z \sim 2-3$, $\log_{10}(\mathrm{M}_\star[\mathrm{M}_\odot]) \sim 9.5$ galaxies from the MOSDEF survey in Sanders et al. (2021). This suggests that YD4 may be more evolved than the extreme O32 ratio galaxies that have been seen in the early Universe. These results therefore pose the question of whether the presence of a Balmer break is indeed connected to observing a low O32 ratio and a high R23 ratio.

5.4. STACK

In the following section we analyse galaxies that lie at a similar position on the R23-O32 plane, to establish whether they also host old stellar populations. However, continuum detections at $\lambda_{\mathrm{rest}} > 0.36\mu$m remain challenging, even in the era of *JWST*. The decrease in the sensitivity of NIRSpec with increasing wavelength often makes identifying Balmer breaks in galaxies difficult. While deeper observations with future *JWST* programs will facilitate the confirmation of Balmer breaks in candidate galaxies, however, these observations are not currently available. We therefore employ a stacking technique, allowing us to identify the presence of a Balmer break within a carefully selected sample of galaxies with similar emission line properties to YD4.

We utilise all galaxies in the public NIRSpec Prism observations by the JADES and CEERS surveys of the GOODS-South and EGS fields. We select our sample based on dust-corrected emission line fluxes taken from Cameron et al. (2023b); Tang et al. (2023); Nakajima et al. (2023); Mascia et al. (2023); Sanders et al. (2023). We select all galaxies at $z > 5.5$, observed with NIRSpec prism, located in the bottom right region of the excitation-ionisation plot, shown in Figure 5.3, similar to YD4:

- $\log_{10}\mathrm{O32} < 0.7$
- $\log_{10}\mathrm{R23} > 0.9$.

We identify a total of 10 galaxies between $5.5 < z < 8.7$ that satisfy these conditions: JADES-GS+53.17582–27.77446, JADES-GS+53.15613–27.77584, JADES-GS+53.11911–27.76080, JADES-GS+53.12259–27.76057, JADES-GS+53.13002–27.77839 (Cameron et al., 2023b), CEERS-1029, CEERS-1023, CEERS-1102 (Tang et al., 2023), CEERS-00403, CEERS-00545 (Nakajima et al., 2023).

We first shift each galaxy's spectrum to a common rest-frame wavelength grid, with a resolution of 0.001 μm by exploiting the SpectRes code from Carnall (2017). We then normalise all spectra by their UV flux at 0.15μm and perform a median stacking. The efficacy of the median stacking approach has been shown by numerous previous works (e.g. Witten et al., 2023; Roberts-Borsani et al., 2024; Harikane et al., 2020). We estimate the uncertainty in our spectra using bootstrap sampling. We produce 1,000 median stacks by randomly selecting N galaxies from our sample allowing for replacement, where N is the number of galaxies in each sample, and median stacking the spectra of the selected galaxies. When a galaxy is selected to be included in the stack we redraw the galaxy's spectrum from Gaussian's centred on the flux in a given wavelength bin and the standard deviation of the Gaussian is given by the uncertainty in the flux. We calculate the median and standard deviation of the 1,000 stacked spectra and report this as the final median stacked spectrum and its associated uncertainty.

The final stacked spectrum can be seen in Figure 5.4. A Balmer break can clearly be identified and we find a break strength of $B = 1.3 \pm 0.2$. This result shows that a low O32 ratio is linked to the presence of a Balmer break. The association between a Balmer break and low O32 ratio is curious because they are each associated with different stellar ages. The Balmer break is produced by older stars ($\gtrsim$50 Myr) whereas strong nebular emission is driven by young stars ($\lesssim$10 Myr). This may imply some direct causal relationship between a previous generation of stars and the young stars, such as metal enrichment of the ISM, or that the older stars are a good proxy for another galactic property. We investigate this association further in Section 5.5.

It is important to note that we are also selecting based on the R23 ratio. Therefore, in order to confirm that the discovery of a Balmer break is not only driven by a metallicity dependence, and hence largely driven by R23, we also stack galaxies with our R23 selection, but with $0.7 < \log_{10}(\mathrm{O32}) < 1.0$, contrasting our typical cut of $\log_{10}(\mathrm{O32}) < 0.7$. The selection criteria can be seen in blue in Figure 5.3 and correspond to a higher ionisation parameter, while falling below the most extreme O32 ratio galaxies seen by *JWST*. Following the same stacking procedure as discussed above in Section 5.4, we find very different spectral properties. The higher O32 and hence higher ionisation parameter stack reveals a multitude of rest-frame UV and optical emission lines as well as a UV-optical continuum

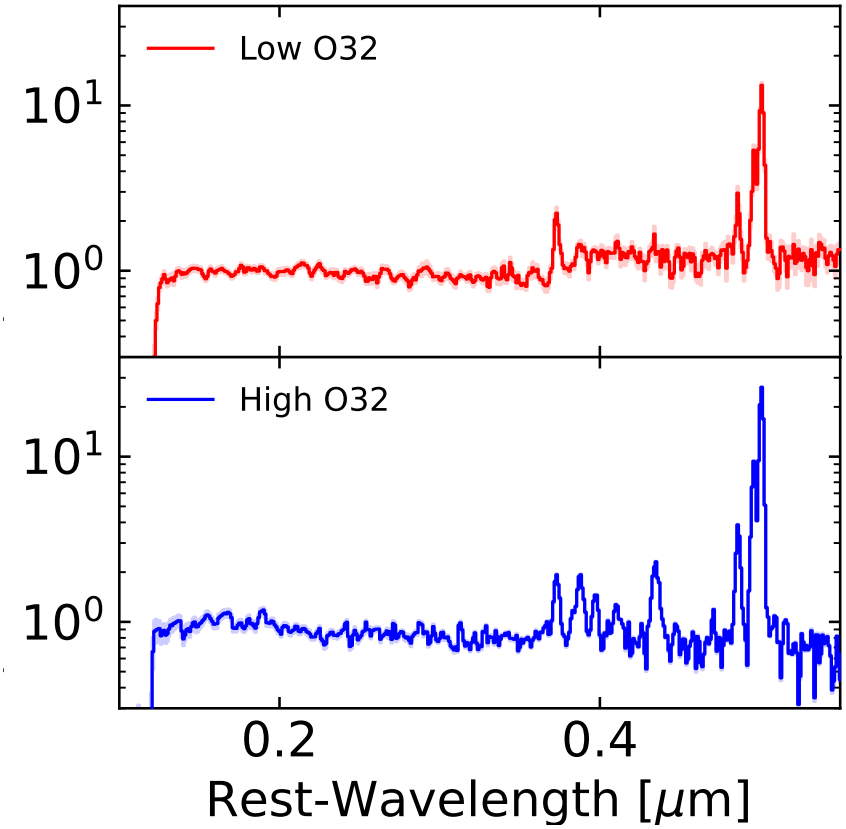


Fig. 5.4 The median stacked spectra of our two selection criteria, shown in Figure 5.3, where red (top panel) indicates those selected with a lower O32 ratio and blue (bottom panel) are those with a higher O32 ratio. The higher ionisation parameter stack shows the presence of multiple optical emission lines and indications of multiple UV emission lines, unlike the lower ionisation parameter stack. The low ionisation parameter stack shows a strong Balmer break, contrary to the flat UV-optical continuum seen in the high ionisation parameter stack.

that is consistent with no break ($B = 1.0 \pm 0.2$), as shown in Figure 5.4. The strength of the Balmer break is shown in Figure 5.2 and appears consistent with a stack of star-forming galaxies at similar redshifts from Roberts-Borsani et al. (2024) and the median galaxy from the SPHINX[20] simulation (discussed further in Section 5.5.3). This strongly implies that the presence of a Balmer break in the low O32 sample is linked to both relatively high metallicity and low ionisation parameter.

5.5. MODELLING

In the previous section we have shown that the emission line ratios and Balmer break in YD4 are peculiar when we compare to the population of high-redshift galaxies. Moreover, when we select galaxies with similarly peculiar emission line ratios in our stack sample, we discover a Balmer break in the median stack. We have justified that, based on its emission line properties, YD4 appears not to be in an AGN phase and that the observed break cannot be explained by accretion disk emission, and for the same reasons, this is also true for our stacked spectrum. Therefore, in the following section we exploit multiple forms of stellar population modelling to constrain the relationship between the SFH and the observed emission line ratios.

5.5.1 Simple stellar population analysis

While we identify a Balmer break in our sample, it remains unclear how the break is connected to the production of the observed O32 and R23 emission line ratios. In order to understand the properties of the stellar populations that are able to produce these ratios we employ the photoionisation code `CLOUDY`. We create grids using the stellar evolution models BPASSv2.2.1 (Eldridge et al., 2017; Stanway & Eldridge, 2018) including binaries and a Chabrier (2003) initial mass function with a high-mass cut off of $300 M_{\odot}$. We use stopping conditions based on electron fraction $e_{\rm frac} = 0.01$ and temperature $T = 3,000$K. We assume spherical geometry with an inner radius of 0.1 pc. We vary the ionisation parameter, metallicity, stellar age and density over the given ranges:

- log$U$ = [-3.0, -2.5, -2.0, -1.5, -1.0]
- $\log_{10}(Z/Z_{\odot})$ = [-2 : -1] in steps of 0.25, [-1 : 0] in steps of 0.1
- ages [Myr] = [1, 3, 5, 7, 10, 25, 50, 100]
- $\log_{10}(\rho[\mathrm{cm}^{-3}])$ = [2 , 3, 4] .

We note that the failure of photoionisation models to recreate the O32 and R23 emission line ratios at a sufficient strength as to match observations has often been noted at high redshifts (e.g. Roberts-Borsani et al., 2024; Cameron et al., 2023b). One potential solution to this is to account for turbulence when modelling emission line ratios. Gray & Scannapieco (2017) find that when they include a turbulent velocity of ~ 40 km/s in their modelling they can increase the R23 ratio by ~ 0.3 dex which would in turn shift our CLOUDY models into the shaded region in Figure 5.3. While our simple CLOUDY models also fail to reproduce the observed O32 and R23 ratios for YD4 and our sample of galaxies at any of the considered stellar ages and densities, extrapolating from trends shown in Figure 5.3 allows us to understand the likely required properties of YD4 and our stack. As discussed in Section 5.3, O32 is a well known tracer of ionisation parameter, and the low O32 observed in YD4 and our stack galaxies relative to the general population implies that the galaxies we study in this paper have a low ionisation parameter of around $\log U \sim -2.5$. The CLOUDY tracks show a turnover with increasing metallicity, occurring at a lower R23 ratio than that seen in our sample of galaxies, however, given the overall trend of increasing metallicity with R23 ratio and a higher R23 ratio in our galaxies than the general population, we conclude that our sample appears to be more chemically enriched than the general population and that the metallicity is likely $\log_{10}(Z/Z_{\odot}) \sim -0.5$ – the values observed at the turnover of R23. We also note that this enrichment is reflected in the observed ALMA dust detection for YD4 and the high dust attenuation predicted by PROSPECTOR, in Section 5.5.2, for both YD4 and the stack. This is similar to the conclusion of Cameron et al. (2023b) who suggest that their galaxies with the highest R23 and lowest O32 appear to be consistent with the turnover in the excitation-ionisation plane seen at $z \sim 0$ (Curti et al., 2020). Cameron et al. (2023b) conclude that this suggests that galaxies with these emission line ratios are likely the most chemically enriched in their sample, with metallicities of $0.3 - 0.5 Z_{\odot}$, in agreement with our conclusion.

Figure 5.3 shows the tracks for CLOUDY models with a stellar age of 5 Myrs and a density of $10^4 \mathrm{cm}^{-3}$. However we note that while changing these parameters can make differences in the exact position of the tracks, often along the direction of changing ionisation parameter, they do not help to reproduce the observed ratios and do not significantly affect the implied metallicity and the relatively low ionisation parameter. The main change in the emission lines is to their equivalent width (EW) which significantly decreases with age.

Previous modelling has shown that the strength of the [O III] emission line, the strongest optical emission line seen in the spectra of YD4 and our stack, falls rapidly with the increasing age of a stellar population. Faisst & Morishita (2024) have shown that

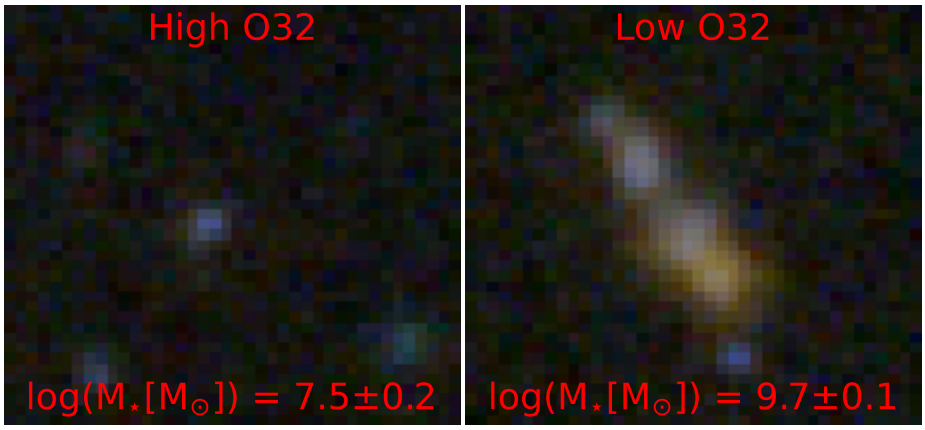


Fig. 5.5 Colour images of two galaxies selected from each of the high O32 (left) and low O32 (right) samples from Section 5.4. The images were obtained by combining F150W (blue), F277W(green) and F444W(red), retrieved from the Dawn *JWST* Archive. The stellar masses reported are taken from Morishita et al. (2024). The high O32 galaxy is more compact and has a lower stellar mass than the low O32 sample, as well as exhibiting bluer colours.

the [O III] emission line luminosity produced by a stellar population decreases by more than two orders of magnitude after just 5 Myrs due to the population's decreasing ionising emissivity. The detection and strength of the optical emission lines in YD4 and the stack therefore imply that these galaxies host stellar populations with ages of $\lesssim 10$ Myrs.

While the production of the emission lines by a young stellar population is not surprising, the requirement for a relatively low log$U$ and relatively high metallicity, compared to other high-redshift galaxies (e.g. Cameron et al., 2023b), is noteworthy. Significant previous work has been undertaken in understanding the trend of increasing ionisation parameter with redshift (e.g. Brinchmann et al., 2008; Bian et al., 2016; Papovich et al., 2022; Liu et al., 2008; Masters et al., 2016). This previous work targeting the evolution in log$U$ by $z \sim 2-3$ indicates that $U \propto \Sigma_{\rm SFR}$, the star formation rate surface density, and this relation has been linked, at least in part, to increasing electron density (Reddy et al., 2023a,b) which is known to evolve with redshift (e.g. Isobe et al., 2023). Unfortunately, the resolution of NIRSpec prism observations makes resolving the [O II] doublet, a key diagnostic of electron density (e.g. Sanders et al., 2016), impossible. However, thanks to analysis by Morishita et al. (2024), we can evaluate the $\Sigma_{\rm SFR}$ using their reported SFRs and effective radii. We find no reduction in the $\Sigma_{\rm SFR}$ of our low O32 sample relative to the higher O32 sample in Section 5.4. However, this relationship has large scatter (Reddy et al., 2023a,b) and our results are likely driven by the minimal overlap between our samples and those of Morishita et al. (2024), thus producing a very small sample size. Moreover, our higher O32 sample does not include the most extreme O32 ratios seen in Figure 5.3.

The galaxies in our sample that do overlap with Morishita et al. (2024) appear to have larger effective radii and stellar masses than both the higher O32 sample, and the general galaxy population at their given redshift. We see a qualitative trend towards more extended structures from NIRCam imaging of all of our low O32 sample. They also show much redder colours than the high O32 stack, relating to their much redder UV slope caused by their more evolved state, an example of which can be seen in Figure 5.5.

We therefore interpret the reduced O32 ratio in our sample as evidence of an enriched ISM and a low ionisation parameter. This is likely driven by a lower SFR surface density, and therefore lower electron density Reddy et al. (2023a). Given the association we find between the low O32 ratio and the presence of a Balmer break, these ISM conditions seem to be driven by the presence of an old stellar population that has chemically enriched the ISM, via supernovae events. Additionally, the Balmer break indicates the presence of an older stellar population and hence a more evolved galaxy which may be more extended and transitioning away from the bursty mode of star formation, and thus creating a reduced SFR surface density (see e.g. Hopkins et al., 2023; McClymont et al., 2024b).

However, in order for a full analysis of the properties driving the low O32 ratio, a larger sample, and high resolution spectroscopy of the [O II] doublet are required, facilitating the de-blending of the [O II] doublet (e.g. Isobe et al., 2023). As it stands, our sample is currently too small to make any such conclusions, however, future *JWST* programs aimed specifically at observing candidate (mini-)quenched galaxies, but moreover, general large surveys, will inevitably unveil a larger population of rejuvenating galaxies.

5.5.2 SED fitting

The detection of the Balmer break within the spectra of YD4 and our stack is noteworthy, not just because it indicates that an old stellar population is present, but implies that the young stellar population is not dominating the SED of these galaxies, likely indicating the presence of very few, very young stars. Constraining what star-formation history is capable of producing these features, and understanding the final chemical and dust enrichment is crucial in understanding the link between the O32 and R23 line ratios and the Balmer break seen in YD4 and our stack.

In this context, we expand our analysis of YD4 and our stacked spectrum by fitting them with the SED-fitting code `PROSPECTOR` (Johnson et al., 2021). We chose to exploit `PROSPECTOR` as a modelling tool thanks to its flexible dust attenuation curve and use of BPASS models. The use of BPASS models (Eldridge & Stanway, 2009) are especially noteworthy in the context of the ionising output of young stellar populations, where the ionising photon emissivity declines rapidly for the Bruzual & Charlot (2003) single-star

models while this decline is more gradual for BPASS models, when we account for the impact of binary stars. As such, binary evolution can have a strong impact on the nebular emission modelling (Xiao et al., 2018; McClymont et al., 2024a) and is therefore important for constraining the exact SFH of galaxies that have ongoing star formation, like those studied in this paper. We fit both the photometric data and the spectroscopic constraints on the emission line fluxes (seen in Table 5.1 for YD4) following a similar approach to Tacchella et al. (2023a). We utilise a non-parametric SFH from Leja et al. (2019). The photometric constraints of YD4 were obtained in the framework of the UNCOVER survey (GO 2561, P.I.:Ivo Labbe) and PSF matched to F444W (Weaver et al., 2024), while for the stack we artificially produce the photometry by passing the spectrum through the transmission curves of the *JWST* filters. We fix the redshift of both YD4 and the stack at the spectroscopic redshift measured in Morishita et al. (2023) ($z = 7.88$) and re-normalise the stack to have the same UV luminosity as YD4. This is motivated by our intention to place the SFHs on comparable scales, we note this will result in stellar masses and formation redshifts that are not representative of the median galaxy in our stack sample, but allows us to make comparisons between the shapes of the SFHs.

We make use of a similar model to that described in Tacchella et al. (2023a), employing a 16-parameter physical model and priors. We fit for the stellar mass, stellar metallicity, gas-phase metallicity, ionisation parameter and dust attenuation. Our dust attenuation modelling includes three-components – a diffuse component over the entire galaxy, an additional component attenuating the light from stars less than 10 Myrs old and finally a flexible slope, following the prescription of Conroy et al. (2009). Our non-parametric SFH employs a continuity prior, described by a Student's t-distribution with $\sigma = 0.3$ and $\nu = 2$, on the relative star formation in each adjacent age bin. We fit nine parameters that control the ratio of the average star-formation rate within ten adjacent bins. These bins span the lookback time range of $0 - 284$ Myrs. We chose to use a coarse binning in time to avoid an over-interpretation of our results with `PROSPECTOR`. While the emission lines allow constraints on the star formation on ~ 10 Myr timescales, the Balmer break constrains star formation on larger timescales (~ 100 Myr), meaning information is only available to constrain the SFH on such scales. As such we exploit bins first covering 0-5 Myrs and 5-10 Myrs time range, while the remaining eight bins are log-spaced to $z = 12$ (lookback time of 284 Myrs). Therefore, when we make statements about the duration of burst and quenched periods, these should be taken as approximate values, noting the bin widths covering these periods.

`PROSPECTOR` models nebular emission line fluxes using the Flexible Stellar Population Synthesis code (FSPS; see Byler et al., 2017), which in turn is based on `CLOUDY` (Ferland

et al., 2013) photoionsation models. This is notable as we show in Figure 5.3 that CLOUDY modelling is unable to reproduce the dust-corrected emission line ratios with a single stellar population. However, the more complex agglomeration of stellar populations that PROSPECTOR employs is able to reproduce the observed emission line fluxes within their associated uncertainties.

The best fit model of YD4 is obtained with the following set of parameters: stellar mass $M_\star$=$(4.1^{+0.2}_{-0.2})\times10^9$ M$_\odot$, stellar metallicity $\log_{10}Z_\star/Z_\odot = -2.84^{+0.1}_{-0.09}$, gas-phase metallicity $\log_{10}Z_{\rm gas}/Z_\odot = -0.31^{+0.01}_{-0.01}$, ionisation parameter $\log U = -2.49^{+0.01}_{-0.01}$, diffuse dust V-band optical depth $\tau = 0.67^{+0.06}_{-0.03}$. The inferred metallicity and ionisation parameter values are in agreement with the values that we predicted from our CLOUDY modelling.

We do not report all of the best fit parameters derived for the fitting of the stack, as our choice of redshift and normalisation, to match the spectrum of YD4 means properties like the stellar mass do not relate to the actual stellar mass of the median galaxy in our stack. However, we note that the implied ionisation parameter ($\log U = -1.93^{+0.07}_{-0.05}$), metallicity ($\log_{10}Z_{\rm gas}/Z_\odot = -0.52^{+0.01}_{-0.01}$) and dust ($\tau = 0.96^{+0.02}_{-0.04}$) are technically statistically inconsistent with those of YD4. However, the errors output from PROSPECTOR are likely significantly underestimated, especially given the relatively large differences in parameters measured with different SED-fitting codes (e.g. Looser et al., 2024). In this particular case, the small uncertainties may be driven by the limited number of models that can reproduce the observed emission line ratios and UV- and optical-continuum. Given these underestimated uncertainties the metallicity, ionisation parameter and dust values of the stack are consistent with the values seen in YD4. These results again strongly imply that these specific values for these parameters are intricately linked to the position in the excitation-ionisation plot and moreover are affected by the presence of an old stellar population, given the inferred SFH discussed below.

The best fit SFH and its uncertainty, achieved by our PROSPECTOR fitting of both YD4 and our stacked spectrum, can be seen in Figure 5.6. These SFHs have been rescaled so that their maximum SFRs are unity. This allows us to compare the shapes of their SFHs. In the SFH of YD4, the first star-forming event seen is a major burst beginning ~ 125 Myrs pre-observation and lasting for nearly 100 Myrs. We then see a 25 Myr period of quiescence, followed by a burst of star formation that reaches 60% of the maximum star formation rate, but its relatively short duration means it forms just 5% of the total stellar mass.

We remind the reader that our investigation of the SFH of the stack is motivated as a comparison to YD4. The shape of the SFH of the stack is notably similar to YD4, with a significant burst finishing ~ 35 Myrs in the past, followed by a lull in star formation.

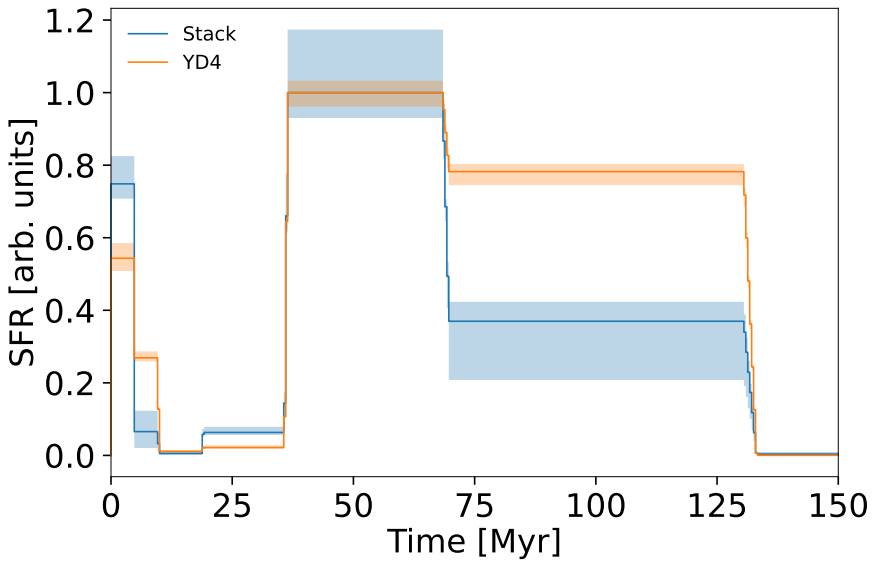


Fig. 5.6 The SFH derived from our PROSPECTOR fitting of the photometry and emission lines of YD4 (orange) and our stacked spectrum (blue). The average SFR within each SFH bin is shown as a flat line across the width of the age bin. The star-formation rates have been normalised by the maximum SFR in the SFH for ease of comparison.

Finally, a recent ($t < 5$ Myrs) burst of star formation is seen. This recent burst is at 75% of the maximum SFR seen in the stack and has a slightly smaller old-stellar population than YD4 forming fewer stars > 75 Myrs ago. However, the recent burst in the SFH of the stack still only forms 6% of the total stellar mass.

In Figure 5.7 we consider the relative contributions to the overall spectrum of YD4 from the four main stellar populations: 0-5 (red), 5-10 (blue), 35-70 (pink) and > 75 Myrs (purple). This comparison allows us to understand why PROSPECTOR preferentially obtains a rejuvenating SFH. The youngest age bin contributes almost all of the nebular emission line flux, while the age bins above 35 Myrs dominate the continuum emission. The 5-10 Myrs bin contributes to the emission line fluxes, however this contribution is significantly smaller than that from the 0-5 Myr bin, even when normalising by the stellar mass in each bin. This decreasing emission line flux with age, and increasing Balmer break strength with age provides an explanation for the rejuvenating solution. Both a young and old stellar population are required to provide the emission lines and Balmer break, respectively. However an intermediate stellar population would both act to drown out the Balmer break, but equally importantly would reduce the EW of the emission lines. In order to reproduce the observed equivalent widths when introducing an intermediate age stellar population, one would need to increase the relative contribution of the youngest age

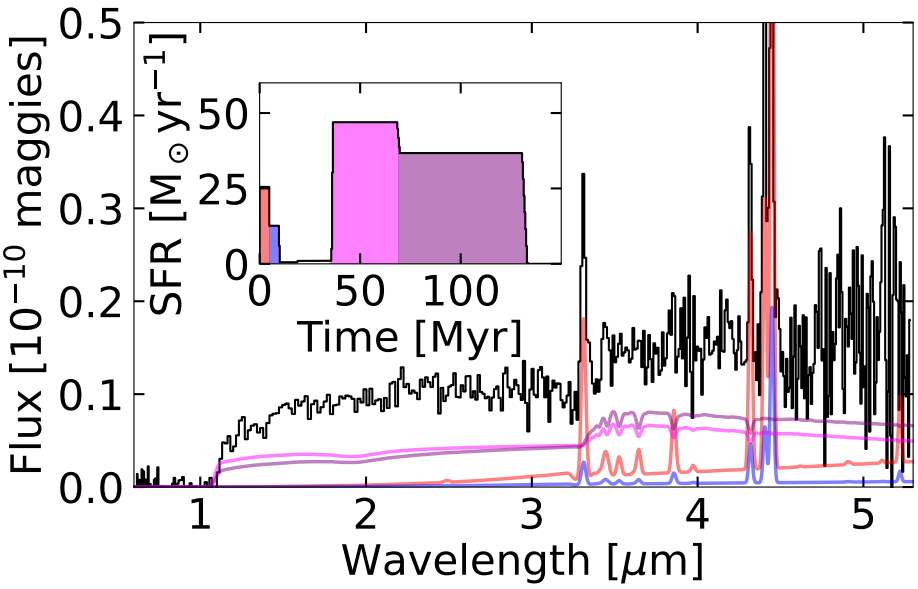


Fig. 5.7 The NIRSpec Prism spectrum of YD4 shown in black. The SFH inferred from PROSPECTOR is shown in the inset panel. The components of the spectrum attributed to the main star formation bins are shown by different colours: 0-5 Myrs (red), 5-10 Myrs (blue), 35-70 Myrs (pink) and > 75 Myrs (purple). The old stellar population (> 35 Myrs) dominates the continuum emission, while the young population (0-10 Myrs) produces all of the observed nebular emission lines.

bins. However, this in turn would act to reduce the Balmer break strength, thus requiring a larger old stellar population. Therefore, a careful ratio of the young to old age stellar populations is required with a very small intermediate age stellar population.

Our PROSPECTOR fitting strongly supports the conclusion we reach using CLOUDY – that these galaxies are chemically enriched and that a low ionisation parameter is seen. Moreover, they provide an explanation for this evolutionary state – that these galaxies have previously hosted a large episode of star formation that has now matured and has hence enriched their ISM. PROSPECTOR also indicates a strong preference for the young stellar population to be a very small fraction of the total stellar mass. This is likely driven by the fact that a relatively low-mass, young stellar population is considerably brighter than older stellar populations of considerably greater stellar mass. This results in an "outshining" effect of the old stellar population (see Papovich et al., 2001; Pforr et al., 2012; Conroy, 2013; Tacchella et al., 2023a; Whitler et al., 2023). Further confounding this problem is the dominance of inverse Balmer jumps in the spectra of high-redshift, young galaxies (e.g. Roberts-Borsani et al., 2024) that will act to cancel out the Balmer break produced by old stellar populations.

In this context, we investigate how a young stellar population can wash out a break, by modelling rejuvenating galaxy spectra with the SED-fitting code PROSPECTOR (Johnson et al., 2021). We create model spectra for a galaxy with the same properties (e.g. gas-phase and stellar metallicity, logU and dust properties) as YD4, as determined by PROSPECTOR, as described above. In order to investigate the "outshining" abilities of young stellar populations, we vary the fraction of the total stellar mass formed within the last 10 Myrs (the duration of the most recent star forming burst). We then enforce a quiescent period with duration $\Delta t_{\text{Quiescent}}$. This is preceded by the first star formation event, a burst lasting for $\Delta t_{\text{Old Pop.}}$. The varying effects of these can be seen in the upper panel of Figure 5.8. The results of this show that regardless of the time spent in a quiescent stage a young stellar population with a mass fraction of 30% can effectively remove the presence of a Balmer break in a galaxy like YD4. With a shorter period of time spent quiescent, the duration over which the old stellar population formed becomes more important given that the strongest Balmer breaks are produced by stars that are $t \gtrsim 50$ Myrs old. However, as we move towards a lengthy period of time spent in a quiescent state, all of the old stellar population produce large Balmer breaks, regardless of how close to the beginning of the quiescent period that they formed, removing any relation with $\Delta t_{\text{Old Pop.}}$.

In the lower panel of Figure 5.8 we emphasise the impact that varying the diffuse dust V-band optical depth, τ, has on the Balmer break. In order to estimate the impact of varying τ we create models using the properties and SFH of YD4 that are a result of our PROSPECTOR

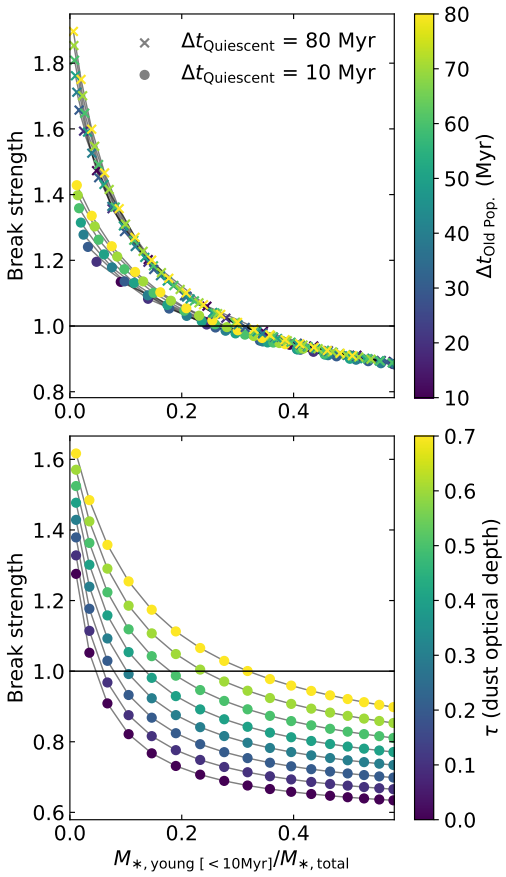


Fig. 5.8 Balmer break strength, as defined by Roberts-Borsani et al. (2024), as a function of the fraction of the total stellar mass produced in the most recent 10 Myr. (Top panel) The effects of changing the time the galaxy spends in a quenched phase (different markers) and the time over which the old stellar population formed (colorbar). The model utilises the galaxy properties of YD4 reported in Section 5.5.2, while varying the SFH. We vary both the time bins, three bins spanning 0–10, 10–(10+$\Delta t_{\rm Quiescent}$) and (10+$\Delta t_{\rm Quiescent}$)–(10+$\Delta t_{\rm Quiescent}$+$\Delta t_{\rm Old.\ Pop.}$) Myrs, and the mass within each bin, $M_{\star,\rm young[<10Myr]}$, 0 and ($M_{\star,\rm total} - M_{\star,\rm young[<10Myr]}$), respectively. When the young stellar population hits ~ 30% of the total stellar mass, it completely outshines the old stellar population. (Lower panel) The impact of dust attenuation (colorbar) on the observed Balmer break strength. We exploit the properties and SFH of YD4 inferred from our PROSPECTOR SED-fitting while varying the dust V-band optical depth and the ratio of stellar mass formed in the two star formation bursts shown in Figure 5.6. Increasing the amount of dust increases the observed Balmer break. Failing to account for dust will therefore bias the age estimate, however, our PROSPECTOR SED-fitting finds a high value of $\tau \sim 0.7$ and hence we are indeed accounting for significant dust attenuation. Given this, the inferred age of YD4 from PROSPECTOR is likely a conservative estimate. An increasing young stellar population reduces the Balmer break strength at a similar rate regardless of the value of τ.

modelling above. As we decrease τ from the value observed in YD4 ($\tau \sim 0.7$) we find a decreasing Balmer break strength that is driven by the decreasing attenuation of the UV continuum relative to the optical continuum. However, what becomes apparent from this plot is that even when no dust is present, a Balmer break would be observed in YD4, but only if there is a negligibly small young stellar population. This dust-free Balmer break is "outshone" very efficiently by a young stellar population with a stellar mass of just $\sim 5\%$. However, given that YD4 has a red UV continuum and ALMA dust detections, we know it hosts a significant mass of dust ($10^{6.5} M_{\odot}$; Hashimoto et al., 2023), thus constraining us to the high optical depth regime. Combining this with the knowledge that YD4 exhibits strong nebular emission lines but also a Balmer break ($B \sim 1.3$), evidencing the presence of a young stellar population and a significant old stellar population, we can identify that YD4 can only have a young stellar population with stellar mass fraction of $\lesssim 10\%$. Reducing the quiescent period length, as in the upper panel of Figure 5.8, can only act to decrease the implied stellar mass of the young stellar population, in order to retain the break strength. However, the presence of strong nebular emission lines helps PROSPECTOR to further constrain the fraction of total stellar mass from the young stellar population to be $\sim 5\%$, as discussed further above.

The implication therefore is that YD4 and our stack host both young and old stellar populations, but are in a very short-lived phase where these are both identifiable. With a larger young stellar population ($\sim 30\%$), the Balmer break seen in YD4 would be entirely lost, while with a slightly more mature young stellar population the emission lines would not be seen.

5.5.3 SPHINX

Given the combination of an evolved galaxy hosting a young stellar population, that is growing up in a chemically enriched and disrupted ISM, we employ a more complex comparison to the SPHINX cosmological radiation-hydrodynamic simulations (Rosdahl et al., 2018, 2022). We make use of the SPHINX^{20} public data release (Katz et al., 2023a). The detailed modelling of the multi-phase ISM in SPHINX is ideally suited to understanding the evolution of emission lines originating from evolved galaxies. In this context, we first consider the ratio between the recent SFR over 10 Myrs and the SFR over 100 Myrs. If our conclusion regarding the required relatively low mass of the young stellar population in these rejuvenating galaxies is accurate we would expect a strong dependence on $\mathrm{SFR}_{10}/\mathrm{SFR}_{100}$ to be seen in the excitation-ionisation plot of SPHINX^{20} $z \gtrsim 6$ galaxies. We show the excitation-ionisation plot, using dust-attenuated emission line ratios, in

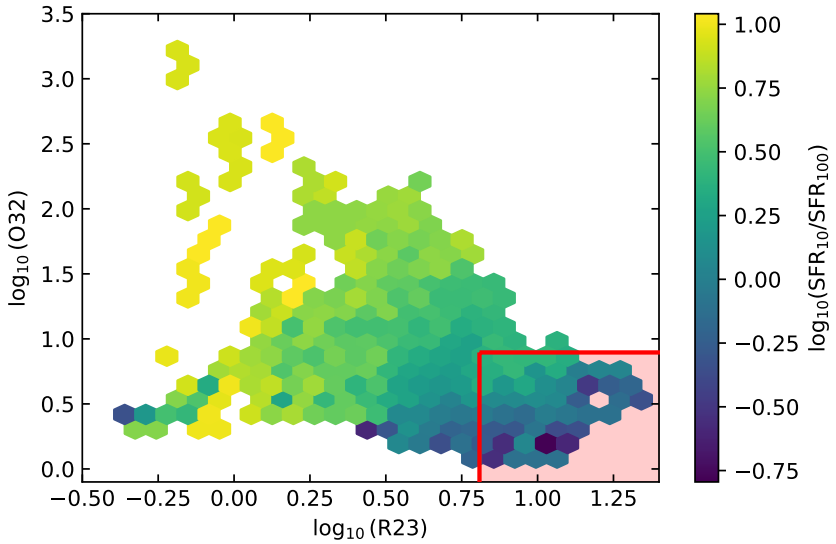


Fig. 5.9 The O32 ratio against R23 ratio of mock observed galaxies, color coded by the average ratio of the recent SFR over 10 Myr and the SFR over 100 Myr in each hexagonal bin. These mock observed galaxies are taken from the SPHINX[20] simulation, which includes the effects of dust attenuation (see Rosdahl et al., 2018, 2022; Katz et al., 2023a). The ratios are shown for all sightlines of all galaxies at $z \geq 6$. There is a clear correlation between a decline in the ongoing SFR (i.e. a low SFR_{10}/SFR_{100}) and the emission line ratios we target, facilitating the observation of the Balmer break. The red shaded area (background) and red box (foreground) indicates the same region as in Figure 5.3 but before accounting for dust corrections. A trend towards a decrease in recent star formation is seen in SPHINX[20] galaxies towards the location of our sample of low O32 and high R23 sample of galaxies, as expected from galaxies that have recently experienced a mini-quenched period.

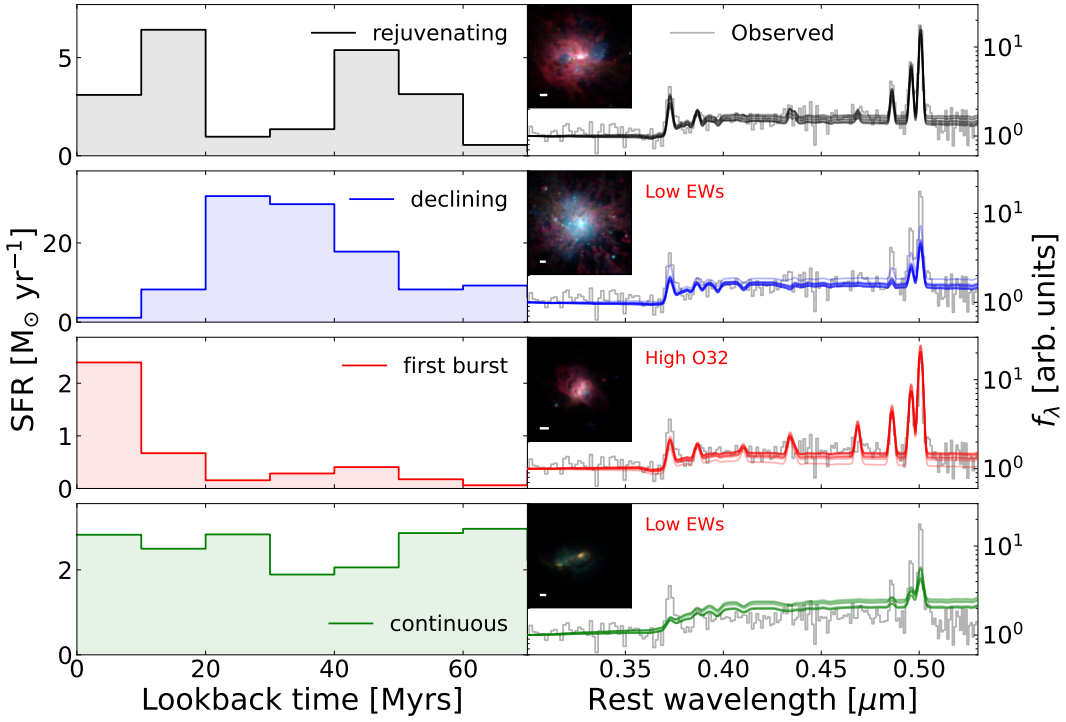


Fig. 5.10 SPHINX simulated galaxies showing the effects of different SFHs on the observed spectrum: rejuvenating (black), declining (blue), first burst (red) and continuous (green). Left panels show the SFH of each galaxy, where the type of SFH is indicated in the legend. Each SFH has been re-binned to show the average SFR within a 10 Myr bin, variations below this resolution will not significantly effect the observed spectrum and hence are not shown for clarity. Right panels show the corresponding spectra of the galaxy along the 10 lines-of-sight available. The spectrum of YD4 is shown in grey for comparison. All spectra are normalised by their UV flux at 0.15 μm. The inset panel shows the RGB image along one line-of-sight, using mock F150W, F277W and F444W images, of the SPHINX galaxies, where the white bar indicates a length of 1 pkpc. The only SFH that recreates both the emission line ratios and equivalent widths seen in YD4 are those that host a recent rejuvenation event. The rejuvenating galaxy and the declining galaxy both show large spatial extents relative to the first burst and continuous SFH galaxies.

Figure 5.9 and demonstrate there is a strong correlation between a decrease in SFR and the emission line ratios we target.

We also wish to assess whether the SPHINX galaxies that fall within the same selection criteria used in Section 5.4 can replicate the observed spectrum of YD4. In Figure 5.10 we present four different regimes of SFHs seen in the SHPINX simulations. The upper panel shows the SFH and spectrum of a rejuvenating galaxy. The SPHINX snapshot catches the galaxy on the downturn of its rejuvenation event, but just recently enough that the emission lines are still visible. We note that catching this galaxy at a slightly earlier epoch, during its burst, would help to boost the EW of the emission lines that are currently slightly lower, but still within the errorbars of the *JWST* spectrum, than those seen in YD4. This SFH is able to replicate the spectrum of YD4 within its uncertainties, both in terms of emission line ratios and EWs, as well as the Balmer break strength. Conversely, the declining SFH produces a spectrum that has emission lines that have much lower EWs than those seen in YD4, but the emission line ratios are well replicated, suggesting that the emission line ratios are indeed driven by the effects on the ISM of an old stellar population. The "first burst" galaxy shows a SFH that is dominated by a single burst of star formation that is ongoing. While the presence of star formation over $\sim 100-200$ Myr before observation accounts for the observed Balmer break, the emission lines from the dominant burst fail to replicate the emission line ratios of YD4, with too little [O II] emission and too much [O III] emission. Finally, the continuous SFH produces a significant Balmer break but fails to reproduce the emission lines. The only galaxy that reproduces the emission line ratios, equivalent widths and Balmer break of YD4 is the galaxy that shows a recent rejuvenation event following a period of quiescence that lasted ~ 20 Myr, in strong agreement with the SFH of YD4 that is produced with `PROSPECTOR` SED-fitting.

However, what is clear is that the young stellar population in the rejuvenating galaxy we took from the SPHINX^{20} simulation has a considerably larger fraction of the stellar mass built up in the $0-70$ Myr history, shown in Figure 5.8, than we derived for YD4. This is driven by the more complex SFH that is modelled in the SPHINX^{20} simulation, which shows an extended SFH over some hundreds of Myrs that helps to build up stellar mass and the Balmer break.

Therefore, we consider the Balmer break strength, using the definition discussed in Section 5.3, seen in SPHINX^{20} galaxies as a function of mass-weighted age in Figure 5.11. The break strength observed in the stellar continuum is, unsurprisingly, higher than that observed in the total galaxy spectrum. This is driven by the presence of inverse Balmer breaks in the nebular component of the total spectrum reducing the total break strength. This effect can also be seen in the break strengths of the BPASS SEDs. These

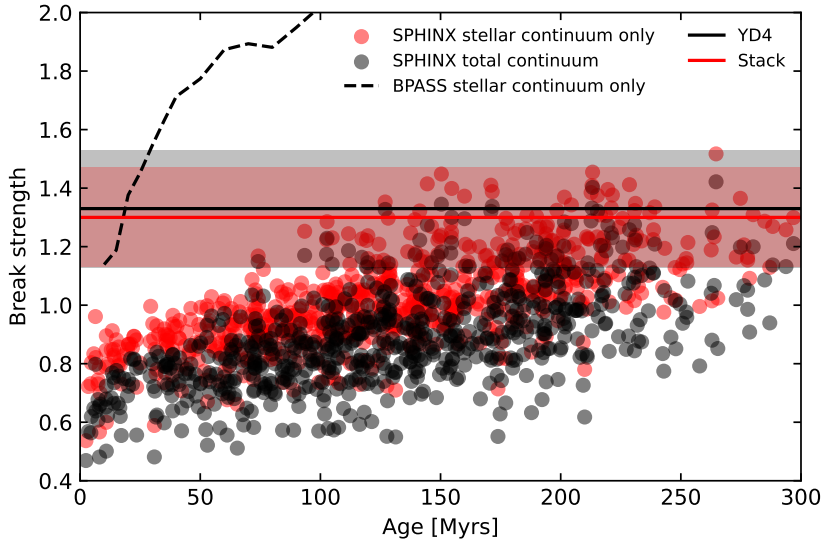


Fig. 5.11 The Balmer break strength as a function of stellar-mass-weighted age in the SPHINX[20] public data release (Rosdahl et al., 2018, 2022; Katz et al., 2023a) at $z \geq 5$. The break strengths seen in the stellar continua (red) and total (stellar and nebular) continua (black) are shown. The Balmer break strength seen in the stellar continuum of the BPASS models with a metallicity of $\log_{10} Z/Z_{\odot} = -0.5$, as a function of the age of the model, is shown by the black dashed line. Finally, the black and red solid lines and shaded regions indicate the break strength and its associated uncertainty for YD4 and the stacked spectrum, respectively.

breaks are considerably stronger than those seen in the SPHINX^{20} galaxies due to our purposeful neglect of nebular emission in our BPASS break strength measurement. These models use a single stellar population of a given age, however the SPHINX^{20} galaxies are accumulations of various stellar populations of different ages. Young stellar populations show no evidence of Balmer breaks in their stellar continua and moreover dominate the total galaxy spectrum. This results in the significant reduction in the break strength of galaxies composed of multiple stellar populations. Thus, when we compare to the more realistic SPHINX^{20} galaxies, the break strengths seen in YD4 and the stack suggest that these galaxies host stellar populations with ages between 75−300 Myrs. The mass weighted ages of YD4 and the stack implied from their SFHs are 80 Myrs and 70 Myrs, respectively. While modelling differences may drive this inconsistency between the ages implied from PROSPECTOR and SPHINX^{20}, the ease for a relatively young stellar population to wash out a Balmer break (shown in Figure 5.11), the effect known as outshining (e.g. Papovich et al., 2001; Pforr et al., 2012; Conroy, 2013; Tacchella et al., 2023a; Whitler et al., 2023), is likely to be playing a crucial role here. This degeneracy means that we are potentially missing significant amounts of stellar mass when we fit these apparently young, bursty galaxies with SED-fitting codes. This is again underlined by the SPHINX^{20} prediction that YD4 and the stack host older stellar populations than is predicted by PROSPECTOR – only the most recent major burst can be identified based on the presence or lack of a Balmer break. This implies that YD4 has a formation redshift of at least $z \sim 9$ and potentially as early as $z \sim 12.5$, underlining the challenges that such bursty star formation histories provide when trying to diagnose the age and stellar mass of high-redshift galaxies.

5.6. DISCUSSION AND CONCLUSION

In this paper we have presented *JWST* NIRSpec PRISM spectra that show both a Balmer break and emission lines. These results are significant in the context of recent detections of (mini-)quenched galaxies at $z \gtrsim 5$ by *JWST* (Looser et al., 2024, 2023; Dressler et al., 2023, 2024; Strait et al., 2023; Trussler et al., 2024a). These galaxies have been observed to host Balmer breaks without the presence of emission lines. We have established that by selecting galaxies in the high metallicity and low ionisation region of the R23-O32 plane, the median galaxy at $z > 5.5$ hosts a Balmer break, in strong contrast to galaxies with higher O32 ratio, as well as the general star-forming galaxy population. We thus conclude that these abnormal emission line ratios are indeed linked to the presence of an old stellar population. CLOUDY modelling helps us to conclude that this appears to be driven by the chemically enriched and disrupted ISM that is produced by these

older stellar populations. These results strongly imply the presence of both an old stellar population (∼ 100 Myrs) producing the Balmer break and a young stellar population ($\lesssim$ 10 Myrs) driving the observed strong nebular emission lines.

These results naturally lead us to query what is the shape of the SFH producing these observed spectral properties. We employ a non-parametric SFH modelling approach, exploiting the SED-fitting code PROSPECTOR. The results of this strongly favour the conclusion that both YD4 and our stack are best described by a rejuvenating SFH. Our modelling of the Balmer break strength strongly constrains the young stellar population to be less than a tenth of the total stellar mass, while the strength of the observed emission lines constrains this fraction to close to 5%. Following this, we exploit the SPHINX^{20} simulations to further constrain the range of SFHs that can replicate the spectral features seen from YD4 and our stack. SPHINX^{20} galaxies show that a rejuvenating SFH is able to match the break strength, emission line ratios and EWs. Alternative SFHs can replicate some aspects of the observed spectrum of YD4 and our stack, but fail to match them all simultaneously.

While the declining SFHs in SPHINX^{20} fail to match the observed spectrum of YD4, we do believe that a carefully tuned gradual decline in star formation could feasibly result in a young stellar population that produces nebular emission lines alongside older stars which produce a Balmer break. Moreover, a galaxy with this star-formation history would host an enriched ISM, which is consistent with the emission line ratios. However, the high observed EW of the emission lines and the strength of the Balmer break place strong constraints on the required young and old stellar populations. As discussed in Section 5.5.2, significant intermediate star formation (10–50 Myr) would act to reduce the EW of emission lines and start to wash out a large Balmer break, in conflict with the observed data. Additionally, the strong preference of our SED fitting for little intermediate star formation may be informed by other constraints, such as the UV and optical slopes. We also note that this more gradual decline scenario is at odds with the rapid quenching timescales which have been inferred for other galaxies in the early Universe (Looser et al., 2024; Gelli et al., 2023; Dome et al., 2024; McClymont et al., 2024b).

All of the results of our modelling and analysis of the spectra of YD4, our stack and SPHINX^{20} galaxies strongly support the premise that we are observing rejuvenation in a galaxy with a chemically enriched, disrupted ISM. We find that a significant population of galaxies exist that host both young and old stellar populations evidencing their complex SFHs. These spectral properties are best explained by galaxies hosting multiple bursts of star formation. Moreover, bursty star formation may need to be invoked to understand the overabundance of UV bright galaxies in the early Universe (e.g. Kravtsov & Belokurov, 2024; Sun et al., 2023b; Mason et al., 2023; Shen et al., 2023; Mirocha & Furlanetto, 2023)

and may in part be a result of the rapid build up of stellar mass through galaxy mergers at extremely high redshifts, as seen in Witten et al. (2024b).

It is notable that the SFH of YD4 and our stacked spectrum, inferred from PROSPECTOR SED-fitting, implies that both must have experienced a quiescent period of ~ 25 Myr. Moreover, the rejuvenating galaxy taken from the SPHINX[20] simulations also shows a reduction in star formation over a 20 Myr timescale. This conclusion about the short timescales of quenching events has been suggested by previous works (e.g. Dome et al., 2024; Looser et al., 2024; McClymont et al., 2024b). The timescales between bursts is regulated by the mode of quenching. New star-formation events are as a result of either the in-fall of gas onto the galaxy, or the return of once feedback-affected gas to a cool and dense state within the galaxy, such that new stars can be formed. As a result, a wide range of quenching mechanisms exist: a lack of gas accretion from the CGM, a lack of gas cooling within the ISM, mechanisms stopping the cold gas from forming stars, the rapid consumption of the available cold gas within a galaxy or the removal of gas from the ISM (see Man & Belli, 2018). The short mini-quenched timescales inferred from our modelling of the spectra of YD4 and our stack suggests that the rapid consumption of all gas or strong AGN feedback scenarios are unlikely. This is further supported by the lack of evidence of AGN activity. Moreover, if mergers were driving the removal or rapid consumption of cold gas within these mini-quenched galaxies, and in turn driving the rejuvenation of star formation, from our quenching timescales we would expect galaxy mergers to be occurring at a frequency of one every few tens of Myr while observational evidence suggests the frequency is an order of magnitude less often (Puskás et al. in prep.). Instead, disruptive feedback driven by star formation or weak AGN feedback is more likely to quench galaxies on these short timescales. Through comparison with multiple simulations Dome et al. (2024) find that $z = 4 - 7$ mini-quenched galaxies are typically in a quenched state for 20-40 Myrs. They identify that this is consistent with the free-fall timescale of the inner halo, which further supports the conclusion that stellar feedback or weak AGN feedback is crucial in producing these bursty SFHs. These processes expel gas from the ISM leaving the galaxy in a mini-quenched phase for a few 10's of Myrs, after the free-fall time of the inner halo, this gas reaccretes and rejuvenates the galaxy.

This growing evidence that galaxies in the early Universe are composed of stellar populations produced by multiple bursts of star formation underlines the importance of accounting for this in our modelling of their SFHs. However, our modelling also shows the fragility of the Balmer break strength as a tracer of old stellar populations, due to the ease at which it can be drowned out by young stars. While the presence of a Balmer break can be taken as evidence of an old stellar population, its absence cannot be used to imply a

lack of one. The effects of stochastic SFHs will act to drown out the optical continuum of old stellar populations and hence will bias us towards underestimating the age and stellar mass of these galaxies in the very early Universe.

CHAPTER 6

CONCLUSIONS

In the previous Chapters I have discussed the properties of galaxies in the early Universe and the conclusions we can draw regarding their environments and the reionisation process. Below I repeat the key science questions outlined in Section 1:

i. Can enough ionising photons escape UV-bright galaxies as to drive reionisation?

ii. What facilitates the escape of Lyα photons from high-redshift galaxies?

iii. What are the environments of high-redshift galaxies?

iv. Do high-redshift galaxies show evidence of bursty star formation histories?

6.1. CONCLUSIONS

In Chapter 2, I exploited existing ground-based (VLT/X-shooter and Keck/MOSFIRE) observations of UV-bright ($M_{UV} \sim -22$) Lyα emitters at $z > 7$ to constrain their role in the reionisation process. While some models have suggested that UV-bright galaxies - "oligarchs" - account for at least 80% of the ionising budget (Naidu et al., 2020), star-forming galaxies are the leading candidates of the reionisation process (e.g. Ocvirk et al. 2020, Trebitsch et al. 2022). Given that ground-based observations of the Lyα line at high-redshifts often push these instruments to their detection limits, I chose to stack their spectra, obtaining a deep stacked spectrum of ~ 25 hours. By exploiting a variety of $f_{\rm esc}$(LyC) diagnostics (the peak separation, the asymmetry of the red peak, the ratio between Lyα and Hβ and the β slope), I conclude that the escape fraction in these three UV-bright $z \geq 7$ galaxies is $f_{\rm esc}$(LyC) $\leq 10\%$, well below the ~ 20% required in the oligarchical models of Naidu et al. (2020). This result is in agreement with more recent works

(Saxena et al., 2024; Mascia et al., 2023), suggesting that the answer to question (i) is that the abundance of UV-faint galaxies are driving the reionisation process, not UV-bright galaxies.

In the pre-*JWST* era, ground-based telescopes identified many high-redshift LAEs (e.g. Zitrin et al., 2015; Stark et al., 2017), however, what facilitates the escape of the Lyα emission from galaxies that are expected to be surrounded by significant neutral hydrogen densities remained poorly understood. The high detection rate of Lyα emission from strong [O III] + Hβ emitters identified in photometry (Roberts-Borsani et al., 2016) had been used to suggest LAEs have extreme ionizing properties (Roberts-Borsani et al., 2023b) and potentially host AGN (Laporte et al., 2017b). In Chapter 3, I investigate how the high-resolution and high-sensitivity imaging capabilities of *JWST*/NIRCam can help to unveil the escape mechanisms of Lyα emission in the early Universe. This reveals that all galaxies in a sample of $z > 7$ Lyα-emitters have close companion galaxies. When comparing LAE properties to on-the-fly radiative transfer magnetohydrodynamical simulations it becomes clear that interactions between galaxies drive bursts of star formation and hence intrinsic Lyα emission, but these interactions also act to create channels cleared of neutral hydrogen facilitating the escape of Lyα emission. This tendency for LAEs to be found with close companion galaxies has been further supported by Witstok et al. (2024c,a); Scholtz et al. (2024); Jung et al. (2024) and provides an answer to question (ii).

With the advent of *JWST* spectroscopy, the highest redshift LAE moved from $z = 8.68$ (Zitrin et al., 2015) to $z = 10.6$ (GN-z11; Bunker et al., 2023b) (and recently $z = 13$; Witstok et al., 2024b). NIRCam imaging of the environment of GN-z11 has revealed a large-scale photometric overdensity (Tacchella et al., 2023b) and NIRSpec IFU spectroscopy shows evidence of Population III stars in its halo (Maiolino et al., 2024b). In Chapter 4, I exploit these NIRSpec IFU observations to show the Lyα emission from GN-z11 is part of an extended halo with a minimum size of 0.8–3.2 kpc, as well as to detect two further emission line, UV-faint galaxy candidates within the IFU FOV. I show that if both the large-scale and local overdensities are confirmed with follow-up observations the environment around GN-z11 represents a protocluster core, similar to those previously identified with *JWST* (Laporte et al., 2022; Hashimoto et al., 2023; Morishita et al., 2023). The evidence that the majority of galaxies are not identified with close companion galaxies compared to the unanimity of LAEs in Chapter 3 and the identification of a candidate protocluster core in Chapter 4 shows the diversity of environments within which we find galaxies, answering question (iii).

Finally, the discovery of high-redshift galaxies hosting Balmer breaks and no nebular emission lines with NIRSpec (Looser et al., 2024; Vikaeus et al., 2024; Trussler et al., 2024a)

has further provoked question (iv), on whether galaxies undergo mini-quenched phases in their SFHs. While multiple simulations predict the existence of bursty SFHs in early galaxies (Fukuda et al., 2000; Immeli et al., 2004; Bournaud et al., 2007; Elmegreen et al., 2009; Faucher-Giguère, 2018; Tacchella et al., 2020), observing potentially mini-quenched galaxies does not give constraints on the frequency or timescales of burst events, offering only the insight that they are currently quenched. Instead, identifying galaxies that are rejuvenating after a recent mini-quenched period allows improved constraints on the burstyness of early galaxies. In Chapter 5 I identify a galaxy that hosts both a Balmer break and emission lines evidencing the presence of both an old and young stellar population and hence a recent rejuvenation event, offering the first spectroscopic confirmation of mini-quenching. SED-fitting of this galaxy allows us to constrain its recent mini-quenched period to be around 20 Myr. Moreover, by building a sample of high-redshift galaxies with similar ISM conditions, assessed by their peculiar emission line ratios, I identify a population of rejuvenating galaxies.

6.2. FUTURE OUTLOOK

While the first two years of observations with *JWST* has pushed back the redshift frontier to $z \sim 14$ (e.g. Curtis-Lake et al., 2023; D'Eugenio et al., 2024; Finkelstein et al., 2024; Arrabal Haro et al., 2023; Carniani et al., 2024; Castellano et al., 2024), it has revealed that some of these highest redshift galaxies have peculiar features that as yet are poorly understood (e.g. Cameron et al., 2024; Witstok et al., 2024b). Many of these unknowns continue to relate to the open questions I outlined in Chapter 1, e.g. the nature of UV-bright galaxies, the evolution of extremely dense environments and constraining the bursty SFHs of early galaxies. In the following section I discuss how current and future telescopes will expand our understanding of these topics, and how state-of-the-art simulations can provide crucial insights.

6.2.1 Wide surveys

Arguably, the most puzzling of the galaxies observed with *JWST* is GN-z11. Originally identified by the *Hubble* Space Telescope, it is one of the brightest galaxies seen at $z \gtrsim 10$. Observations of this remarkably luminous galaxy with *JWST*'s NIRCam and NIRSpec instruments have been used to claim that it hosts: an AGN (Maiolino et al., 2024a); extended Lyα emission (Bunker et al., 2023b; Scholtz et al., 2024); Pop. III stars in its halo (Maiolino et al., 2024b); peculiar stellar populations (Cameron et al., 2023a; Marques-Chaves et al.,

2024); and resides within both a small scale (Scholtz et al., 2024) and large scale (Tacchella et al., 2023b) overdensity. While GN-z11 is clearly one of the most interesting objects that has been observed at the highest-redshifts and placing stronger constraints on its properties is key in understanding its evolutionary pathway and the contribution of AGN and environment on its observed properties, the crucial question is whether similarly UV-bright, high-redshift galaxies show the same properties as GN-z11, or whether it is simply an outlier. The field therefore finds itself at a stage where large, relatively shallow depth surveys ($m_{\rm AB} \sim 26$) now hold the key to identifying these rare objects.

The Euclid[1] deep survey will reach 5σ depths of $m_{\rm AB} \sim 26.4$ AB magnitude, covering $\sim 53{\rm deg}^2$ (Euclid Collaboration et al., 2022). UV luminosity functions suggest that this should reveal of order 10 $z \sim 10$ galaxies with a comparable UV luminosity to GN-z11. It should be noted that the near-infrared spectrometer and photometer (NISP) offers only three wide-band imaging filters. At best, only two of which have sufficient wavelength coverage to detect emission from $z \sim 10$ galaxies, making photometric identification challenging, however, successful candidates were identified with *HST* with single filter detections (Bouwens et al., 2011; Ellis et al., 2013; Oesch et al., 2016), and as such, with careful selection, Euclid should be able to identify GN-z11 mimics. Utilising *JWST* to perform spectroscopic and deep-imaging follow-up of these galaxies and their environments, will allow us to understand if GN-z11 is representative of high-redshift UV-bright galaxies.

The GTO survey on the next generation VLT instrument MOONS will provide spectroscopy for a few thousand $z = 6-8$ galaxies (Maiolino et al., 2020). Given the $\sim 0.6-1.7\mu$m wavelength range of MOONS, and its ability to simultaneously obtain spectra for ~ 1000 objects within 500 square arcminutes, it is capable of revealing the relationship between Lyα emission and environment at high redshifts. Moreover, the high spectral resolution of MOONS will facilitate modelling of the Lyα emission line profile and hence potential mapping of ionised regions.

The second generation, multi-object and multi-integral field spectrograph, MOSAIC, at the extremely large telescope (ELT), will offer a similar science case to MOONS, but with significantly increased sensitivity. So much so that it can spectroscopically confirm the Lyα emission seen from GN-z11, that took *JWST* 10 hours to confirm (Bunker et al., 2023b), in just a few 10's of minutes. MOSAIC will therefore be a powerful instrument in confirming LAEs and in studying very overdense regions and how these evolve at high-redshift. Moreover, the IFU instruments will facilitate the resolved study of LAEs,

[1] Euclid is a space telescope aimed at mapping 1/3 of the sky with four wide-band filters across the visible and near-infrared wavelengths

confirming whether the Lyα halo seen around GN-z11 (discussed in Section 4) is frequently observed around AGN candidate galaxies.

6.2.2 *JWST*

The NIRCam WFSS is highly efficient at spectroscopically confirming strong emission line galaxies (e.g. Meyer et al., 2024), as such it is a crucial tool for constraining the environment of spectroscopically confirmed high-redshift galaxies, especially in the context of extremely overdense environments, such as protocluster cores. Protoclusters are believed to grow stellar mass in an inside-out manner between $z = 10 - 5$ (Chiang et al., 2017). Therefore, high-redshift protocluster cores should be evolved, highly star-forming environments. One such protocluster core candidate is within the SMACS 0723 lensing field at $z = 7.66$ (Laporte et al., 2022). The NIRCam imaging within this field does not include medium band filters, and as discussed in Chapter 1, this makes disentangling nebular and stellar continuum emission challenging, and thus SED-fitting is highly uncertain. Recent NIRCam WFSS observations with the F444W filter, covering the three strongest contaminating emission lines [O III] and Hβ, as part of the *JWST* Cycle 2 GO (ID: 4043, PI: Witten) will enable spectroscopic confirmation of the protocluster core and detailed spectro-photometric fitting of resident galaxies. Further identification of extremely overdense environments and resident galaxies' properties will allow for improved observational constraints on their formation and evolution.

Given protocluster cores at high redshifts are expected to host some of the most evolved stellar populations, they appear obvious locations to search for Balmer break galaxies. Indeed the spectroscopically confirmed protocluster core at $z = 7.88$ in the Abell 2744 lensing field (Hashimoto et al., 2023; Morishita et al., 2023) hosts the rejuvenating galaxy YD4, discussed in Chapter 5. Identifying more mini-quenched and rejuvenating galaxies in the EoR offers the potential to begin to constrain the duty cycle of high-redshift galaxies by confirming the quenched timescales of these galaxies. Future *JWST* proposals targeting mini-quenched galaxies (e.g. OASIS ID:5997 PI:Looser) as well as the ever increasing number of spectroscopically observed high-redshift galaxies, will inevitably identify larger samples of Balmer break galaxies in the near future.

6.2.3 Simulations

Large-volume cosmological simulations are beginning to include sufficiently advanced physical models which enable us to study the relationship between large scale structure (over several Mpcs) and galaxy properties, e.g. THESAN (Kannan et al., 2022; Smith et al.,

2022a; Garaldi et al., 2022) and SPHINX (Rosdahl et al., 2022; Katz et al., 2023a). Protocluster cores are expected to have radii of 10's of pkpc at high redshifts (Chiang et al., 2017), as such it is possible to identify some of the most overdense regions in simulations such as THESAN with a box length of 95.5 cMpc, and compare to the overdense structures identified with *JWST*. By relating galaxy counts, stellar mass and star-formation rates in observed and simulated protocluster cores one can utilise these simulations to constrain the dark matter halo masses, underlying faint galaxy populations and future evolutionary pathways of these overdense structures. Moreover, with the improved constraints on galaxy properties in protoclusters, such as their star-formation histories, neutral gas reservoirs and ISM properties (see Chapter 5 and Section 6.2.2) comparisons to simulations will enable a significant improvement in our understanding of the relationship between these properties and their larger-scale structures.

An additional crucial aspect of current state-of-the-art simulations are "zoom-in" cosmological simulations that resolve the ISM. The increased resolution of "zoom-in" simulations allows gas to collapse to much higher densities before it becomes unresolved, potentially driving burstier SFHs (Hu et al., 2023). Such burstier SFHs have been observed in "zoom-in" simulations (e.g. Sun et al., 2023b) and are believed to be related to high sSFR and relatively low gravitational potentials facilitating more efficient evacuation of gas at high redshifts (Hopkins et al., 2023). Existing "zoom-in" simulations, such as those presented in Chapter 3 and Bennett et al. (2024), and the upcoming THESAN-Zoom simulations, that includes a multi-phase ISM model with on-the-fly radiative transfer and a live dust model (Kannan et al. in prep.), may help to further reveal the properties of the bursty SFHs observed at high-redshift.

6.3. Final Remarks

The wealth of data provided by *JWST* has revolutionised the high-redshift field, moving from tens of spectroscopically confirmed galaxies at $z > 7$, to several thousand at the time of writing. High-redshift observations have moved from tentative single line detections, to continuum observations and the potential for robust, cutting-edge science has grown accordingly. We have already reached a stage where large samples of galaxies are revealing population properties at high-redshift, but ever larger samples are required. The wide-field imaging capabilities of *Euclid* and multi-object spectroscopy of MOONS will provide further answers to many of the questions put forward at the beginning of this thesis, regarding reionisation, Lyα escape and extremely overdense environments. The sensitivity and resolution of next generation telescopes and instruments such as ELT/MOSAIC will

inevitably place further constraints on these processes and structures. However, with the wealth of robust observational data of high-redshift galaxies, comparison to realistic simulations is crucial in understanding the implications of these observational results. A combination of state-of-the-art "zoom-in" simulations (e.g. *Azahar* and THESAN-Zoom) and larger-scale cosmological simulations (e.g. SPHINX and THESAN) are, as yet, underutilised tools by observational astronomers. Given all this, the field finds itself in an exceptional moment where observations and simulations are sufficient to radically alter our understanding of galaxies in the very early Universe, with the tantalising prospect of future simulations and next generation instruments around the corner.

References

Alpher R. A., Herman R. C., 1949, Physical Review, 75, 1089

Arrabal Haro P., et al., 2023, Nature, 622, 707

Arribas S., et al., 2024, Astronomy & Astrophysics, 688, A146

Asada Y., et al., 2023, MNRAS, 523, L40

Atek H., Richard J., Kneib J.-P., Schaerer D., 2018, MNRAS, 479, 5184

Austin D., et al., 2024, arXiv e-prints, p. arXiv:2404.10751

Bacon R., et al., 2010. San Diego, California, USA, p. 773508, doi:10.1117/12.856027, http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.856027

Bacon R., et al., 2021, Astronomy & Astrophysics, 647, A107

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

Baker J. G., Menzel D. H., 1938, ApJ, 88, 52

Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5

Barkana R., Loeb A., 2004, ApJ, 609, 474

Battaglia G., et al., 2005, MNRAS, 364, 433

Becker R. H., et al., 2001, AJ, 122, 2850

Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57

Bennett J. S., Sijacki D., Costa T., Laporte N., Witten C., 2024, MNRAS, 527, 1033

Benson A. J., Sugiyama N., Nusser A., Lacey C. G., 2006, MNRAS, 369, 1055

Bergamini P., et al., 2023, ApJ, 952, 84

Berman L., Trumpler R. J., 1936, Monthly Notices of the Royal Astronomical Society, 96, 890

Bertin E., Arnouts S., 1996, åps, 117, 393

Bezanson R., et al., 2022, arXiv e-prints, p. arXiv:2212.04026

Bian F., Kewley L. J., Dopita M. A., Juneau S., 2016, ApJ, 822, 62

Binggeli C., et al., 2019, MNRAS, 489, 3827

Binney J., 1977, ApJ, 215, 483

Blaizot J., et al., 2023, MNRAS, 523, 3749

Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517

Boera E., Becker G. D., Bolton J. S., Nasir F., 2019, ApJ, 872, 101

Bolan P., et al., 2022, MNRAS, 517, 3263

Bond J. R., Efstathiou G., Silk J., 1980, Phys. Rev. Lett., 45, 1980

Borthakur S., Heckman T. M., Leitherer C., Overzier R. A., 2014, Science, 346, 216

Bosman S. E. I., et al., 2022, MNRAS, 514, 55

Bournaud F., Elmegreen B. G., Elmegreen D. M., 2007, ApJ, 670, 237

Bouwens R. J., Illingworth G. D., Blakeslee J. P., Franx M., 2006, ApJ, 653, 53

Bouwens R. J., et al., 2010, ApJ, 709, L133

Bouwens R. J., et al., 2011, ApJ, 737, 90

Bouwens R. J., et al., 2014, ApJ, 793, 115

Bouwens R. J., et al., 2015, ApJ, 803, 34

Bouwens R. J., Smit R., Labbé I., Franx M., Caruana J., Oesch P., Stefanon M., Rasappu N., 2016, ApJ, 831, 176

Bouwens R. J., Illingworth G., Ellis R. S., Oesch P., Stefanon M., 2022, The Astrophysical Journal, 940, 55

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, Monthly Notices of the Royal Astronomical Society, 370, 645

Bowler R. A. A., et al., 2014, MNRAS, 440, 2810

Bowler R. A. A., Jarvis M. J., Dunlop J. S., McLure R. J., McLeod D. J., Adams N. J., Milvang-Jensen B., McCracken H. J., 2020, MNRAS, 493, 2059

Boyett K., et al., 2024, Nature Astronomy, 8, 657

Bradač M., et al., 2024, ApJ, 961, L21

Bradley L. D., et al., 2012, ApJ, 760, 108

Brammer G. B., et al., 2012, ApJS, 200, 13

Brinchmann J., Pettini M., Charlot S., 2008, MNRAS, 385, 769

Bromm V., 2013, Reports on Progress in Physics, 76, 112901

Bruton S., Lin Y.-H., Scarlata C., Hayes M. J., 2023, The Astrophysical Journal Letters, 949, L40

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

Bryan G. L., Norman M. L., 1998, ApJ, 495, 80

Bunker A. J., et al., 2023a, arXiv e-prints

Bunker A. J., et al., 2023b, Astronomy & Astrophysics, 677, A88

Byler N., Dalcanton J. J., Conroy C., Johnson B. D., 2017, ApJ, 840, 44

Böker T., et al., 2022, åp, 661, A82

Böker T., et al., 2023, PASP, 135, 038001

Calvi R., Dannerbauer H., Arrabal Haro P., Rodríguez Espinosa J. M., Muñoz-Tuñón C., Pérez González P. G., Geier S., 2021, MNRAS, 502, 4558

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682

Cameron A. J., Katz H., Rey M. P., Saxena A., 2023a, Monthly Notices of the Royal Astronomical Society, 523, 3516

Cameron A. J., et al., 2023b, åp, 677, A115

Cameron A. J., Katz H., Witten C., Saxena A., Laporte N., Bunker A. J., 2024, MNRAS

Cantalupo S., Porciani C., Lilly S. J., 2008, ApJ, 672, 48

Cantalupo S., Arrigoni-Battaia F., Prochaska J. X., Hennawi J. F., Madau P., 2014, Nature, 506, 63

Capak P. L., et al., 2011, Nature, 470, 233

Carnall A. C., 2017, arXiv e-prints, p. arXiv:1705.05165

Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, MNRAS, 480, 4379

Carniani S., et al., 2024, arXiv e-prints, p. arXiv:2405.18485

Casey C. M., 2016, ApJ, 824, 36

Castellano M., et al., 2010, åp, 524, A28

Castellano M., et al., 2016, ApJ, 818, L3

Castellano M., et al., 2022, Astronomy & Astrophysics, 662, A115

Castellano M., et al., 2024, The Astrophysical Journal, 972, 143

Chabrier G., 2003, PASP, 115, 763

Chabrier G., 2005, in Corbelli E., Palla F., Zinnecker H., eds, Astrophysics and Space Science Library Vol. 327, The Initial Mass Function 50 Years Later. p. 41, doi:10.1007/978-1-4020-3407-7_5

Champagne J. B., et al., 2021, ApJ, 913, 110

Chanchaiworawit K., et al., 2019, ApJ, 877, 51

Chandrasekhar S., 1931, ApJ, 74, 81

Charlot S., Fall S. M., 1993, ApJ, 415, 580

Chen Z., Stark D. P., Mason C., Topping M. W., Whitler L., Tang M., Endsley R., Charlot S., 2024, MNRAS, 528, 7052

Chiang Y.-K., Overzier R., Gebhardt K., 2013, ApJ, 779, 127

Chiang Y.-K., Overzier R. A., Gebhardt K., Henriques B., 2017, ApJ, 844, L23

Choustikov N., et al., 2024, MNRAS, 529, 3751

Conroy C., 2013, ARA&A, 51, 393

Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486

Costa T., Sijacki D., Trenti M., Haehnelt M. G., 2014, MNRAS, 439, 2146

Costa T., Arrigoni Battaia F., Farina E. P., Keating L. C., Rosdahl J., Kimm T., 2022, MNRAS, 517, 1767

Croton D. J., et al., 2006, MNRAS, 365, 11

Cullen F., et al., 2024, MNRAS, 531, 997

Curti M., Cresci G., Mannucci F., Marconi A., Maiolino R., Esposito S., 2017, Monthly Notices of the Royal Astronomical Society, 465, 1384

Curti M., Mannucci F., Cresci G., Maiolino R., 2020, Monthly Notices of the Royal Astronomical Society, 491, 944

Curti M., et al., 2022, Monthly Notices of the Royal Astronomical Society, 518, 425

Curtis-Lake E., et al., 2023, Nature Astronomy

D'Eugenio F., et al., 2024, arXiv e-prints, p. arXiv:2404.06531

Davies F. B., et al., 2018, The Astrophysical Journal, 864, 142

De Barros S., et al., 2017, åp, 608, A123

Dicke R. H., Peebles P. J. E., Roll P. G., Wilkinson D. T., 1965, ApJ, 142, 414

Dijkstra M., 2014, Publ. Astron. Soc. Australia, 31, e040

Dijkstra M., Loeb A., 2008, MNRAS, 391, 457

Dome T., Tacchella S., Fialkov A., Ceverino D., Dekel A., Ginzburg O., Lapiner S., Looser T. J., 2024, MNRAS, 527, 2139

Donnan C. T., et al., 2023, MNRAS, 518, 6011

Dressler A., et al., 2023, ApJ, 947, L27

Dressler A., et al., 2024, ApJ, 964, 150

Duan Q., et al., 2024, arXiv e-prints

Dubois Y., Commerçon B., 2016, åp, 585, A138

Dubois Y., Commerçon B., Marcowith A., Brahimi L., 2019, åp, 631, A121

Duval F., Schaerer D., Östlin G., Laursen P., 2014, Astronomy & Astrophysics, 562, A52

Einasto J., Saar E., Kaasik A., Chernin A. D., 1974, Nature, 252, 111

Eisenstein D. J., et al., 2023, arXiv e-prints, p. arXiv:2306.02465

Eldridge J. J., Stanway E. R., 2009, MNRAS, 400, 1019

Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, Publ. Astron. Soc. Australia, 34, e058

Ellis R. S., 2022, When Galaxies Were Born. The Quest for Cosmic Dawn

Ellis R. S., et al., 2013, ApJ, 763, L7

Elmegreen B. G., Lada C. J., 1977, ApJ, 214, 725

Elmegreen B. G., Elmegreen D. M., Fernandez M. X., Lemonias J. J., 2009, ApJ, 692, 12

Elson R. A. W., Fall S. M., Freeman K. C., 1989, ApJ, 336, 734

Endsley R., et al., 2022, MNRAS, 517, 5642

Endsley R., Stark D. P., Whitler L., Topping M. W., Chen Z., Plat A., Chisholm J., Charlot S., 2023, Monthly Notices of the Royal Astronomical Society, 524, 2312

Euclid Collaboration et al., 2022, åp, 662, A112

Faisst A. L., Morishita T., 2024, The Astrophysical Journal, 971, 47

Fall S. M., Efstathiou G., 1980, Monthly Notices of the Royal Astronomical Society, 193, 189

Fan X., Carilli C. L., Keating B., 2006, ARA&A, 44, 415

Farcy M., Rosdahl J., Dubois Y., Blaizot J., Martin-Alvarez S., 2022, MNRAS, 513, 5000

Farina E. P., et al., 2019, ApJ, 887, 196

Faucher-Giguère C.-A., 2018, MNRAS, 473, 3717

Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, PASP, 110, 761

Ferland G. J., et al., 2013, Rev. Mex. Astron. Astrofis., 49, 137

Finkelstein S. L., et al., 2013, Nature, 502, 524

Finkelstein S. L., et al., 2015, ApJ, 810, 71

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

Finkelstein S. L., et al., 2023, The Astrophysical Journal Letters, 946, L13

Finkelstein S. L., et al., 2024, The Astrophysical Journal Letters, 969, L2

Flury S. R., et al., 2022, ApJS, 260, 1

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Fosbury R. A. E., et al., 2003, The Astrophysical Journal, 596, 797

Frenk C. S., White S. D. M., Efstathiou G., Davis M., 1985, Nature, 317, 595

Fukuda H., Habe A., Wada K., 2000, ApJ, 529, 109

Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, ApJ, 613, 1

Furtak L. J., et al., 2024, Nature, 628, 57

Galli D., Palla F., 1998, åp, 335, 403

Gamow G., 1949, Reviews of Modern Physics, 21, 367

Garaldi E., Kannan R., Smith A., Springel V., Pakmor R., Vogelsberger M., Hernquist L., 2022, MNRAS, 512, 4909

Garel T., Blaizot J., Rosdahl J., Michel-Dansac L., Haehnelt M. G., Katz H., Kimm T., Verhamme A., 2021, MNRAS, 504, 1902

Gazagnes S., Chisholm J., Schaerer D., Verhamme A., Izotov Y., 2020, åp, 639, A85

Gelli V., Salvadori S., Ferrara A., Pallottini A., Carniani S., 2023, ApJ, 954, L11

Goerdt T., Dekel A., Sternberg A., Ceverino D., Teyssier R., Primack J. R., 2010, MNRAS, 407, 613

Gordon K. D., Clayton G. C., Misselt K. A., Landolt A. U., Wolff M. J., 2003, ApJ, 594, 279

Goto H., et al., 2021, The Astrophysical Journal, 923, 229

Governato F., Willman B., Mayer L., Brooks A., Stinson G., Valenzuela O., Wadsley J., Quinn T., 2007, Monthly Notices of the Royal Astronomical Society, 374, 1479

Governato F., et al., 2010, Nature, 463, 203

Gray W. J., Scannapieco E., 2017, ApJ, 849, 132

Greig B., Mesinger A., Bañados E., 2019, Monthly Notices of the Royal Astronomical Society, 484, 5094

Gunn J. E., Gott J. Richard I., 1972, ApJ, 176, 1

Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633

Guseva N. G., Izotov Y. I., Thuan T. X., 2006, The Astrophysical Journal, 644, 890

Guseva N. G., Izotov Y. I., Papaderos P., Fricke K. J., 2007, Astronomy & Astrophysics, 464, 885

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759

Haiman Z., Thoul A. A., Loeb A., 1996, ApJ, 464, 523

Hainline K. N., et al., 2024, The Astrophysical Journal, 964, 71

Hamilton D. R., 1940, Physical Review, 58, 122

Harikane Y., et al., 2019, ApJ, 883, 142

Harikane Y., Laporte N., Ellis R. S., Matsuoka Y., 2020, ApJ, 902, 117

Harikane Y., et al., 2022, The Astrophysical Journal, 929, 1

Harikane Y., Nakajima K., Ouchi M., Umeda H., Isobe Y., Ono Y., Xu Y., Zhang Y., 2024, ApJ, 960, 56

Harrington J. P., 1973, MNRAS, 162, 43

Hashimoto T., et al., 2018, Nature, 557, 392

Hashimoto T., et al., 2023, The Astrophysical Journal Letters, 955, L2

Hayes M. J., Scarlata C., 2023, The Astrophysical Journal Letters, 954, L14

Hayes M., Laporte N., Pelló R., Schaerer D., Le Borgne J.-F., 2012, MNRAS, 425, L19

Heintz K. E., et al., 2024a, arXiv e-prints

Heintz K. E., et al., 2024b, arXiv e-prints, p. arXiv:2404.02211

Heintz K. E., et al., 2024c, Science, 384, 890

Henden N. A., Puchwein E., Shen S., Sijacki D., 2018, MNRAS, 479, 5385

Hennawi J. F., Prochaska J. X., Cantalupo S., Arrigoni-Battaia F., 2015, Science, 348, 779

Henyey L. G., Greenstein J. L., 1941, ApJ, 93, 70

Hoag A., et al., 2019, The Astrophysical Journal, 878, 12

Hopkins P. F., et al., 2023, MNRAS, 525, 2241

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

Hu C.-Y., et al., 2023, ApJ, 950, 132

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Hubble E. P., 1936, Realm of the Nebulae

Hui L., Gnedin N. Y., 1997, MNRAS, 292, 27

Hunter D. A., Shaya E. J., Holtzman J. A., Light R. M., O'Neil Earl J. J., Lynds R., 1995, ApJ, 448, 179

Immeli A., Samland M., Westera P., Gerhard O., 2004, ApJ, 611, 20

Inoue A. K., Shimizu I., Iwata I., Tanaka M., 2014, MNRAS, 442, 1805

Inoue A. K., et al., 2018, Publications of the Astronomical Society of Japan, 70, 55

Ishigaki M., Kawamata R., Ouchi M., Oguri M., Shimasaku K., Ono Y., 2018, The Astrophysical Journal, 854, 73

Isobe Y., Ouchi M., Nakajima K., Harikane Y., Ono Y., Xu Y., Zhang Y., Umeda H., 2023, ApJ, 956, 139

Izotov Y. I., Schaerer D., Thuan T. X., Worseck G., Guseva N. G., Orlitová I., Verhamme A., 2016, MNRAS, 461, 3683

Izotov Y. I., Thuan T. X., Guseva N. G., 2017, MNRAS, 471, 548

Izotov Y. I., Schaerer D., Worseck G., Guseva N. G., Thuan T. X., Verhamme A., Orlitová I., Fricke K. J., 2018a, MNRAS, 474, 4514

Izotov Y. I., Worseck G., Schaerer D., Guseva N. G., Thuan T. X., Fricke Verhamme A., Orlitová I., 2018b, MNRAS, 478, 4851

Jakobsen P., et al., 2022, åp, 661, A80

Jeans J. H., 1902, Philosophical Transactions of the Royal Society of London Series A, 199, 1

Jin X., et al., 2023, The Astrophysical Journal, 942, 59

Johnson B. D., Leja J., Conroy C., Speagle J. S., 2021, ApJS, 254, 22

Jones G. C., et al., 2024, Astronomy & Astrophysics, 683, A238

Jung I., et al., 2020, The Astrophysical Journal, 904, 144

Jung I., et al., 2022, arXiv e-prints, p. arXiv:2212.09850

Jung I., et al., 2024, ApJ, 967, 73

Kakiichi K., Gronke M., 2021, ApJ, 908, 30

Kannan R., Garaldi E., Smith A., Pakmor R., Springel V., Vogelsberger M., Hernquist L., 2022, MNRAS, 511, 4005

Katz H., Laporte N., Ellis R. S., Devriendt J., Slyz A., 2019, MNRAS, 484, 4054

Katz H., et al., 2022, arXiv e-prints, p. arXiv:2211.04626

Katz H., et al., 2023a, The Open Journal of Astrophysics, 6, 44

Katz H., et al., 2023b, MNRAS, 518, 270

Katz H., et al., 2024, arXiv e-prints

Kaurov A. A., Gnedin N. Y., 2014, ApJ, 787, 146

Keating L. C., Weinberger L. H., Kulkarni G., Haehnelt M. G., Chardin J., Aubert D., 2020, MNRAS, 491, 1736

Keating L. C., Bolton J. S., Cullen F., Haehnelt M. G., Puchwein E., Kulkarni G., 2024, MNRAS, 532, 1646

Kennicutt Robert C. J., 1998, ARA&A, 36, 189

Kerutt J., et al., 2022, åp, 659, A183

Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, ARA&A, 57, 511

Kimm T., Cen R., 2014, ApJ, 788, 121

Kimm T., Katz H., Haehnelt M., Rosdahl J., Devriendt J., Slyz A., 2017, MNRAS, 466, 4826

Klessen R. S., Glover S. C. O., 2023, ARA&A, 61, 65

Kocevski D. D., et al., 2024, arXiv e-prints, p. arXiv:2404.03576

Konno A., et al., 2014, The Astrophysical Journal, 797, 16

Konno A., et al., 2018, Publications of the Astronomical Society of Japan, 70, S16

Kravtsov A., Belokurov V., 2024, arXiv e-prints

Kriek M., et al., 2006, The Astrophysical Journal, 645, 44

Kroupa P., 2001, MNRAS, 322, 231

Kuiper G. P., 1938, ApJ, 88, 472

Kulkarni G., Keating L. C., Haehnelt M. G., Bosman S. E. I., Puchwein E., Chardin J., Aubert D., 2019, MNRAS, 485, L24

Labbé I., et al., 2023, Nature, 616, 266

Laporte N., et al., 2017a, ApJ, 837, L21

Laporte N., Nakajima K., Ellis R. S., Zitrin A., Stark D. P., Mainali R., Roberts-Borsani G. W., 2017b, ApJ, 851, 40

Laporte N., et al., 2019, MNRAS, 487, L81

Laporte N., Meyer R. A., Ellis R. S., Robertson B. E., Chisholm J., Roberts-Borsani G. W., 2021, MNRAS, 505, 3336

Laporte N., Zitrin A., Dole H., Roberts-Borsani G., Furtak L. J., Witten C., 2022, åp, 667, L3

Laporte N., Ellis R. S., Witten C. E. C., Roberts-Borsani G., 2023, MNRAS, 523, 3018

Larson R. B., 1974, Monthly Notices of the Royal Astronomical Society, 169, 229

Larson R. B., 1985, MNRAS, 214, 379

Larson R. B., 1998, MNRAS, 301, 569

Larson R. L., et al., 2023, The Astrophysical Journal Letters, 953, L29

Laursen P., Sommer-Larsen J., Andersen A. C., 2009, ApJ, 704, 1640

Laursen P., Sommer-Larsen J., Razoumov A. O., 2011, ApJ, 728, 52

Leclercq F., et al., 2017, åp, 608, A8

Leitet E., Bergvall N., Hayes M., Linné S., Zackrisson E., 2013, åp, 553, A106

Leja J., Carnall A. C., Johnson B. D., Conroy C., Speagle J. S., 2019, ApJ, 876, 3

Lemaître G., 1927, Annales de la Société Scientifique de Bruxelles, 47, 49

Leonova E., et al., 2022, MNRAS, 515, 5790

Li A., Draine B. T., 2001, ApJ, 554, 778

Li Q., Narayanan D., Davé R., 2019, MNRAS, 490, 1425

Li Q., et al., 2024, arXiv e-prints

Lifshitz E. M., 1946, Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki, 16, 587

Liu X., Shapley A. E., Coil A. L., Brinchmann J., Ma C.-P., 2008, ApJ, 678, 758

Long A. S., et al., 2020, ApJ, 898, 133

Looser T. J., et al., 2023, arXiv e-prints, p. arXiv:2306.02470

Looser T. J., et al., 2024, Nature, 629, 53

Lotz J. M., et al., 2017, ApJ, 837, 97

Lynden-Bell D., 1967, MNRAS, 136, 101

Madau P., 1995, ApJ, 441, 18

Madau P., Haardt F., Rees M. J., 1999, ApJ, 514, 648

Mainali R., et al., 2018, MNRAS, 479, 1180

Maiolino R., Mannucci F., 2019, The Astronomy and Astrophysics Review, Volume 27, Issue 1, article id. 3, >187</NUMPAGES>pp., 27, 3

Maiolino R., et al., 2020, The Messenger, 180, 24

Maiolino R., et al., 2024a, Nature, 627, 59

Maiolino R., et al., 2024b, Astronomy & Astrophysics, 687, A67

Maji M., et al., 2022, Astronomy & Astrophysics, 663, A66

Man A., Belli S., 2018, Nature Astronomy, 2, 695

Marques-Chaves R., et al., 2024, åp, 681, A30

Martin-Alvarez S., Slyz A., Devriendt J., Gómez-Guijarro C., 2020, MNRAS, 495, 4475

Martin-Alvarez S., Katz H., Sijacki D., Devriendt J., Slyz A., 2021, MNRAS, 504, 2517

Martin-Alvarez S., Sijacki D., Haehnelt M. G., Farcy M., Dubois Y., Belokurov V., Rosdahl J., Lopez-Rodriguez E., 2023, Monthly Notices of the Royal Astronomical Society, 525, 3806

Mascia S., et al., 2023, åp, 672, A155

Mason C. A., Gronke M., 2020, MNRAS, 499, 1395

Mason C. A., et al., 2019, Monthly Notices of the Royal Astronomical Society, 485, 3947

Mason C. A., Trenti M., Treu T., 2023, MNRAS, 521, 497

Masters D., Faisst A., Capak P., 2016, ApJ, 828, 18

Mateo M., 1988, ApJ, 331, 261

Matthee J., Sobral D., Gronke M., Paulino-Afonso A., Stefanon M., Röttgering H., 2018, åp, 619, A136

Matthee J., et al., 2024, ApJ, 963, 129

McClymont W., et al., 2024a, arXiv e-prints, p. arXiv:2403.03243

McClymont W., et al., 2024b, arXiv e-prints, p. arXiv:2405.15859

McConachie I., et al., 2022, ApJ, 926, 37

McGaugh S. S., 1991, The Astrophysical Journal, 380, 140

McGreer I. D., Mesinger A., D'Odorico V., 2015, MNRAS, 447, 499

McKee C. F., Ostriker E. C., 2007, Annual Review of Astronomy and Astrophysics, 45, 565

McLeod D. J., McLure R. J., Dunlop J. S., 2016, Monthly Notices of the Royal Astronomical Society, 459, 3812

McLure R. J., et al., 2013, MNRAS, 432, 2696

Menzel D. H., Baker J. G., 1937, ApJ, 86, 70

Mesinger A., Aykutalp A., Vanzella E., Pentericci L., Ferrara A., Dijkstra M., 2015, Monthly Notices of the Royal Astronomical Society, 446, 566

Mestel L., 1952, Monthly Notices of the Royal Astronomical Society, 112, 598

Meurer G. R., Heckman T. M., Calzetti D., 1999, ApJ, 521, 64

Meyer R. A., Laporte N., Ellis R. S., Verhamme A., Garel T., 2021, MNRAS, 500, 558

Meyer R. A., et al., 2024, arXiv e-prints, p. arXiv:2405.05111

Michel-Dansac L., Blaizot J., Garel T., Verhamme A., Kimm T., Trebitsch M., 2020, åp, 635, A154

Miralda-Escudé J., 1998, ApJ, 501, 15

Mirocha J., Furlanetto S. R., 2023, MNRAS, 519, 843

Mitchell P. D., Blaizot J., Cadiou C., Dubois Y., Garel T., Rosdahl J., 2021, MNRAS, 501, 5757

Mo H. J., Mao S., White S. D. M., 1998, Monthly Notices of the Royal Astronomical Society, 295, 319

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution. Cambridge University Press

Morales A. M., Mason C. A., Bruton S., Gronke M., Haardt F., Scarlata C., 2021, The Astrophysical Journal, 919, 120

Morishita T., Abdurro'uf Hirashita H., Newman A. B., Stiavelli M., Chiaberge M., 2022, The Astrophysical Journal, 938, 144

Morishita T., et al., 2023, ApJ, 947, L24

Morishita T., et al., 2024, ApJ, 963, 9

Mortlock D. J., et al., 2011, Nature, 474, 616

Murray N., Quataert E., Thompson T. A., 2005, ApJ, 618, 569

Naidu R. P., Tacchella S., Mason C. A., Bose S., Oesch P. A., Conroy C., 2020, ApJ, 892, 109

Nakajima K., Maiolino R., 2022, MNRAS, 513, 5134

Nakajima K., Ouchi M., Isobe Y., Harikane Y., Zhang Y., Ono Y., Umeda H., Oguri M., 2023, ApJS, 269, 33

Navarro J. F., Eke V. R., Frenk C. S., 1996, MNRAS, 283, L72

Neufeld D. A., 1990, ApJ, 350, 216

Neyer M., et al., 2024, MNRAS, 531, 2943

Ocvirk P., et al., 2020, MNRAS, 496, 4087

Oesch P. A., et al., 2015, ApJ, 804, L30

Oesch P. A., et al., 2016, The Astrophysical Journal, 819, 129

Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, The Astrophysical Journal, 855, 105

Oesch P. A., et al., 2023, Monthly Notices of the Royal Astronomical Society, 525, 2864

Offner S. S. R., Clark P. C., Hennebelle P., Bastian N., Bate M. R., Hopkins P. F., Moraux E., Whitworth A. P., 2014, in , Protostars and Planets VI. University of Arizona Press, doi:10.2458/azu_uapress_9780816531240-ch003, http://muse.jhu.edu/books/9780816598762/9780816598762-9.pdf

Oke J. B., Gunn J. E., 1983, ApJ, 266, 713

Oppenheimer J. R., Volkoff G. M., 1939, Physical Review, 55, 374

Ormerod K., et al., 2023, Monthly Notices of the Royal Astronomical Society, 527, 6110

Osterbrock D. E., Ferland G. J., 2006, Astrophysics of gaseous nebulae and active galactic nuclei

Ostriker J. P., Peebles P. J. E., Yahil A., 1974, ApJ, 193, L1

Ouchi M., et al., 2009, ApJ, 706, 1136

Ouchi M., et al., 2010, The Astrophysical Journal, 723, 869

Ouchi M., et al., 2018, Publications of the Astronomical Society of Japan, 70, S13

Ouchi M., Ono Y., Shibuya T., 2020, Annual Review of Astronomy and Astrophysics, 58, 617

Papovich C., Dickinson M., Ferguson H. C., 2001, ApJ, 559, 620

Papovich C., et al., 2022, ApJ, 937, 22

Partridge R. B., Peebles P. J. E., 1967, ApJ, 147, 868

Peebles P. J., 1966, Phys. Rev. Lett., 16, 410

Peebles P. J. E., 1968, ApJ, 153, 1

Peebles P. J. E., 1969, ApJ, 155, 393

Peebles P. J. E., 1982, ApJ, 263, L1

Peebles P. J. E., Dicke R. H., 1968, ApJ, 154, 891

Peimbert M., Costero R., 1969, Boletin de los Observatorios Tonantzintla y Tacubaya, 5, 3

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266

Pentericci L., et al., 2011, ApJ, 743, 132

Pentericci L., et al., 2014, ApJ, 793, 113

Pentericci L., et al., 2016, ApJ, 829, L11

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Perlmutter S., et al., 1999, ApJ, 517, 565

Perna M., et al., 2023, arXiv e-prints, p. arXiv:2304.06756

Pettini M., Pagel B. E. J., 2004, Monthly Notices of the Royal Astronomical Society, 348, L59

Pforr J., Maraston C., Tonini C., 2012, MNRAS, 422, 3285

Planck M., 1901, Annalen der Physik, 309, 553

Planck Collaboration et al. 2016, åp, 594, A13

Planck Collaboration et al., 2020, åp, 641, A6

Polletta M., et al., 2021, åp, 654, A121

Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14

Press W. H., Schechter P., 1974, ApJ, 187, 425

Raiter A., Schaerer D., Fosbury R. A. E., 2010, Astronomy & Astrophysics, 523, A64

Rauscher B. J., et al., 2012, in Holland A. D., Beletic J. W., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 8453, High Energy, Optical, and Infrared Detectors for Astronomy V. p. 84531F, doi:10.1117/12.926089

Reddy N. A., et al., 2023a, ApJ, 951, 56

Reddy N. A., Topping M. W., Sanders R. L., Shapley A. E., Brammer G., 2023b, The Astrophysical Journal, 952, 167

Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541

Reines A. E., Nidever D. L., Whelan D. G., Johnson K. E., 2010, The Astrophysical Journal, 708, 26

Rezzolla L., Most E. R., Weih L. R., 2018, The Astrophysical Journal Letters, 852, L25

Rieke M. J., et al., 2023a, PASP, 135, 028001

Rieke M. J., et al., 2023b, The Astrophysical Journal Supplement Series, 269, 16

Riess A. G., et al., 1998, AJ, 116, 1009

Rigby J., et al., 2023, PASP, 135, 048001

Roberts M. S., Rots A. H., 1973, åp, 26, 483

Roberts-Borsani G. W., et al., 2016, ApJ, 823, 143

Roberts-Borsani G. W., Ellis R. S., Laporte N., 2020, MNRAS, 497, 3440

Roberts-Borsani G., et al., 2022, The Astrophysical Journal Letters, 938, L13

Roberts-Borsani G., et al., 2023a, Nature, 618, 480

Roberts-Borsani G., et al., 2023b, The Astrophysical Journal, 948, 54

Roberts-Borsani G., et al., 2024, arXiv e-prints, p. arXiv:2403.07103

Robertson B. E., 2022, ARA&A, 60, 121

Robertson B. E., et al., 2013, ApJ, 768, 71

Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, ApJ, 802, L19

Rosdahl J., Teyssier R., 2015, MNRAS, 449, 4380

Rosdahl J., et al., 2018, MNRAS, 479, 994

Rosdahl J., et al., 2022, MNRAS, 515, 2386

Rosen A., Bregman J. N., 1995, ApJ, 440, 634

Rubin V. C., Ford W. K. J., Thonnard N., 1978, ApJ, 225, L107

Rubin V. C., Ford W. K. J., Thonnard N., 1980, ApJ, 238, 471

Rybicki G. B., Lightman A. P., 1979, Radiative processes in astrophysics. Wiley

Salim S., Narayanan D., 2020, Annual Review of Astronomy and Astrophysics, 58, 529

Salpeter E. E., 1955, ApJ, 121, 161

Sanders R. L., et al., 2016, ApJ, 816, 23

Sanders R. L., et al., 2021, The Astrophysical Journal, 914, 19

Sanders R. L., Shapley A. E., Topping M. W., Reddy N. A., Brammer G. B., 2023, ApJ, 955, 54

Saxena A., et al., 2023, Astronomy & Astrophysics, 678, A68

Saxena A., et al., 2024, Astronomy & Astrophysics, 684, A84

Schaerer D., 2002, Astronomy & Astrophysics, 382, 28

Schaerer D., de Barros S., 2009, åp, 502, 423

Schaller G., Schaerer D., Meynet G., Maeder A., 1992, åps, 96, 269

Schechter P., 1976, ApJ, 203, 297

Schenker M. A., Stark D. P., Ellis R. S., Robertson B. E., Dunlop J. S., McLure R. J., Kneib J.-P., Richard J., 2012, ApJ, 744, 179

Schenker M. A., Ellis R. S., Konidaris N. P., Stark D. P., 2014, ApJ, 795, 20

Schmidt B. P., et al., 1998, ApJ, 507, 46

Schneider R., Salvaterra R., Ferrara A., Ciardi B., 2006, MNRAS, 369, 825

Scholtz J., et al., 2023, arXiv e-prints

Scholtz J., et al., 2024, Astronomy & Astrophysics, 687, A283

Schroeder J., Mesinger A., Haiman Z., 2013, Monthly Notices of the Royal Astronomical Society, 428, 3058

Sharma M., Theuns T., Frenk C., Bower R. G., Crain R. A., Schaller M., Schaye J., 2017, MNRAS, 468, 2176

Shen X., Vogelsberger M., Boylan-Kolchin M., Tacchella S., Kannan R., 2023, MNRAS, 525, 3254

Sijacki D., Springel V., Haehnelt M. G., 2009, MNRAS, 400, 100

Silk J., 1977, The Astrophysical Journal, 211, 638

Simmonds C., et al., 2024a, arXiv e-prints

Simmonds C., et al., 2024b, MNRAS, 527, 6139

Skelton R. E., et al., 2014, ApJS, 214, 24

Smith B., Sigurdsson S., Abel T., 2008, MNRAS, 385, 1443

Smith A., Safranek-Shrader C., Bromm V., Milosavljević M., 2015, MNRAS, 449, 4336

Smith A., Kannan R., Garaldi E., Vogelsberger M., Pakmor R., Springel V., Hernquist L., 2022a, MNRAS, 512, 3243

Smith A., et al., 2022b, Monthly Notices of the Royal Astronomical Society, 517, 1

Song M., Finkelstein S. L., Livermore R. C., Capak P. L., Dickinson M., Fontana A., 2016, ApJ, 826, 113

Springel V., Di Matteo T., Hernquist L., 2005, The Astrophysical Journal, 620, L79

Stanway E. R., Eldridge J. J., 2018, MNRAS, 479, 75

Stark D. P., Ellis R. S., Ouchi M., 2011, ApJ, 728, L2

Stark D. P., et al., 2015, MNRAS, 454, 1393

Stark D. P., et al., 2017, MNRAS, 464, 469

Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, ApJ, 462, L17

Steinmetz M., Navarro J. F., 1999, ApJ, 513, 555

Stiavelli M., et al., 2023, ApJ, 957, L18

Strait V., et al., 2023, ApJ, 949, L23

Sun F., et al., 2023a, The Astrophysical Journal, 953, 53

Sun G., Faucher-Giguère C.-A., Hayward C. C., Shen X., Wetzel A., Cochrane R. K., 2023b, ApJ, 955, L35

Tacchella S., Forbes J. C., Caplar N., 2020, MNRAS, 497, 698

Tacchella S., et al., 2022, ApJ, 927, 170

Tacchella S., et al., 2023a, MNRAS, 522, 6236

Tacchella S., et al., 2023b, ApJ, 952, 74

Tacchella S., et al., 2024, arXiv e-prints

Tan J. C., McKee C. F., 2004, ApJ, 603, 383

Tang M., et al., 2023, MNRAS, 526, 1657

Teyssier R., 2002, åp, 385, 337

Teyssier R., Fromang S., Dormy E., 2006, Journal of Computational Physics, 218, 44

Thies I., Kroupa P., 2007, ApJ, 671, 767

Thomson J. J. J. J., 1906, Conduction of electricity through gases / by J.J. Thomson., 2nd ed. edn. Cambridge University Press, Cambridge

Tilvi V., et al., 2020, ApJ, 891, L10

Tolman R. C., 1939, Physical Review, 55, 364

Toomre A., 1977, Mergers and Some Consequences. https://ui.adsabs.harvard.edu/abs/1977egsp.conf..401T

Toshikawa J., et al., 2018, PASJ, 70, S12

Totani T., Kawai N., Kosugi G., Aoki K., Yamada T., Iye M., Ohta K., Hattori T., 2006, Publications of the Astronomical Society of Japan, 58, 485

Totani T., et al., 2014, Publications of the Astronomical Society of Japan, 66, 63

Trebitsch M., Hutter A., Dayal P., Gottlöber S., Legrand L., Yepes G., 2022, Monthly Notices of the Royal Astronomical Society, 518, 3576

Treu T., et al., 2022, ApJ, 935, 110

Trinca A., Schneider R., Valiante R., Graziani L., Ferrotti A., Omukai K., Chon S., 2024, MNRAS, 529, 3563

Trussler J. A. A., et al., 2023, MNRAS, 525, 5328

Trussler J., et al., 2024a, arXiv e-prints

Trussler J. A. A., et al., 2024b, MNRAS, 527, 11627

Tweed D., Devriendt J., Blaizot J., Colombi S., Slyz A., 2009, åp, 506, 647

Umeda H., Ouchi M., Nakajima K., Harikane Y., Ono Y., Xu Y., Isobe Y., Zhang Y., 2023, arXiv e-prints, p. arXiv:2306.00487

Vale A., Ostriker J. P., 2004, MNRAS, 353, 189

Van Den Bosch F. C., Yang X., Mo H. J., 2003, Monthly Notices of the Royal Astronomical Society, 340, 771

Vanzella E., et al., 2008, åp, 478, 83

Vanzella E., et al., 2023, Astronomy & Astrophysics, 678, A173

Veilleux S., Osterbrock D. E., 1987, ApJS, 63, 295

Venemans B. P., et al., 2005, åp, 431, 793

Venturi G., et al., 2024, arXiv e-prints, p. arXiv:2403.03977

Verhamme A., Schaerer D., Maselli A., 2006, Astronomy & Astrophysics, 460, 397

Verhamme A., Orlitová I., Schaerer D., Hayes M., 2015, åp, 578, A7

Vikaeus A., et al., 2024, MNRAS, 529, 1299

Vito F., Di Mascia F., Gallerani S., Zana T., Ferrara A., Carniani S., Gilli R., 2022, MNRAS, 514, 1672

Wagoner R. V., Fowler W. A., Hoyle F., 1967, ApJ, 148, 3

Walborn N. R., Maíz-Apellániz J., Barbá R. H., 2002, AJ, 124, 1601

Wang T., et al., 2016, ApJ, 828, 56

Wang F., et al., 2020, The Astrophysical Journal, 896, 23

Wang X., et al., 2024a, The Astrophysical Journal Letters, 967, L42

Wang B., et al., 2024b, The Astrophysical Journal Letters, 969, L13

Weaver J. R., et al., 2024, ApJS, 270, 7

Weinberger L. H., Haehnelt M. G., Kulkarni G., 2019, MNRAS, 485, 1350

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Whitler L. R., Mason C. A., Ren K., Dijkstra M., Mesinger A., Pentericci L., Trenti M., Treu T., 2020, Monthly Notices of the Royal Astronomical Society, 495, 3602

Whitler L., Stark D. P., Endsley R., Leja J., Charlot S., Chevallard J., 2023, MNRAS, 519, 5859

Wien W., 1896, Annalen der Physik, 294, 662

Wilkins S. M., Stanway E. R., Bremer M. N., 2014, MNRAS, 439, 1038

Wilkins S. M., et al., 2024, MNRAS, 527, 7965

Williams H., et al., 2023, Science, 380, 416

Wise J. H., Abel T., 2008, ApJ, 685, 40

Wisotzki L., et al., 2018, Nature, 562, 229

Witstok J., et al., 2024a, arXiv e-prints

Witstok J., et al., 2024b, arXiv e-prints

Witstok J., et al., 2024c, Astronomy & Astrophysics, 682, A40

Witten C. E. C., et al., 2022, MNRAS, 516, 3254

Witten C. E. C., Laporte N., Katz H., 2023, ApJ, 944, 61

Witten C., et al., 2024a, arXiv e-prints

Witten C., et al., 2024b, Nature Astronomy, 8, 384

Wood K., Reynolds R. J., 1999, ApJ, 525, 799

Xiao L., Stanway E. R., Eldridge J. J., 2018, MNRAS, 477, 904

Yan H., Wang L., Ma Z., Hu L., 2023, ApJ, 947, L1

Yee H. K. C., De Robertis M. M., 1991, ApJ, 381, 386

Yusef-Zadeh F., Morris M., White R. L., 1984, ApJ, 278, 186

Zackrisson E., Rydberg C.-E., Schaerer D., Östlin G., Tuli M., 2011, ApJ, 740, 13

Zackrisson E., Inoue A. K., Jensen H., 2013, ApJ, 777, 39

Zackrisson E., et al., 2017, ApJ, 836, 78

Zheng Z., Coil A. L., Zehavi I., 2007, The Astrophysical Journal, 667, 760

Zheng W., et al., 2014, ApJ, 795, 93

Zhu Y., et al., 2022, The Astrophysical Journal, 932, 76

Zitrin A., et al., 2015, ApJ, 810, L12

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

de Marchi G., Paresce F., 2001. p. MS 05 51, https://ui.adsabs.harvard.edu/abs/2001AGM...
.18S0551D

Übler H., et al., 2024, Monthly Notices of the Royal Astronomical Society, 531, 355