



THE UNIVERSITY *of* EDINBURGH  
**School of Physics  
and Astronomy**

MPhys Project Report

# Hunting for Emission lines in High Redshift Galaxies with JWST

Blythe Fernandes

Supervisor: Prof. Ross McLure, Dr. Derek McLeod

*Submitted for the 40pt MPhys Project course PHYS11016*

25 March 2024

## Abstract

The ultra-deep imaging provided by the recently launched JWST allows for a deeper analysis to improve our current understanding for the epoch of Cosmic Reionisation; by studying the properties of galaxies that were responsible for it. Using data from select filters of the JADES survey, we are able to identify galaxies with strong H $\alpha$  and O [III] emission at  $z \approx 5$  and 7, respectively. This analysis is additionally supported with data from HST. By studying the strength of these emission lines it is possible to estimate the ionising photon production efficiency and the ionised photon escape fraction for a galaxy population. The overall aim of the analysis is to determine whether or not the galaxy sample at high redshift was capable of producing enough ionising photons to explain cosmic reionisation.

[**Keywords:** *High Redshift Galaxies, H-Alpha, Oxygen-III, Photon Escape Fraction, Photon Ion Efficiency*]

## Personal Statement

This project report seeks to analyse particular properties using data from the recently launched JWST which probes the period of time when Cosmic Reionisation began. I started working on this project in the month of September and for the first 2 weeks, I read various papers to understand the importance for looking at galaxies in this period. Thereafter, I had weekly meetings with my supervisor throughout the academic year; to monitor progress, receive any feedback on data analysis conducted and, to ask any questions regarding the tasks. We also discussed ideal strategies and steps that should be taken for data analysis.

Additional meetings with my 2<sup>nd</sup> supervisor and a member of staff, that work with high redshift galaxies, were also arranged during the year to get further advise on various elements on the project. Some of these were conducted to maintain the weekly schedule for meetings where my main supervisor was not available.

I documented all of my findings in a notebook throughout the project to keep track of the different errors and methods that worked and equally didn't work. Minutes for each of the meetings were also recorded for future reference.

A major challenge I experienced during the project was a complication with the computational memory. The JADES survey happened to be significantly large and therefore, required higher RAM space to run analysis using Python.

During the first semester, the focus was to calibrate the dataset and extract objects to a catalogue. During the winter break, I refined the calibration process for a more optimised and efficient use of system resources. I also worked on the code further, using modules ‘pytorch’ (PyTorch Developers n.d.) and ‘multiprocessing’ (Python Software Foundation n.d.) to work around the computational problem. During the second semester, the next step in the project was to process the catalogue for redshifts. This was one of the main steps in the project to which I dedicated about 4 weeks of work.

Lastly, I spent about 4 to 6 weeks using the redshifts estimates with linear relationships that had been established in previous literature work. The results allowed me to make deductions regarding the galactic sample and the results were compared to literature findings to see how well they resembled each other.

Another challenge that was prevalent during the middle of the second semester were complications with the operating system on my personal computer, which I had been using for this project. During this time, the laptop was sent back and forth between myself and the repair depot. A couple of weeks were dedicated towards this issue, however this is still ongoing. Due to this reason the the plan for the project was pushed slightly behind; however, progress on the project continued. I moved my work to the Linux Machines and I got a laptop loan from the library to continue working on my project.

The information on the filters used from JADES makes use of the Spanish Virtual Observatory at <https://svo.cab.inta-csic.es>.

## Acknowledgments

Firstly, I would like extend my sincere gratitude to my supervisor, Prof. Ross McLure, for all the productive and insightful meetings throughout this project. His explanations on the various concepts were very clear and I have been personally inspired by his work ethic. I am very grateful to Prof. R. McLure for taking me under his wing and mentoring me in the science of galaxies. I would also like to thank him for being incredibly understanding and patient with the complicated circumstances that were encountered regarding my personal computer. I have learnt a great deal from Prof. R. McLure and I will forever carry the experiences that I have gained through our interactions onto my future endeavours.

I would also like to thank Dr. Derek McLeod for aligning all the images that were used in this project. He works closely with configuring and analysing large datasets from JWST, and worked on this before I began the project so that I would be able to make an immediate start once I came along. I would also like to thank him for making time to meet with me to answer any questions I had on the project.

Having worked extensively on EAzY, I would also like to thank Dr. Ryan Begley for his time with helping me set it up and with understanding how it works. Dr. R. Begley explained and provided me with configuration parameters for the software that were suitable for my project. I made use of this software to run my catalogue for photometric redshift estimates. Dr. R. Begley also provided me with specific python methods and tools that were useful for proper analysis in my project.

At the beginning of the project, I also received help from Dr. David McKain with installing GAIA on my personal laptop. I greatly appreciate Dr. D. McKain's support and patience with installing a software that is not commonly done on a Windows OS.

I would also like to thank the developers of the main python modules that I used in this project. These are namely numpy, astropy, matplotlib, pandas, sep and eazy-py.

Last but not least, a heartfelt thank you to my family who have always been there for me and kept me motivated. They help me remind myself why I decided to study galaxies and, always encourage me to reach for the stars.

# Contents

<b>Personal Statement</b>	<b>ii</b>
<b>Acknowledgments</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Background Theory</b>	<b>2</b>
2.1 Epoch of Cosmic Reionisation . . . . .	2
2.2 Emission Lines . . . . .	4
2.2.1 Hydrogen Alpha ( $H\alpha$ ) . . . . .	4
2.2.2 Doubly Ionised Oxygen ( $O[III]$ ) . . . . .	4
2.3 JWST and JADES Survey . . . . .	5
<b>3 Inspecting the JADES Dataset</b>	<b>8</b>
3.1 Object Features Observed . . . . .	10
<b>4 Calibrating the Data</b>	<b>10</b>
4.1 PSF Homogenisation . . . . .	10
4.2 Constructing Convolution Kernels . . . . .	13
4.3 Mapping Background Noise . . . . .	15
<b>5 Building a Catalogue</b>	<b>16</b>
5.1 Deriving the Error on the Flux . . . . .	16
5.2 Gathering Flux Data . . . . .	18
<b>6 Results and Discussion</b>	<b>19</b>
6.1 Determining Redshifts . . . . .	19
6.2 Equivalent Widths of the Emission Lines . . . . .	22

6.2.1	Ionising Photon Production Efficiency . . . . .	25
6.3	UV Spectral Slope . . . . .	26
6.3.1	Photon Escape Fraction . . . . .	29
6.4	Understanding Photon Ionisation . . . . .	30
<b>7</b>	<b>Conclusion</b>	<b>31</b>
7.1	Summary . . . . .	31
7.2	Areas for Improvement . . . . .	32
<b>References</b>		<b>33</b>

# 1 Introduction

This report will go through the processes and methods that were taken to achieve photometric redshifts for a sample of galaxies at  $z \approx 5$  and  $7$ . These redshifts are then used to calculate other properties for the galaxy population with known relationships from previous work.

We preface the content in this report, in section 2, by giving the motivation behind the research. Getting a better understanding of the universe during the epoch of cosmic reionisation helps us understand how our current universe came to be. The James Webb Space Telescope (JWST) is now able to see further back in time because of its ability to observe at longer wavelengths. It is within this period in time where we are able to probe the faint-end of the galaxy luminosity function.

This report will describe aspects of the JADES dataset in section 2.3 and 3 that are used for the analysis and conclusions of this project. This dataset is supported with the Hubble Space Telescope, Ultra Deep Field (HST UDF) in addition to JADES. From typical Spectral Energy Distributions (SEDs) documented in previous work such as Förster Schreiber et al. 2020, the analysis is focused on 2 different strong emissions. These emission lines can tell us about useful properties such as the star-formation rate, the ionised photon production efficiency and the photon escape fraction of a galaxy population.

In section 4, we delve into the processes that taken to create a full catalogue from the objects detected in the image. Two steps that were taken to support the scientific analysis used calibration methods on the image flux and on the background. Convolution is also performed on the images to resolve the differences that arises from observing data at different mean wavelengths; to keep the all data consistent for analysis. Detection methods from Source Extractor were used to find all the objects within the dataset. And finally, the cuts that were applied to filter the catalogue to objects that have low magnitude values; since our focus is to look at distant galaxies.

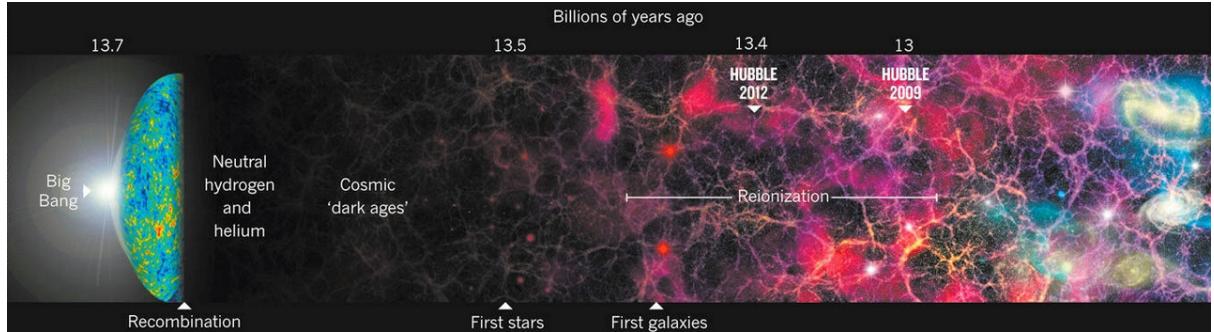
Once a calibrated catalogue is achieved, in section 5, this catalogue is passed through a photometric redshift software which provides redshifts estimates based on the flux values of the objects that best fit the templates used. It uses a linear combination of templates to minimise  $\chi^2$ .

The redshifts found are then used to find values for the various properties of interest in section 6.

Any modules that were used or any code described in this report will use this ‘font’ and therefore should be understood as such unless stated otherwise.

## 2 Background Theory

### 2.1 Epoch of Cosmic Reionisation



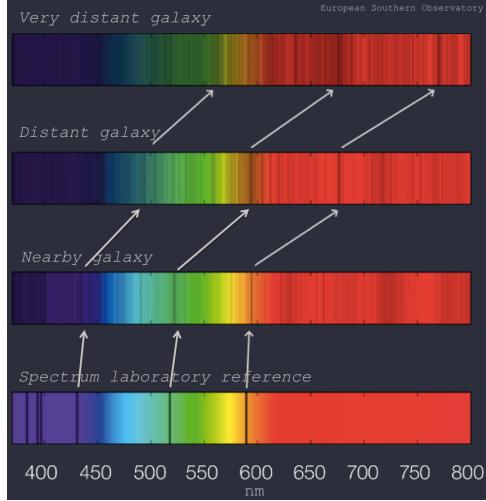
**Figure 2.1:** A theoretical model based on our current understanding of the early universe spanning from the beginning of time until about a billion years after the Big Bang. Formation of the first astronomical objects (including galaxies) can be seen at  $z \approx 20$ . Image taken from [ESA 2016](#).

Up until now, the understanding for stars and galaxies has come observations of the galaxy-rich universe we currently live in using successful telescopes such as the Hubble Space Telescope (HST). However, our knowledge on the astrophysics of how the universe came to be as we currently see it, is not very well understood ([Sterling et al. 1994](#)).

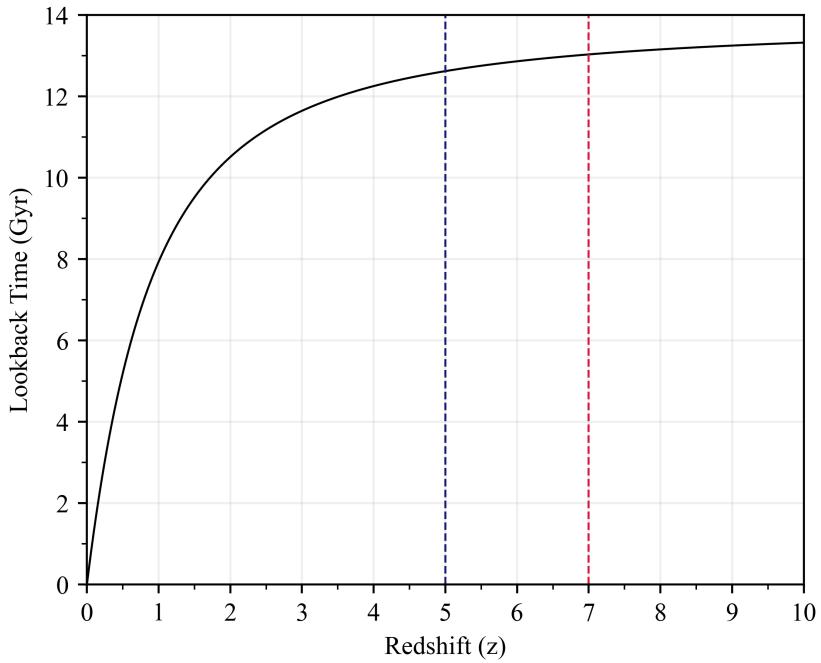
Theory suggests that stars a hundred times the mass of our Sun and, which are millions of times brighter, burn for only a few billions of years and then explode. The ultraviolet (UV) radiation from these allowed Hydrogen to be ionised by splitting the atoms into electrons and protons. From this point until about a billion years later is known as the “Epoch of Cosmic Reionisation”. During this period in the timeline of the universe, a small fraction of this matter condensed to form the first galaxies. The neutral Hydrogen was reionised by the growing radiation emitting from the formation of first stars and galaxies. Measurements from the cosmic microwave background radiation suggested that the midpoint of cosmic reionisation occurred at  $z \approx 8$ . The light emitting off these galaxies had spread throughout the Intergalactic medium (IGM) ([Planck Collaboration et al. 2020](#)).

Different methods have been used to measure the change in the amount of neutral hydrogen in the IGM over time to provide an overview of the timeline for cosmic reionisation. These methods indicate that neutral hydrogen within the IGM reionised at  $6 \lesssim z \lesssim 10$ . This is thought to be due to the ionising photon radiation emitted from star-forming galaxies ([Robertson et al. 2015](#); [Trebitsch et al. 2022](#); [Naidu et al. 2020](#)).

To understand the early history of the universe, we have to look at distant galaxies which, due to the finite speed of light, we observe them as they would have appeared to be in the distant past. This concept known as the lookback time is critical when deciding how far into the distant universe we need to observe (Centre for Astrophysics and Supercomputing n.d.). Figure 2.3 shows how far back in time we would be looking into the past from the present time of observation.



**Figure 2.2:** Wavelength observed in the laboratory rest frame would be doppler shifted due to photons emitting from the source that are in motion as they travel to the observer from a distant. The wavelengths will be larger, the longer the distance of observation. Image taken from University of Iowa Department of Physics and Astronomy [2024](#).



**Figure 2.3:** Lookback time is calculated using astropy to show the relationship between time and redshift. It gives us an idea how far back in time can we go with this dataset.

Figure 2.1 shows the evolution of the universe through time and equally through decreasing redshift. HST can observe until  $z \approx 10$  which is during the period of cosmic reionisation (Bouwens, Illingworth, et al. [2011](#)). To investigate the universe further back in time, we need to look at higher redshifts. Due to doppler shift (United States Navy Department [1969](#); Doppler [1846](#)), we will need to account for the difference of emission lines between observed and and rest-frame (laboratory) wavelengths. For this, the redshift is considered,

and is defined as the relationship between the observed ( $\lambda_{obs}$ ) and rest frame ( $\lambda_{rest}$ ) wavelengths:

$$z = \frac{\Delta\lambda}{\lambda_{rest}} = \frac{\lambda_{obs} - \lambda_{rest}}{\lambda_{rest}} = \frac{v}{c} \quad (1)$$

In this equation,  $c$  is a constant defined as the speed of light and  $v$  is the recessional velocity that represents the speed of the flux emitted from the object, relative to the observer.

Since the launch of James Webb Space Telescope (JWST) (McElwain et al. 2023; Rigby et al. 2023), more observations can be carried out which would give us a glimpse into galaxy evolution at its early stages (ESA 2024). Deeper observations of the universe reveals to us the earliest epochs of reionisation. The study for the epoch of Cosmic Reionisation and how it impacted, to form the universe we know today is an active field of research.

## 2.2 Emission Lines

The emission lines used for this analysis are the brightest emitters seen in galaxies. These are the H $\alpha$  and O[III] known at the laboratory wavelengths and can be estimated with equation 1 to find the wavelengths we need to observe. Additionally, we need to consider the filter in which the emission line is seen in.

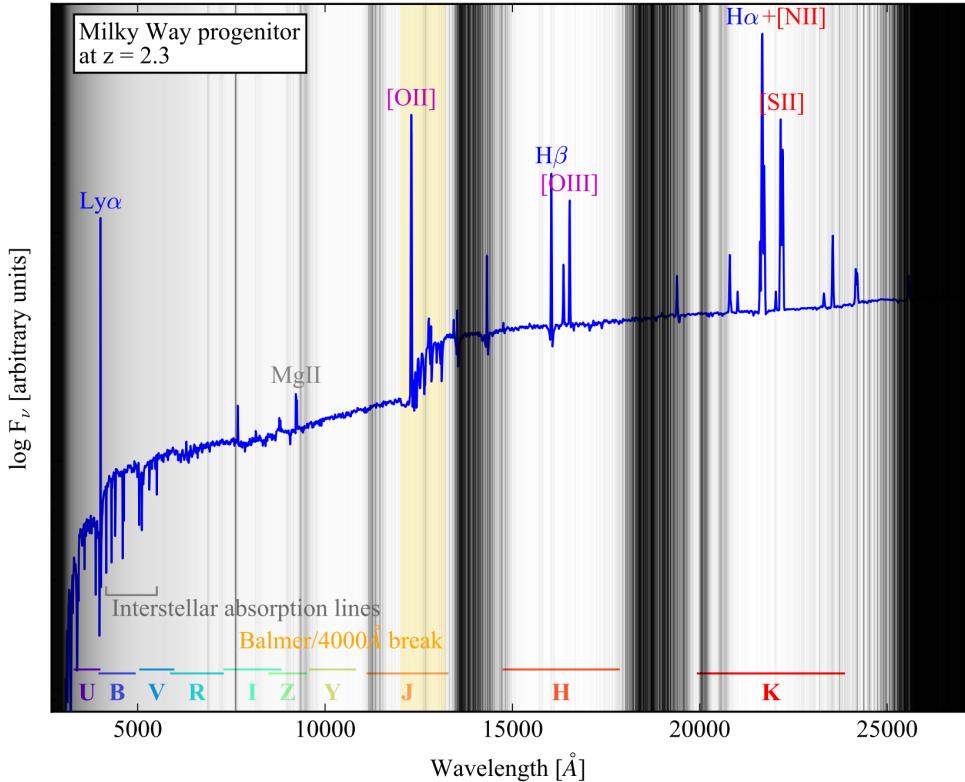
### 2.2.1 Hydrogen Alpha (H $\alpha$ )

H $\alpha$  is known to be a great tracer for star formation rate within a stellar population (Kennicutt 1992). Young massive stars that are born within these galaxies ionise the gas around them, creating a HII region around the star that consists of ionised Hydrogen gas. A large fraction of this gas starts to recombine and, creates a balance between ionisation and recombination within this layer. Due to the presence of recombination, a mass production of Hydrogen occurs and the intensity or luminosity of this emission line is proportional to the amount of hydrogen present. This in-turn is visible through the emission of the H $\alpha$  in spectra.

Understanding the intensity of the H $\alpha$  emission is essential to understand how quickly the stellar population was formed. There are many models that suggest when they began and ended but the journey in between is of particular interest.

### 2.2.2 Doubly Ionised Oxygen (O[III])

When a galaxy population is dominated by massive young stars, the galaxy becomes very blue and very hot. This is due to the mass production of hard, high energy ionising ph-



**Figure 2.4:** SED of a Star-forming Galaxy (SFG). Some of the brightest emission lines that can be seen in SFG are the O[III] and H $\alpha$ . The black lines scattered across the plot show wavelengths that cannot be seen from the ground. The spectra starts to fall off leftward of the Ly $\alpha$  (Lyman Alpha Break) at the rest-frame wavelength 1216 Å. Image taken from Förster Schreiber et al. 2020.

tons. These photons are the source for the O[III] emission. This emission supports the star formation rate (SFR) that can be deduced from the H $\alpha$  emission (Suzuki et al. 2016).

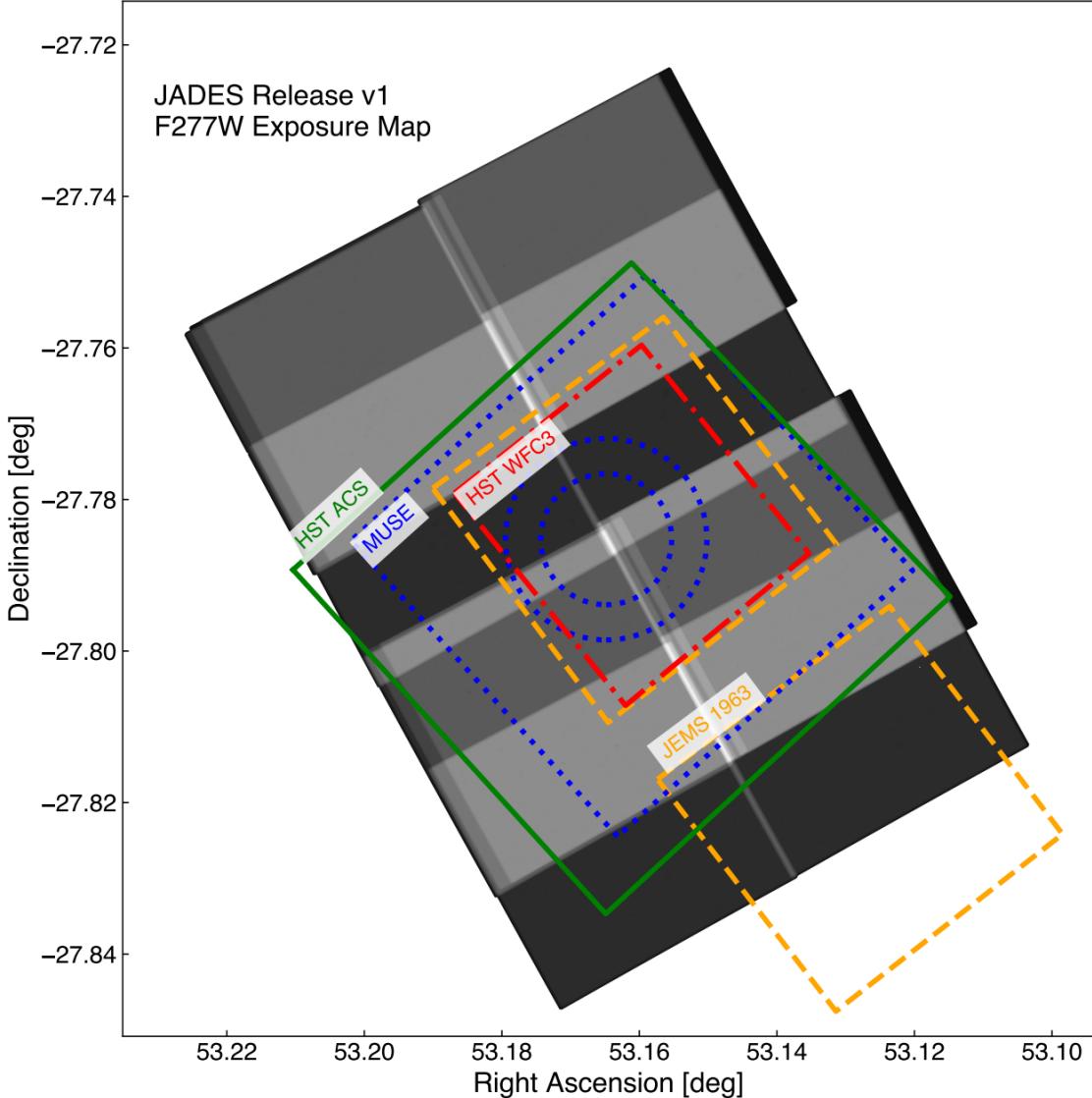
To find the wavelengths at which these emission lines can be observed at, equation 1 can be used. The H $\alpha$  is known to be at 6563 Å and The O[III] is known at 5007 Å.

## 2.3 JWST and JADES Survey

Among the many instruments that support JWST's observations, this report uses the data from is the Near-InfraRed Camera (NIRCam, Space Telescope Science Institute 2019) which is supported with spectroscopic data from NIRSpec, Space Telescope Science Institute 2023.

The filters analysed in this project are listed in the legend of figure 2.6 that use from the Near infrared Camera instrument of the JWST except for the first 5 which are taken from the Hubble Space telescope. All of mean wavelengths of the filters from HST are wide-band filters, 435W, 606W, 775W, 814W, barring one, 850LP which is the only long pass (LP) filter. LP means that the range of this filter is longer than the wide band range.

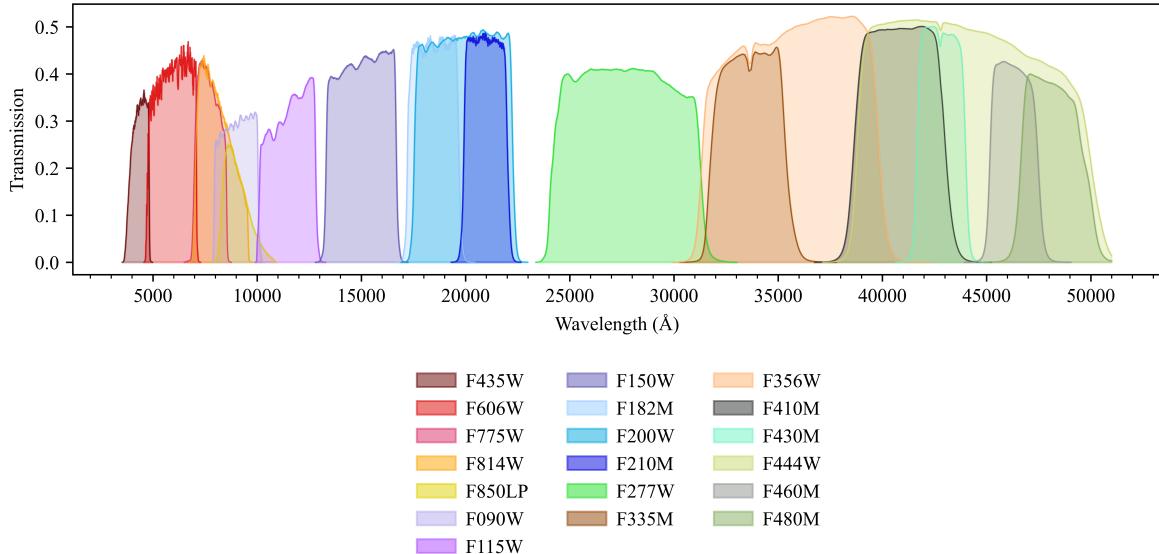
Out of the filters from JADES, the 182M and 210M are the medium (M) band filters, while



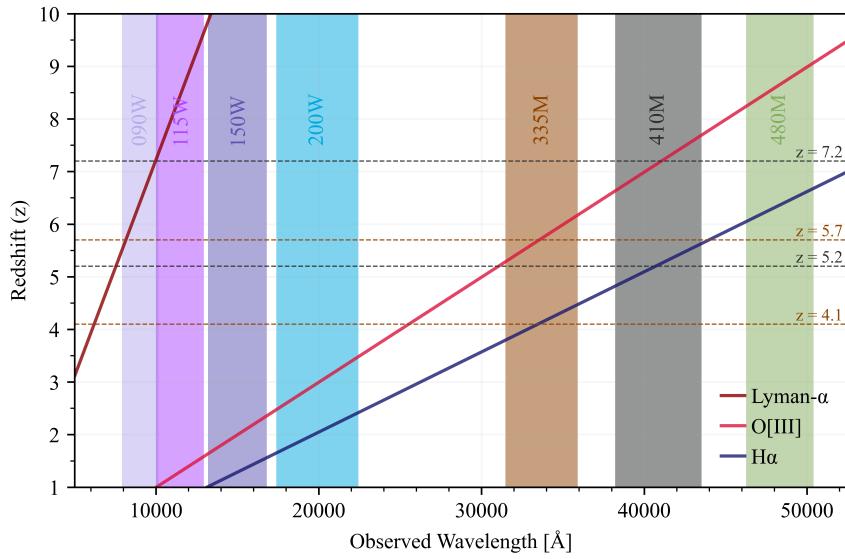
**Figure 2.5:** The JADES telescope field layout shown along with the field layouts of other surveys within this pointing. This graph shows the location of the HST ACS (Hubble Space Telescope Advanced Camera for Surveys) Ultra Deep Field (Beckwith et al. 2006) which are used for analysis. This graph is taken from Rieke et al. 2023.

the others fall in the wide (W) band filter category. Filters 090W, 115W, 150W, 180M, 200W and 210M are labelled as short wavelength filters due to its spatial distance between the peaks. From the same instrument, the long wavelengths selected to study are the 277W, 335M, 356W, 410M, 430M, 444W, f460, f480. Figure 2.6 shows the filter transmission profiles that indicate which filters cover a certain range of wavelengths.

This report looks into a galaxy population that have redshifts,  $z$  of about 5 or 7. The JADES image set was chosen for this project because the area looked at has been studied and extensively looked at previously. There would be enough data researched and documented to compare and understand the performance of the JADES telescope. This field of view is



**Figure 2.6:** Transmission Profiles for the filters used in this project. The data for this plot is taken from [SVO Filter Profile Service 2019](#) (Rodrigo, Solano, and Bayo 2012; Rodrigo and Solano 2020). The data for the 5 filters shown in the legend are for the HST ACS. The rest of the filters are for the NIRCam instrument of JWST. Filters that succeed with the letter ‘W’ stand are labelled wide-band filter; those with ‘M’ are medium-band filters. ‘LP’ stands for low-pass filter.



**Figure 2.7:** Using the known rest frame wavelength of these emission lines, this plot shows theoretically where an emission line could be observed the different filters at a given redshift. The emission lines of interest can be observed at unique redshifts for each filter. The filter widths are approximated between 10% of the maximum transmission within the filter. At  $z \approx 5.2$  and  $7.2$ , H $\alpha$  and O[III] can be seen peaking in the 410M filter, respectively. Also, at  $z = 5.2$ , the Lyman break known at  $1216 \text{ \AA}$  is seen just before the 090W filter band and for  $z = 7.2$ , it is seen just before the 115W filter band. These emission lines can also be seen peaking in the 335M filter at  $z \approx 4.1$  and  $5.7$  respectively.

**Table 2.1:** Each filter has a given literature value for the effective Wavelength,  $\lambda_{\text{eff}}$  and effective width,  $W_{\text{eff}}$  of the filter. These values correspond to the filters shown in figure 2.6. The data tabled here is taken from [SVO Filter Profile Service 2019](#) (Rodrigo, Solano, and Bayo 2012; Rodrigo and Solano 2020)

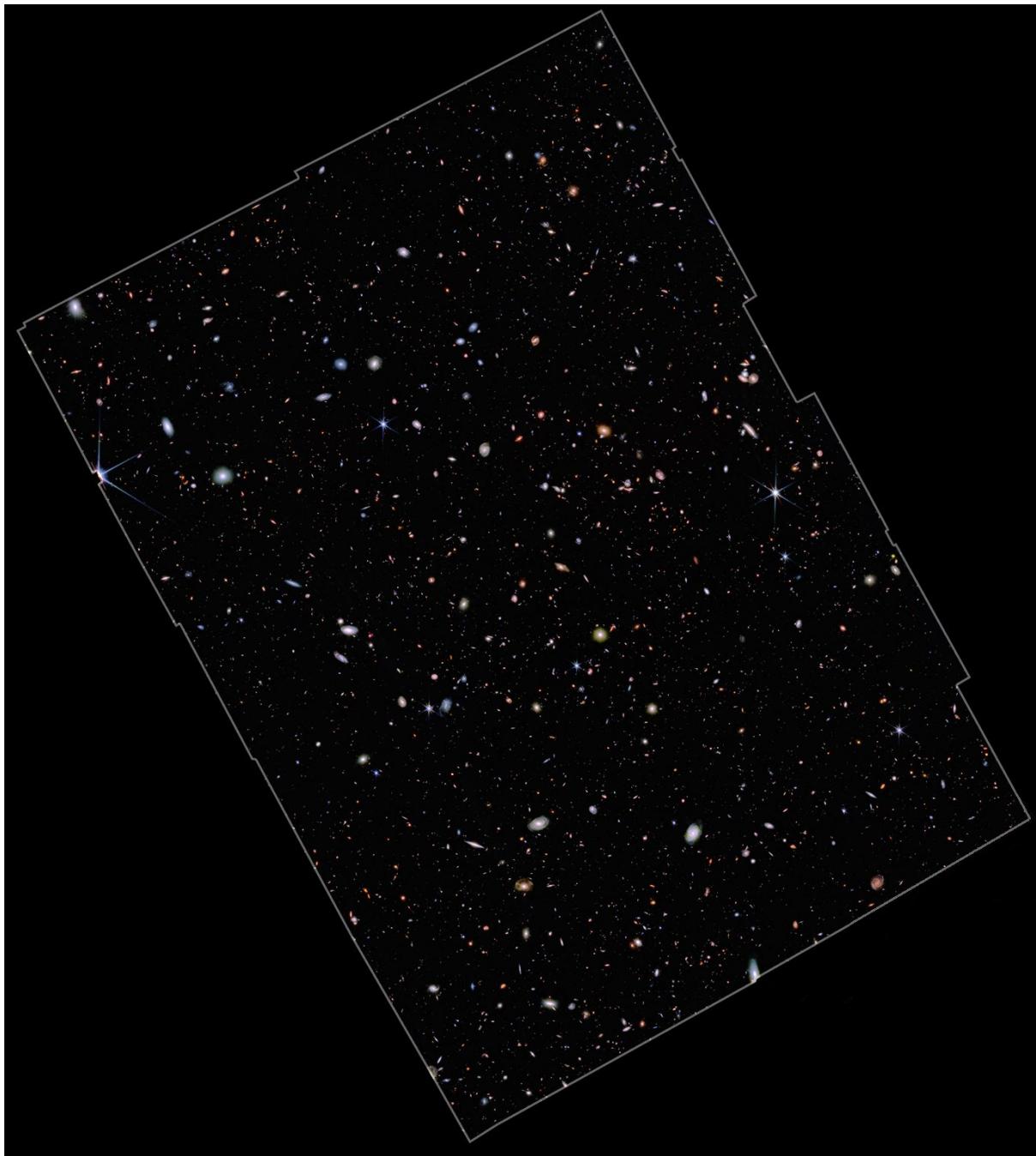
Filter	$\lambda_{\text{eff}}$ [Å]	$W_{\text{eff}}$ [Å]
435W	4341.62	821.68
606W	5809.26	1771.88
775W	7652.44	1379.43
814W	7973.39	1888.66
850LP	9004.99	1321.98
090W	8984.98	1772.74
115W	11433.62	2055.13
150W	14872.56	2890.43
182M	18388.83	2250.81
200W	19680.41	4190.39
210M	20908.35	2055.38
277W	27278.58	6614.61
335M	33537.23	3389.42
356W	35287.04	7239.30
410M	40723.18	4262.86
430M	42784.79	2295.14
444W	43504.26	10676.00
460M	46269.86	2309.00
480M	48139.11	3141.37

popularly known as the GOODS or the HST. Figure 2.7 shows where we would be able to observe a specific wavelength at a specific redshift in a given filter.

The amount of sky area that is covered by JADES is  $\approx 25$  square arcminutes and the spatial resolution of the images is  $\approx 0.03''$  per pixel. The zero-point of the data is  $23.9 \mu\text{Jy}$ . For each of the filters in JADES, the values for the effective wavelength and the effective filter width, useful for analysis are presented in table 2.1.

### 3 Inspecting the JADES Dataset

To understand what the data looks like visually, the image software, GAIA (Draper 2016), distributed by Starlink (East Asian Observatory 2023) was used to accomplish this. The data is stored in the form of .fits files; making the data easy to view with GAIA. Figure



**Figure 3.1:** Pseudo-coloured images using the data taken from JWST showing the region for the JADES survey field. This image has been generated from the data collected using NIRCam. This image uses the light from 115W as blue, 200W for green, and 444W for red. Image taken from NASA James Webb Space Telescope [n.d.](#)

3.1 shows a coloured image of this data used for the analysis in this project.

As expected, in some of the filters, the regions shown are different from one another (for reference see figure 2.5) since the focus goals for individual telescopes and instruments are different. Some filter images do not cover an identical region to that of the majority wide

band filters. This meant that objects found in one filter were not always found in others and only those found in all filter could be considered for analysis.

Prior to the start of this project, the images were aligned to make sure that despite how different the survey fields are from one another, the centroids of the same objects would be found at approximately the similar  $x$  and  $y$  coordinates within the image given a certain tolerance. This tolerance was noted to be roughly 25 pixels. The centre of the objects were checked to make sure these were aligned before any analysis was performed.

After the images had been realigned prior to the start of this project, the shift caused some residual flux to remain around the edges of the images. This needed to be taken into consideration when extracting the relevant objects at the various stages when `sep` was used to extract flux measurements.

The images were also noted to have significant extra space surrounding the data where the detector did not read anything in that space. Since, the data is stored as `.fits` files, cutting this region out of the image could not be easily done. The work around this issue was to simply removing this particular dataset space form the array using Python. This was implemented within the data analysis step wherever relevant.

### 3.1 Object Features Observed

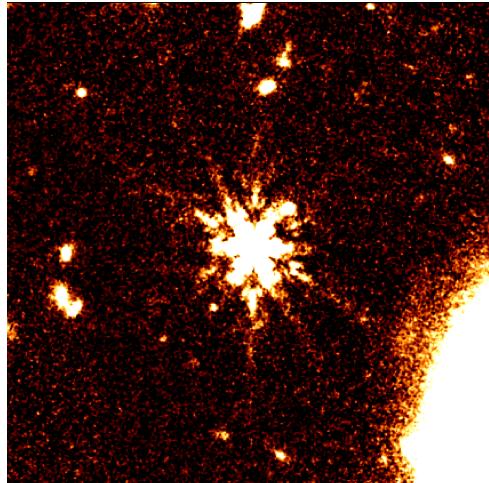
Going through the filters in order of increasing wavelength (as shown in figure 2.6) we can note that the spatial resolution decreases. Objects with diffraction spikes are known as point-source bodies and the light emitted from these objects originate from one source. Therefore, these objects can be known as the Stars within the image. More rounded objects that do not have this feature would be galaxies or more complex bodies with more than one source of flux.

Due to the different effective wavelengths and effective filter widths, for a given fixed exposure time, some objects are not visible in the dataset or can not be considered due to saturation. Some of the objects in the data have dead or black pixels close to the centroids. This occurs when in the emitting light, over a long exposure time, get over-exposes the detector and hence, saturates the data at that pixel point. Such objects found in any single filter were ignored to ensure consistent results throughout. This feature was typically seen in objects that were significantly larger than other objects in the image.

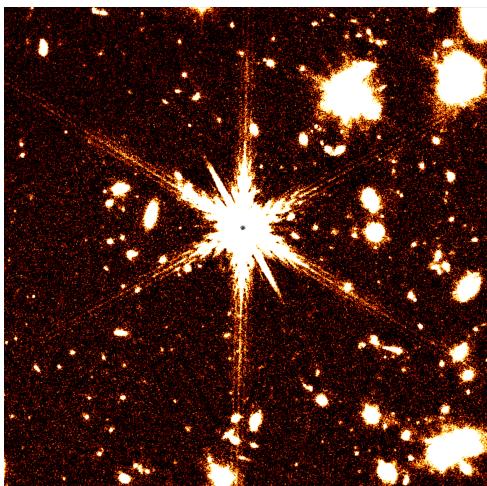
## 4 Calibrating the Data

### 4.1 PSF Homogenisation

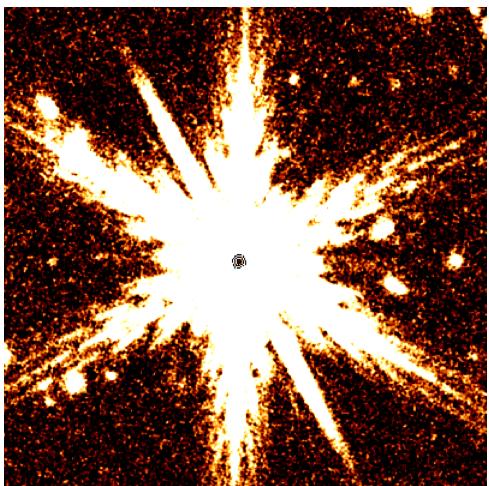
Firstly, The images need to be analysed to understand how much flux within an aperture can be detected to accept it at  $5\sigma$  significance. This is done by conducting a PSF (Point Spread Function) homogenisation. To perform this calibration, an ideal star is chosen from the



(a)



(b)



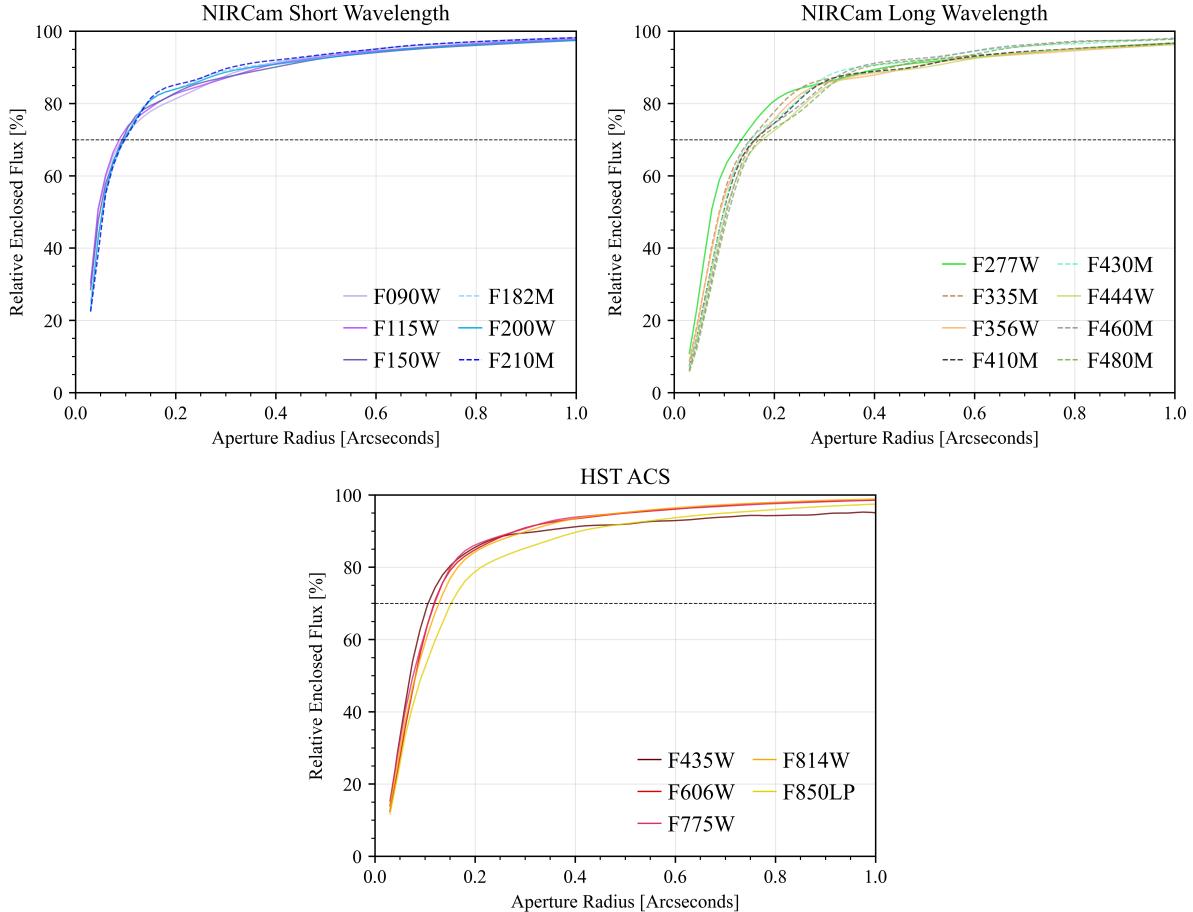
(c)

**Figure 4.1:** The diffraction spikes seen on these objects of focus tell us that these are stars. Shown in the top row, is the image of the ideal star that was chosen for calibration. The bottom row, shows an example of a bad pick for a star. In (b), the star is clearly surrounded by a lot of objects and would cause contamination with flux readings. Zooming into the centre of (b), shown in (c), there are saturated pixels and this star can not be used for calibration.

survey in order to determine the appropriate aperture radius. The star is chosen such that it was isolated and that there was no flux bleeding into the aperture from nearby sources. This was done by visually looking for a star using GAIA in the 444W filter. Since this filter observes at the longest wavelength amongst the wide filters. The resolution in this image would be the worst and therefore making the choice for an ideal star in this filter would ensure that it was seen in other filters. The star was also examined for any overexposed pixels that would potentially interfere with flux measurements.

To decide which aperture is most effective, a curve of growth is produced. This describes the flux output that is collected at each aperture of increasing radius. The increment used is 0.5 pixels that is approximately 0.015" for this dataset.

This process was done with the python module `sep` (Barbary 2016), which is based on the software SExtractor (Source Extractor, Bertin et al. 1996). `astropy` (Astropy Collaboration, Robitaille, et al. 2013; Astropy Collaboration, Price-Whelan, Sipocz, et al. 2018; Astropy Collaboration, Price-Whelan, Lim, et al. 2022) was used to fetch, organise and manage the data held in the `fits` files.



**Figure 4.2:** Flux output derived from Source Extractor for each filter at varying aperture radius. Radius increment step is selected at 0.5. The flux output of the chosen star is RA and DEC (03h 32m 35.102s, -27° 48' 44.60").

Figure 4.2 shows that there is a quick, exponential rise in the flux that is measured. From this plot, the flux of the star is concentrated in the centre. The curve flattens out at larger radii. This is because majority of the object flux has been collected and we are now adding in the sky background noise which is smaller in value compared to the flux of the star.

Although, this noise is smaller, it still contributes to the flux and hence, we need to consider an aperture size that maximises the S/N ratio; meaning that we would like to have more signal flux than background noise collected within the aperture. For the 444W filter, only about  $\approx 70\%$  of the total flux is considered to reduce the risk some of background noise being calculated into the flux measurement. Although, we can accept 70% for this filter, not all the other curves that describe the other filters have 70% of the total flux within the same radius of aperture.

## 4.2 Constructing Convolution Kernels

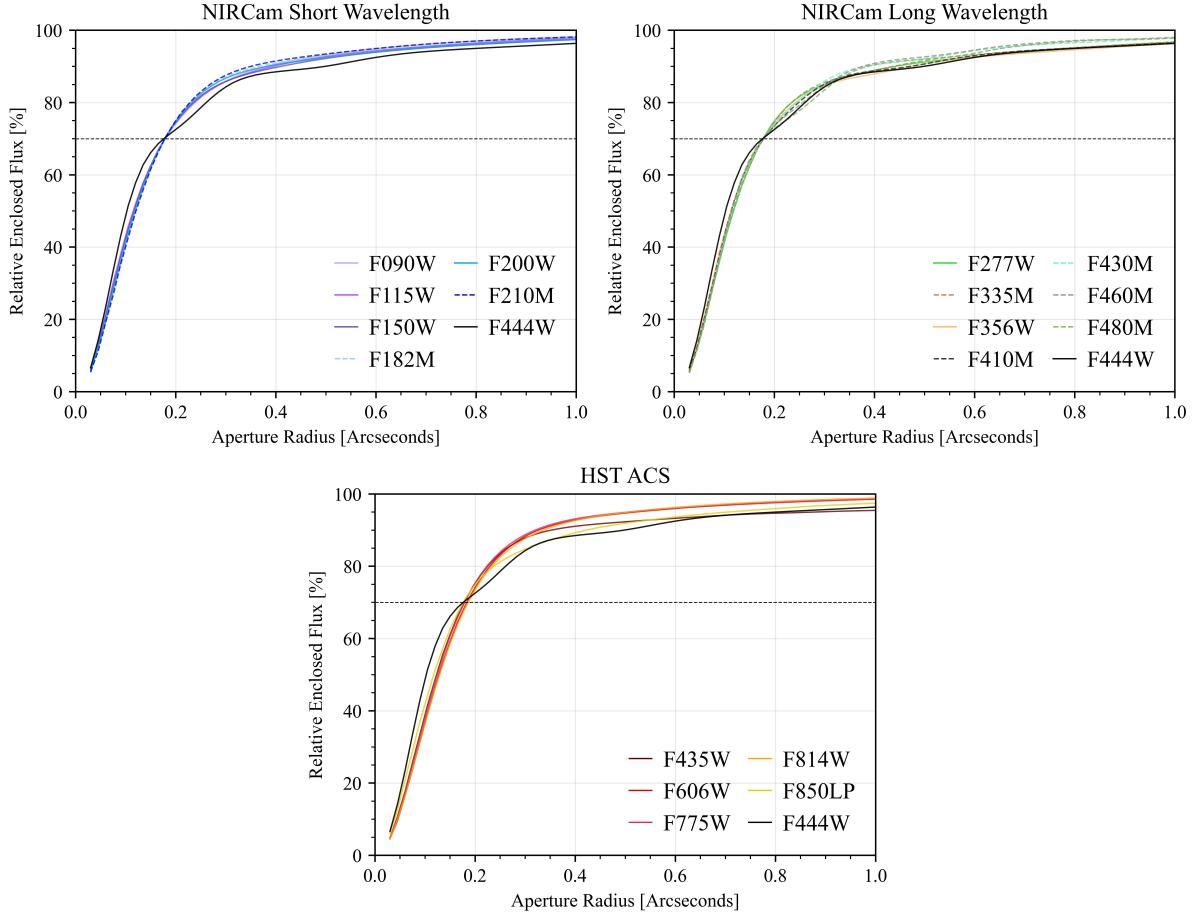
The moffat profile that is based upon the lorentzian distribution that used to describe the flux output of a star. Some researchers use a gaussian distribution profile however, the reason for choosing moffat over gaussian is due to the extra degrees of freedom available, that help us to best describe the flux output. Its particular importance in astrophysics is due to its ability to accurately PSFs, whose wings cannot be accurately portrayed by either a Gaussian or Lorentzian function. The form of the Moffat distribution is,

$$\text{PSF}(r) = \frac{\beta - 1}{\pi\alpha^2} \left[ 1 + \left( \frac{r}{\alpha} \right)^2 \right]^{-\beta} \quad (2)$$

Gaussian distributions can be recreated as a moffat distribution with equation 2. The width of the distribution is the only parameter that governs the shape of a gaussian while a moffat only has 2; one that changes the tail of the distribution and the other that changes how defined the peak is. The parameters,  $\alpha$  and  $\beta$  found for each filter to insert into equation 2 are listed in table 4.1. These values bring the flux of the star to the same value for the same radius of aperture at 70%.

**Table 4.1:** Alpha and Beta parameters chosen for each filter to create convolution kernels with respect to F444W. Data seen in the filter listed here will match the data in F444W at the selected aperture at the flux acceptance value, 70%.

Filter	$\alpha$	$\beta$
435M	3.250	2.2
606W	3.200	2.3
775W	3.300	2.3
814W	3.300	2.3
850LP	2.100	2.2
090W	3.100	2.2
115W	3.220	2.2
150W	3.225	2.2
182M	3.325	2.2
200W	3.325	2.2
210M	3.400	2.2
277W	2.665	2.2
335M	2.100	2.2
356W	1.800	2.2
410M	1.500	2.2
430M	1.800	2.2
460M	1.400	2.2
480M	1.000	2.2



**Figure 4.3:** The plots show the enclosed flux output from Source Extractor for each of the selected filters. The top row are for the JADES dataset while the bottom row is for the HST UDF field. The data here is collected after convolving all the data to match the f444w filter using the parameters in table 4.1 with equation 2.

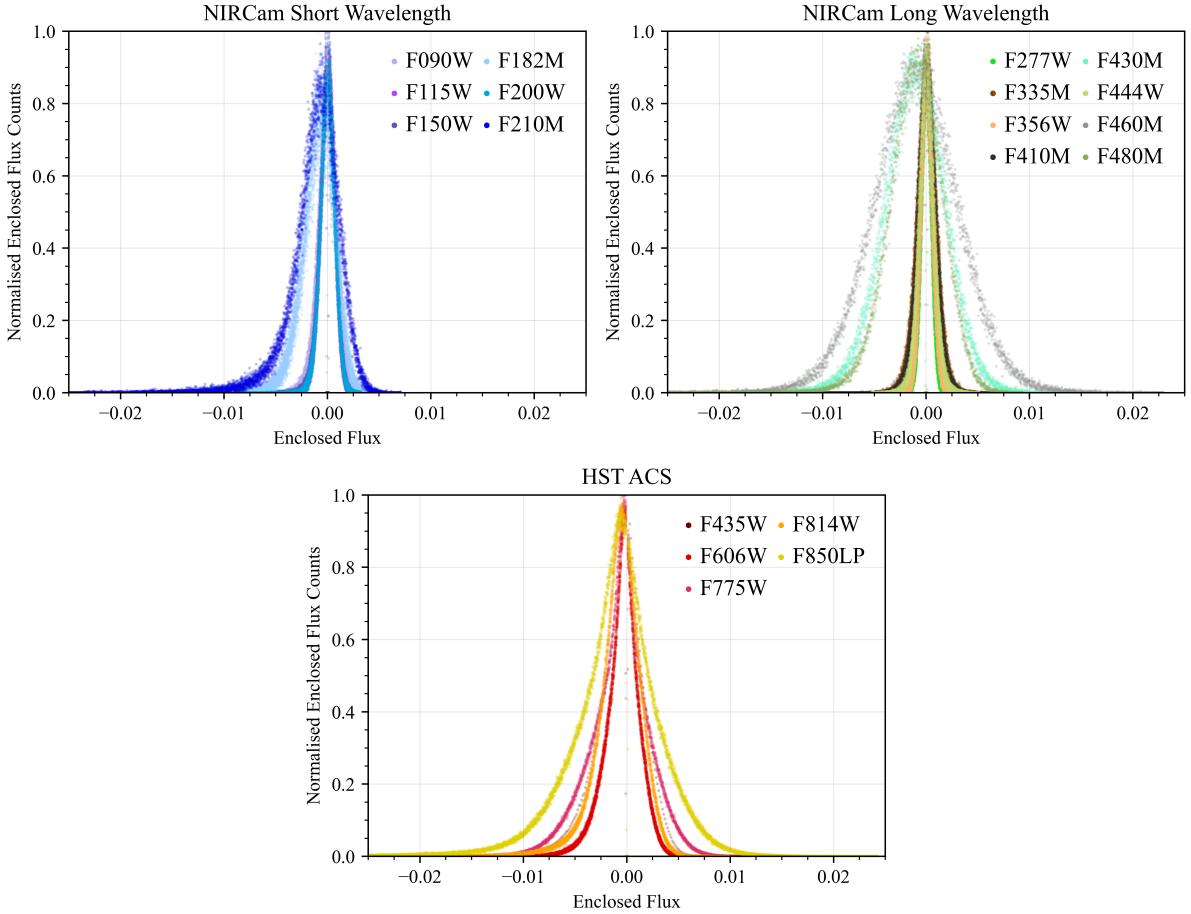
Firstly, to find a suitable aperture size where the overall calculated flux is around 70% of the total flux, the images need to convolved with respect to a single filter in order to normalise the flux output calculated. Using an aperture size that obtained the desired 70% of the total flux in one filter could give us a smaller percent of the flux in other filters. Convolving the images allows us to purposely blur the data so that the flux in the chosen filter image will match the new image data in the convolved images.

For this project, this process is done by keeping the 444W filter constant and convolving the other filters with respect to this filter. Using the moffat distribution profile (equation 2) with specific kernel values (refer to table 4.1) for each convolution kernel, the flux output can be summed in each aperture. The convolution kernels are set in such a way that we expect 70% of the total flux at the same aperture radius. Figure shows the change in the flux within the aperture for each filter with changing radius.

The initial value used as an estimate for the aperture size was set at  $\approx 5$  pixels in radius which translates to  $0.175''$  on the JADES images. This was chosen for a specific star selected

in section 4.1, based on it's location in the image compared to it's neighbouring stars and the size of the star. This was done to ensure a fast calibration process and to only consider objects that haven't been saturated due to observational reasons.

### 4.3 Mapping Background Noise



**Figure 4.4:** Background noise collected from a filtered grid of apertures at 70% of the total flux. The distributions are shown for each of the filters; HST (bottom row) has a border distribution compared to JADES (Top row) due to the spatial resolution that is proportional to the wavelength of the filter.

Since we focus on looking at distinct galaxies in this project, and thus faint galaxies, the error on the flux, in theory, comes from the background noise of the data. This calibration helps us to understand what the sensitivity is of the dataset. This is done by drawing up a set grid of apertures evenly spaced out across the entire image. These apertures are then filtered through the following criteria by considering the flux contained in them. The filtration is as follows:

1. Apertures that are exactly 0 are removed from the grid.

2. Apertures that sit just outside the edge of the detection region are also removed. This was done due to the realignment mentioned in section 3 to ensure a clean straight line along the edge of the data.
3. Apertures that were on top of objects were also removed.

The last step of the filtration is done using the segmentation map for each of the datasets. The segmentation map is produced by default with `sep`. When Source Extractor sees an object, the pixels on this map where the object lies are set to the id number of the catalogue, while any other pixels are set to zero. This makes it useful for us to find any apertures that sit on objects by simply picking apertures that have a sum-in-aperture that is greater than zero.

Once all the apertures have been carefully selected, their fluxes can be plotted as a distribution for all the filters. This distribution is expected to be a Gaussian centred at 0 since background noise would be significantly smaller than the flux produced by objects. We also expect the probability of getting positive and negative values would be somewhat equally distributed as noise ideally fluctuates.

The distribution of the curves show a strong peak in the centre however, there is a long tail at the ends of the distribution. This is due to artifacts such as ghosts and wisps that source extractor did not recognise during the aperture filtration process. Some of the distributions have a narrower width, corresponding to the data from wide filters.

After the data collected from the telescope is background subtracted, theoretically, the images should be flat. However, this subtraction isn't always perfect and there is some residual background flux that remains which is the noise within the images that has to be accounted for during calibration. At the mean wavelengths of the filters (refer to table 2.1), this could be from zodiacal light from dust within the solar system that is being reflected by the Sun.

## 5 Building a Catalogue

### 5.1 Deriving the Error on the Flux

The widths of these distributions can now be used to find the sigma estimates that set the error on flux of the dataset. This would ideally be done by taking the standard deviation of the distribution. However, as discussed, we have long tails at the ends of the distribution, which are the outliers in the dataset—caused by image artifacts. To achieve a more robust estimate, a median absolute deviation (MAD) statistic,  $\sigma_{\text{MAD}}$ , is calculated and set as the  $\pm 1$  sigma error. This estimate is found using the `mad_std` function from the Python module `astropy`.<sup>1</sup> The true sigma,  $\sigma$ , and the MAD estimate, follow a linear relation that is defined as:

$$\sigma \approx 1.4826 \sigma_{\text{MAD}}$$

---

<sup>1</sup>Values calculated using `astropy.stats.mad_std`

The magnitudes that correspond to these error values can be found with the commonly known magnitude equation.

$$m_{\text{filter}} = -2.5 \times \log(f_{\text{filter}}) + ZP \quad (3)$$

In Equation 3,  $f_{\text{filter}}$  is the flux point of the object within any single filter and  $ZP$  is the known zero-point of the dataset (refer to Section 2.3). The errors and magnitudes found for this dataset are tabulated in table 5.1.

**Table 5.1:** The MAD estimates and the Limiting Magnitudes found for each filter. The MAD estimates are taken as the flux errors for each filter. The limiting magnitude gives us a cut-off for detected objects.

Filter	MAD estimate, $\sigma_{MAD}$	Limiting Magnitude ( $\mu\text{Jy}$ )
435M	$2.101 \times 10^{-3}$	28.85
606W	$1.529 \times 10^{-3}$	29.19
775W	$2.520 \times 10^{-3}$	28.65
814W	$1.909 \times 10^{-3}$	28.95
850LP	$3.567 \times 10^{-3}$	28.27
090W	$9.343 \times 10^{-4}$	29.73
115W	$7.674 \times 10^{-4}$	29.94
150W	$7.982 \times 10^{-4}$	29.90
182M	$1.580 \times 10^{-3}$	29.16
200W	$7.829 \times 10^{-4}$	29.91
210M	$2.058 \times 10^{-3}$	28.87
277W	$5.881 \times 10^{-4}$	30.23
335M	$1.012 \times 10^{-3}$	29.64
356W	$6.650 \times 10^{-4}$	30.10
410M	$9.715 \times 10^{-4}$	29.68
430M	$3.017 \times 10^{-3}$	28.45
444W	$8.057 \times 10^{-4}$	29.89
460M	$4.142 \times 10^{-3}$	28.11
480M	$2.603 \times 10^{-3}$	28.61

The limiting Magnitudes gives an indication to how faint we can reach before we would have to stop considering any sources detected passed this limit as flux objects. These values tell us where to truncate the catalogue based on the reasoning.

## 5.2 Gathering Flux Data

A catalogue is built by looking at all the filters in dual image mode with Source Extractor. In the python version of the software, this mode isn't available and hence, a process for this was constructed using `sep`. This process utilises a detection image and a photometry image. The detection image is used to identify sources in the image and then record the  $(x, y)$  co-ordinates. Apertures are then placed at these co-ordinates onto the photometry image to collect the sum-in-aperture as the flux value. The detection image doesn't change while the photometry image is set by cycling through the all the filters.

The 356W filter is set as the detection image because this is a wide filter. Certain physical processes emit or absorb light at this specific wavelength, including the dust surrounding stars that emit thermal radiation at Infra-Red (IR) wavelengths. From table 2.1 the  $\lambda_{\text{eff}}$  for the 356W filter is 35287.04 Å, making this filter quite red. This makes the filter mass sensitive and ideal for detecting objects.

Due to the reasons listed in section 3, the objects in some filters would be offset and this needs to be implemented into the process. A square tolerance pixel box was created to search for the same object and record their flux.

The JADES dataset has been established to have large file sizes. Within the code that was used to perform the flux extraction, the data was first read for a list of aperture coordinates of the objects that Source Extractor detected at first glance. The image was then cut into parts that were individually processed for flux values to make efficient use of the computational power available.

The catalogue was then complied into a .txt file with an ASCII format.<sup>2</sup> The file contained the flux values and the errors on the flux as well. Errors were constant across the column as these were found for each unique filter. The magnitudes were found using equation 3 for the 356W filter and added to the catalogue. Additional information included in the catalogue were the RA and Dec, and the  $(x, y)$  co-ordinates of each object detected.

A python function was then created and used to read-out the data to the file. Within the function, the catalogue was managed as the variable datatypes: `dict` and `pandas.DataFrame`<sup>3</sup>. Two cuts were made on the catalogue:

- Sigma-Cut: Any flux values in the 356W filter for an object that were less than 5 times the MAD estimate in this filter were removed. This is done to clean the catalogue of any that potential noise values that may have been recognised as objects.
- Magnitude-Cut: Using the magnitudes, any objects that were found to have values greater than the limiting magnitude of 356W filter were also removed from the catalogue. This was done to only consider faint objects.

---

<sup>2</sup>Data entries in the file are separated with spaces or tabs.

<sup>3</sup>The Pandas Development Team n.d.

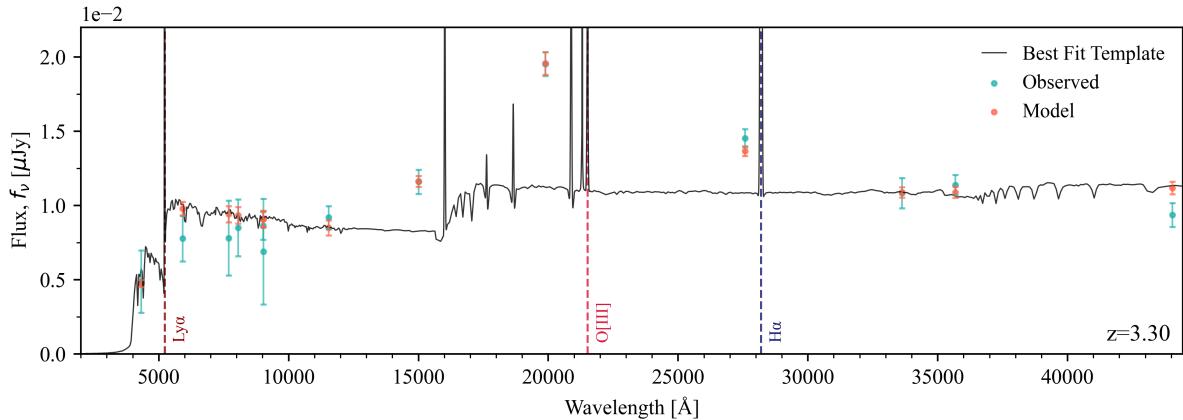
## 6 Results and Discussion

### 6.1 Determining Redshifts

To derive redshifts estimates on the flux data, the catalogue is passed through a photometric redshift code known as EAZY, Brammer et al. 2008; Brammer, Gabriel 2015. This program is used as its python equivalent, eazy-py for this project. It is used to derive redshifts since we do not have information on the spectroscopic redshifts associated with all the objects in the catalogue. It uses a linear combinations of given templates and, the commonly known  $\chi^2$  minimisation technique to make estimates. The algorithm doesn't involve the use of any machine learning techniques but instead, uses a redshift quality parameter to estimate the reliability of the redshift estimate it makes. This program is known to work very well on established public databases (Brammer et al. 2008).

The software was setup to match the filters known in this project to the filters in the JADES and HST ASC<sup>4</sup> filters that were included by default within EAZY. The errors were attached to the corresponding filters as well.

Nine different templates were used to fit the data which included a ‘high EW’ and a ‘dusty galaxy’ template. The data was then fitted using these template to a grid of redshifts that were created by EAZY. The redshift corresponding to the lowest value of  $\chi^2$  is then set as the estimate.

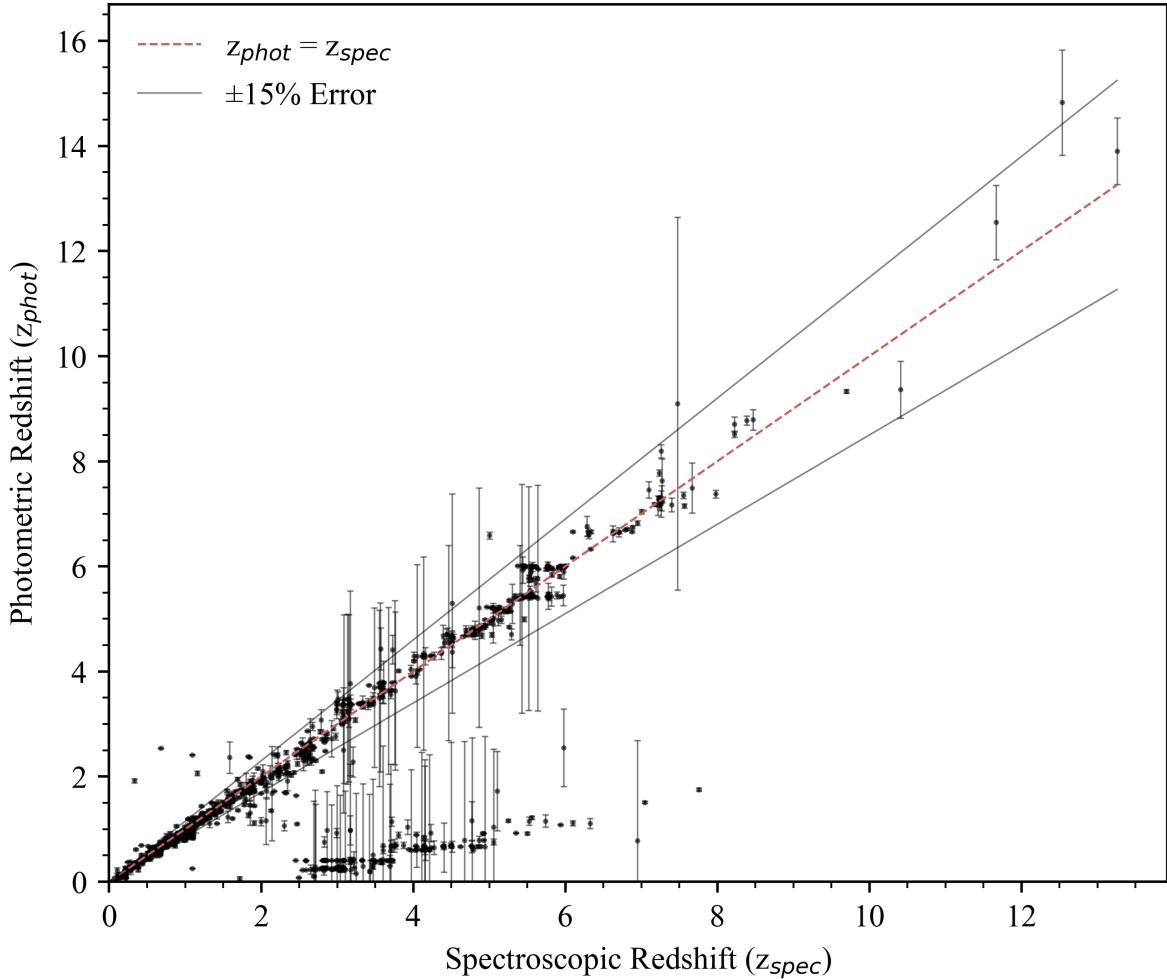


**Figure 6.1:** Best fitted object with the lowest  $\chi^2$  value. The best fit template is the overall model template that is used to fit this data built within the program. The wavelengths marked here are labelled as they would have appeared in the rest-frame of observation. This is calculated using equation 1. The model data points (in orange) are transformed from the observed data (in teal). The observed data is plotted as the flux value recorded at each filter.

Before the catalogue is passed into EAZY, filters that don’t show any objects at certain coordinates were removed from the catalogue. This was done to better optimise the program to construct the best fitting template for the data.

<sup>4</sup>Advanced Survey for Cameras

Figure 6.1 shows an the best fitted SED for a galaxy at  $z \approx 3.3$ . The redshift is estimated from the best fitted template.

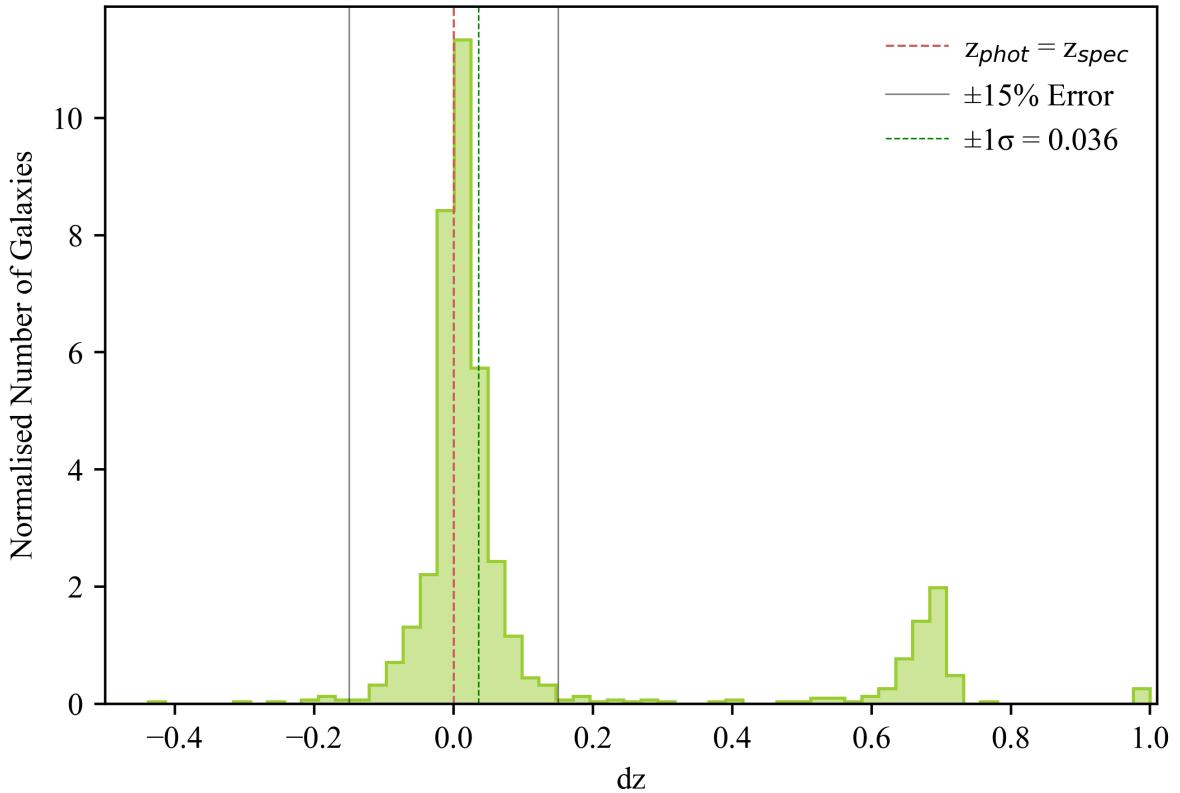


**Figure 6.2:** The photometric and spectroscopic redshifts plotted against each other shows well EAZY fit the data for redshifts. The objects plotted here are those that only had spectroscopic redshifts.

The scatter plot, figure 6.2, shows the spread of photometric redshift estimates against their spectroscopic redshifts that taken from Eisenstein et al. 2023; Oesch et al. 2021. The sources were matched using TOPCAT by using the the in-built angular separation method. This method calculates the difference in angle between 2 objects using RA and DEC. Each object in the spectroscopic catalogue was cross-matched with each object from the photometric catalogue discussed in section 5 within a tolerance of 0.2".

The error on the sources in figure 6.2 is given as the  $1\sigma$  error by evaluating it as the average of the difference between the 50<sup>th</sup> and 16<sup>th</sup>, and the difference between the 84<sup>th</sup> and 50<sup>th</sup> percentiles of the final probability distribution that EAzY calculates.

Figure 6.2 further displays how well the fitting had been done. Majority of the sources lie on the 1:1 line with a few exceptions. This statement can be reinforced with figure 6.3. The



**Figure 6.3:** This distribution is a reflection of figure 6.2. Here, we can better understand how well the redshift was estimated with EAzY. The coloured lines describe the same parameters presented in figure 6.2. Sources that lie between  $-0.15$  and  $0.15$  can be accepted fits. Also, shown is the  $\pm 1\sigma$  error on data.

y-axis shows the distribution for the  $dz$  statistic and is defined by the following equation:

$$dz = \frac{z_{\text{spec}} - z_{\text{phot}}}{(1 + z_{\text{spec}})}$$

The figure shows a peak Gaussian distribution centered at 0 (corresponding to the  $z_{\text{phot}} = z_{\text{spec}}$  line). The width,  $\sigma$  of the distribution is a measure for the scatter in the data shown in figure 6.2 along the  $z_{\text{phot}} = z_{\text{spec}}$  line (in red). The grey lines outline a boundary where beyond this region,  $|dz| \geq 0.15$ . Any sources that fall within this region are known as catastrophic outliers. Here sources have underestimated redshifts i.e., they have higher spectroscopic values since they may have not been fitted well with EAzY.

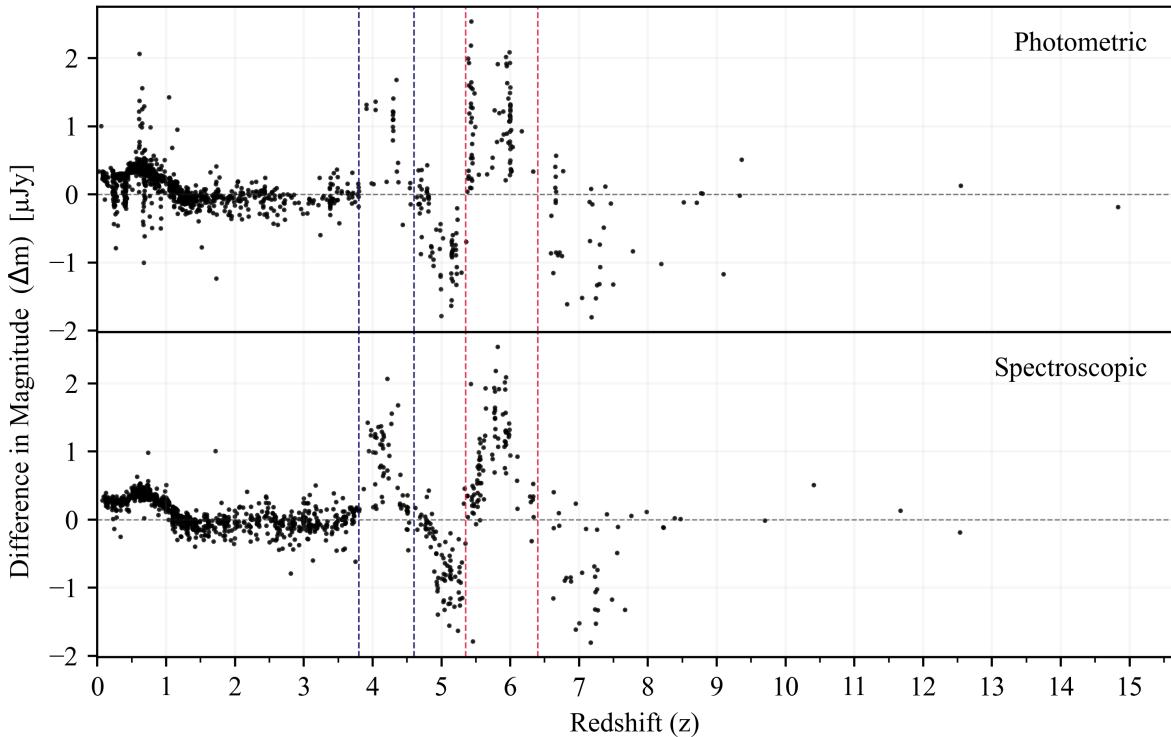
This could be due to the specific configuration of the program that may have not allowed EAzY to best calculate an estimate for these sources. Figure 6.3 supports these statements as there is a smaller second peak at  $dz \approx 0.7$ . The fit could be improved by adjusting the zero-points before running the catalogue again through EAzY.

## 6.2 Equivalent Widths of the Emission Lines

Having verified the integrity of the redshift fit, the colour magnitudes and their corresponding equivalent widths can now be derived. Firstly, we establish the 410M filter as the filter where emission lines are to be observed along with an underlying continuum and the 335M filter for the estimate on the continuum emission.

Ideally, the continuum and the emission lines would be taken from within the same filter, but due to the short time that can be spent on this project we have taken this approach. Since the 410M filter was chosen for identifying emission lines, we have to take the next closest filter band as the continuum filter. From figure 2.6, the 335M and 410M are one of the long wavelength medium band filters and from figure 2.7, redshifts do not have to be established at very high values to find objects that emit the emission lines H $\alpha$  and O[III].

Using this theory, we can now evaluate magnitudes from the chosen filters that can be used to later identify emission lines. The magnitudes are found using equation 3 which are then used to subtract the continuum from the emission line. Therefore doing  $\Delta m = 410M - 355M$ , we would be able to clearly observe where in the spectrum, the emission line occur.



**Figure 6.4:** The  $x$ -axis shows the difference in the magnitude values,  $\Delta m$  calculated in the 410M filter and the 335M filter. The top plot shows the spread of the photometric redshift estimates while the bottom shows the literature values for the same spread. The coloured lines mark where we have isolated sources pertaining to a specific redshift and emission line corresponding to the grey lines in figure 2.7.

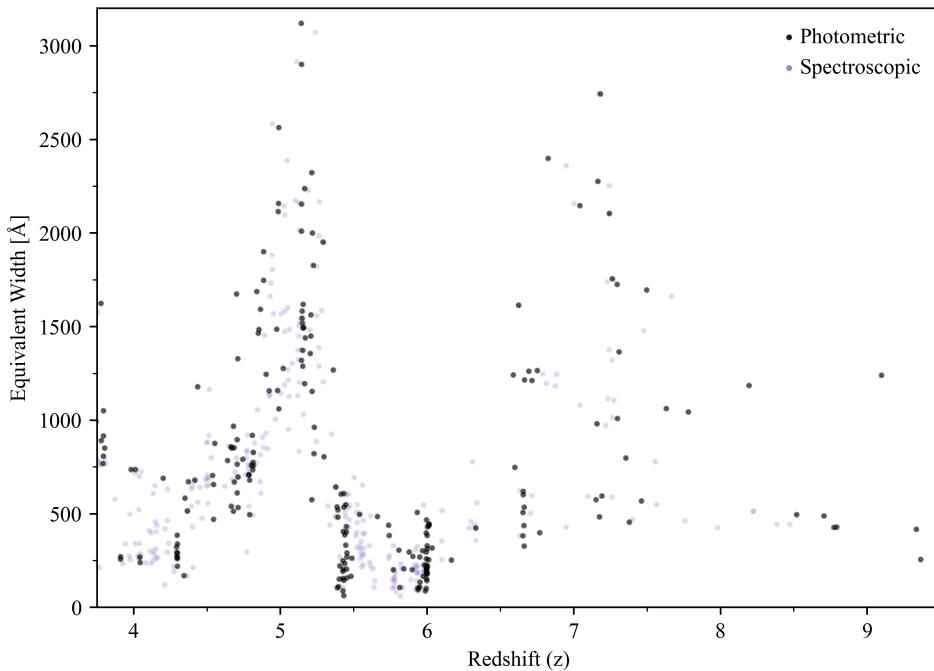
The calculated difference is shown against redshifts in figure 6.4. There are two distinct

peaks at  $z = 4.2$  and  $5.8$ . These refer to the  $\text{H}\alpha$  and  $\text{O}[\text{III}]$  emissions as seen for the  $335\text{M}$  filter in figure 2.7. The 2 dips right behind the emissions are not absorption features and are present in the scatter due to the magnitude subtraction of the filters. These dips actually represent the same emission lines (from left to right) that are observed in the  $410\text{M}$  filter.

The strength of an emission line in any given filter is found by measuring the equivalent width (EW) of the line. The equivalent width is defined as the luminosity of the line divided by the luminosity of the continuum. We try to find how much of the continuum is actually part of the emission. In figure 6.4, the grey line is where the continuum lies and the coloured lines shows the emission lines of interest, from which we will derive the EW.

This is done by finding the area under the emission line within the width of the filter it is observed in. The edges of the filter are defined as  $\lambda_1$  and  $\lambda_2$  and the difference of these values,  $\Delta\lambda$  gives us the width of the filter. Table 2.1 lists the known effective widths,  $W_{\text{eff}}$  for each filter that can be used in equation 4. The derivation is as follows:

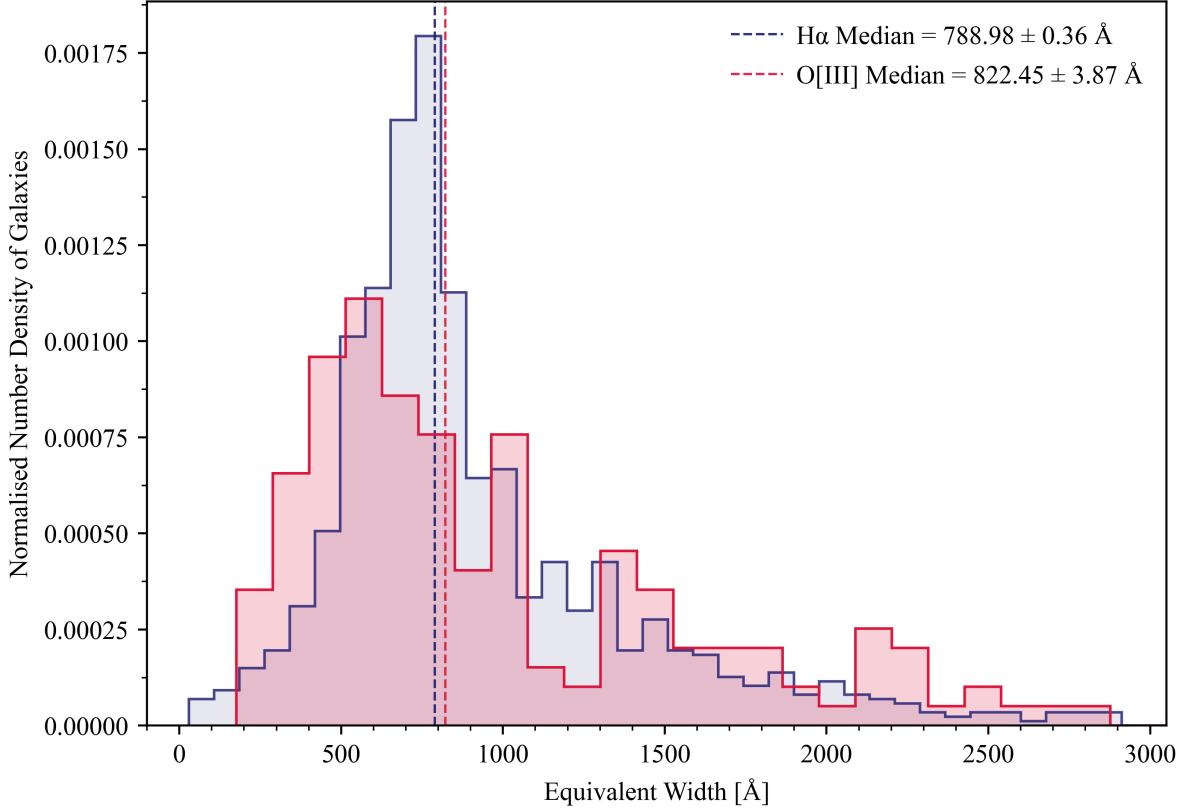
$$\begin{aligned} \text{EW} &= \int_{\lambda_1}^{\lambda_2} \left( \frac{f_\lambda - f_c}{f_c} \right) d\lambda \\ &= \int_{\lambda_1}^{\lambda_2} \left( \frac{f_\lambda}{f_c} - 1 \right) d\lambda \\ &= \Delta\lambda \left( \frac{\langle f_\lambda \rangle}{\langle f_c \rangle} - 1 \right) \end{aligned} \quad (4)$$



**Figure 6.5:** Objects are isolated at specific redshift ranges of interest to focus our analysis with galaxies that are strong emitters. The spectroscopic (literature) data is plotted in the background to show that the fitted distribution with EAzY mimics the literature scatter.

where in the above,  $f_\lambda$  is the flux value of the emission plus continuum and  $f_c$  is the flux of the continuum. Using this, we can find the EW as,  $\text{EW} = 4262.86 \times \left( \frac{f_{410\text{M}}}{f_{335\text{M}}} - 1 \right)$ . These values would have units of angstroms.

The analysis is focused on the galaxies seen in the 410M filter and therefore, redshift cuts are applied to consider galaxies at  $4.6 \leq z \leq 5.35$  and at  $6.75 \leq z \leq 7.55$ . These correspond to the H $\alpha$  and O[III] emissions, respectively. Figure 6.5 shows the EW of these isolated objects as a function of redshift and there is a strong H $\alpha$  emission seen at  $z = 5$  in addition to a wider yet intense peak emission at  $z = 7.2$ .



**Figure 6.6:** Shown here are the EW distributions for H $\alpha$ (in blue) and O[III](in red). The number of sources for O[III] are less compared to H $\alpha$  and hence the distribution is more staggered.

The H $\alpha$  and O[III] EW distributions in figure 6.6 are skewed gaussians that describe the number of galaxies that are found at specific EWs. The H $\alpha$  is higher than the O[III] due to the number of galaxies found within the respective redshift constraints. As higher redshifts are reached, it gets harder for EAzY to be able to make proper estimates since we know very little about them. As discussed earlier, EAzY is built with existing photometric data and does not learn from this. It simply fits the data to a value it deems best.

There is a spread of EWs from (0 – 3000) Å however, there are a higher number of galaxies with comparatively lower EWs. Those galaxies which have extreme EWs, ( $> 1000$ ) Å, occur when the galactic processes are dominated by very young stellar populations. It is only in these conditions where we are able to see prominent optical emissions (H $\alpha$  and O[III])

(Stark 2016).

### 6.2.1 Ionising Photon Production Efficiency

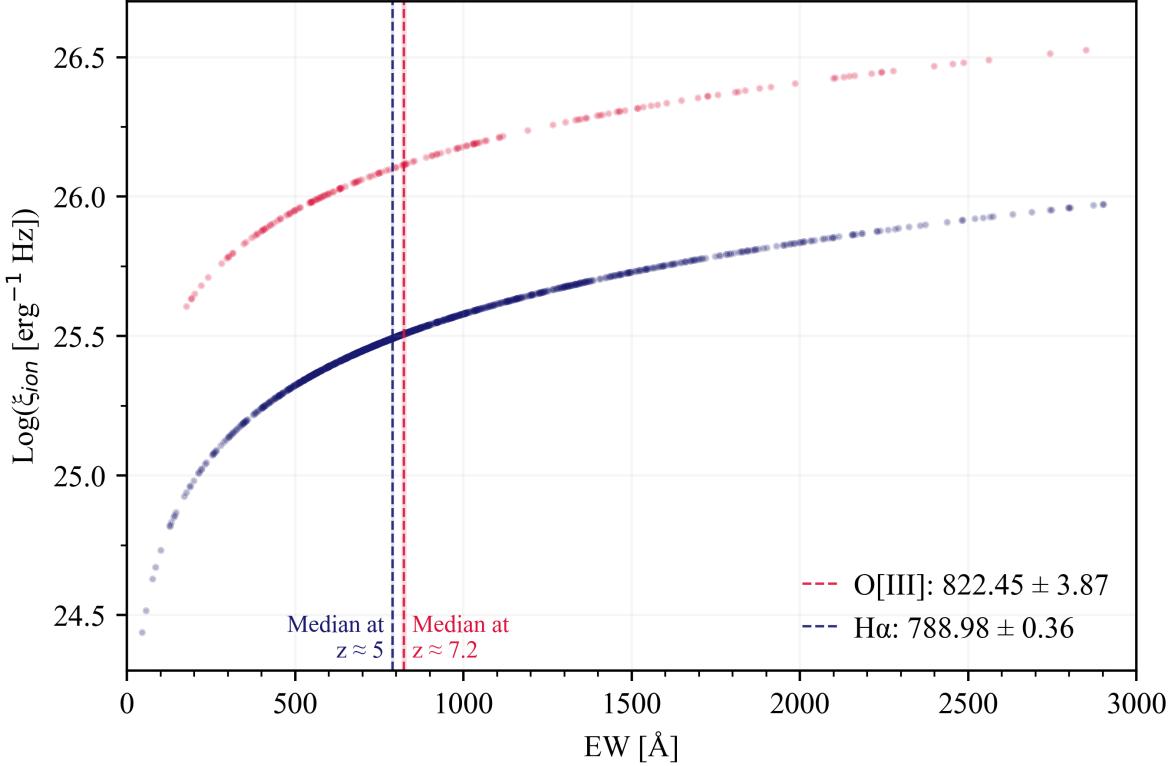
Our EW values can tell us about the ionising photon production efficiency,  $\xi_{\text{ion}}$  by using a known linear relationship known as the Calzetti UV dust attenuation law (Calzetti et al. 1994). For this analysis, the derived relationships are used, taken from Tang et al. 2019.

For O[III],

$$\log(\xi_{\text{ion}}) = 0.76 \times \log(\text{EW O[III]}) + 23.27$$

and for H $\alpha$ ,

$$\log(\xi_{\text{ion}}) = 0.85 \times \log(\text{EW H}\alpha) + 23.03$$



**Figure 6.7:** This plot shows the distribution of photon production rates for the same range of EWs shown in figure 6.6. The dotted lines on the plot are the EW medians found for each of the redshift samples. O[III] is found to have a slightly higher median than H $\alpha$ .

The distribution for the range of EWs from figure 6.6 is presented in figure 6.7. There is linear increase in the ion efficiency rate as the number of atoms available to ionise that can release a photon increases. The number of galaxies found at higher EWs and equally high

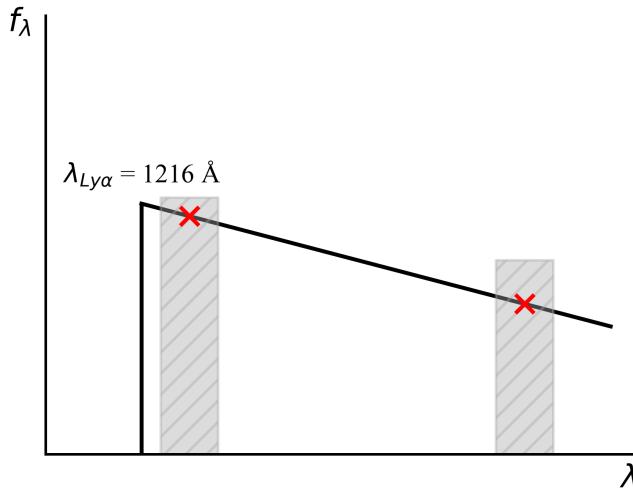
photon efficiency rates decreases. This likely means that these galaxies are just beginning their reionisation phase.

The trend follows the same patterns shown in figure 8 of Tang et al. 2019 that uses a sample of galaxies between  $z \approx 1.3 - 2.4$ . It also predicts that for galaxies at higher redshifts the same trend applies. The median for a sample of galaxies at  $z > 7$  (Stark, Ellis, et al. 2016) is higher at 900 Å for O[III]. The median found in this analysis is  $822 \pm 3.87$  Å which is fairly similar for a sample at  $z = 7.2$ . In Tang et al. 2019, sources found with O[III] EWs  $> 800$  Å, the production efficiency had a median of  $25.58 \text{ erg}^1 \text{ Hz}$ ; the median found here is roughly  $25.8 \text{ erg}^1 \text{ Hz}$ .

The findings demonstrate that the same physical mechanics that drive stellar formation in galaxies is largely similar to those galaxies in the not-so-distant nearby universe ( $z \approx 1.3 - 2.4$ ).

### 6.3 UV Spectral Slope

The UV spectral slope is determined by amount of dust within the galaxy. We expect this value to be intrinsically  $-2$  for a galaxy that has just been formed with no dust present in them. For galaxies at low redshifts, there is a correlation between the redness of the slope and how many photons escape during ionisation. This is measured by looking at the escape fraction.



**Figure 6.8:** A schematic representation of the UV spectral slope that is used to extrapolate  $\beta_{\text{UV}}$  values for a galaxy population. This feature is not very distinct but present in figure 6.1. The grey shaded regions represent filters from which the flux values would be collected. The Lyman- $\alpha$  break occurs at rest-frame wavelength of 1216 Å.

To derive the slope of the UV stellar continuum, we begin by assuming that the slope follows a power law; the flux being proportional to a factor of the wavelength.

$$f_\lambda \propto \lambda^\beta$$

Figure 6.8 shows how the spectral slopes falls as the given relation. There is a drop-off that occurs as well at the Lyman- $\alpha$  break. Therefore, to find  $\beta_{\text{UV}}$ , we need to extrapolate the flux at two points on the slope. The first being just after the Lyman break and the second at a point that is sufficiently further away from this break. Taking the ratio of the fluxes at these two points we get the following,

$$\begin{aligned}\frac{f_\lambda}{f_{\lambda_0}} &= \left(\frac{\lambda}{\lambda_0}\right)^\beta \\ f_\lambda &= f_{\lambda_0} \left(\frac{\lambda}{\lambda_0}\right)^\beta \\ f_\nu &= f_{\lambda_0} \left(\frac{\lambda}{\lambda_0}\right)^\beta \frac{c}{\nu^2} \\ f_\nu &= f_{\lambda_0} \left(\frac{\nu_0}{\nu}\right)^\beta \frac{c}{\nu^2}\end{aligned}$$

From the above we can infer that,

$$f_\nu \propto \frac{1}{\nu^\beta \nu^2} = \nu^{-(\beta+2)} \quad (5)$$

From equation 3, a relationship for the difference in magnitudes between 2 filters can be derived. This relationship is defined as follows,

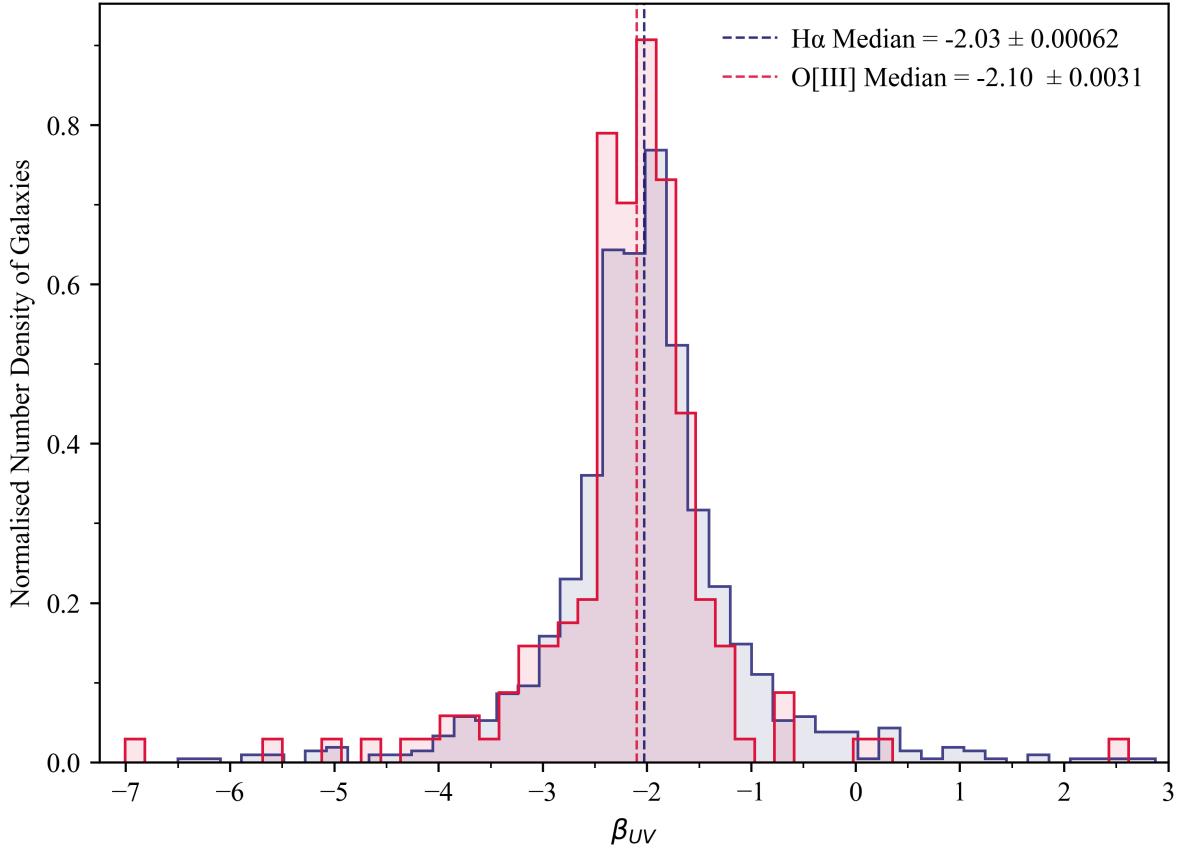
$$m_1 - m_2 = -2.5 \log\left(\frac{f_{\nu_1}}{f_{\nu_2}}\right) \quad (6)$$

Substituting equation 5 into equation 6 we arrive at an equation that allows to calculate the UV spectral slope,  $\beta_{\text{UV}}$  with magnitudes and fluxes.

$$\begin{aligned}m_1 - m_2 &= -2.5 \log\left(\frac{\nu_1}{\nu_2}\right)^{-(\beta+2)} \\ m_1 - m_2 &= -2.5 \log\left(\frac{\nu_1}{\nu_2}\right)^{-(\beta+2)} \\ m_1 - m_2 &= 2.5(\beta+2) \log\left(\frac{\lambda_2}{\lambda_1}\right) \\ \beta + 2 &= \frac{m_1 - m_2}{2.5 \log\left(\frac{\lambda_2}{\lambda_1}\right)} \\ \beta_{\text{UV}} &= \frac{m_1 - m_2}{2.5 \log\left(\frac{\lambda_2}{\lambda_1}\right)} - 2\end{aligned} \quad (7)$$

Equation 7 can be rewritten to account for dust attenuation if these factors,  $A_{\lambda_1}$  and  $A_{\lambda_2}$ , are known as intrinsic  $\beta$ ,

$$\beta_{\text{int}} = \frac{(m_1 - A_{\lambda_1}) - (m_2 - A_{\lambda_2})}{2.5 \log\left(\frac{\lambda_2}{\lambda_1}\right)} - 2$$



**Figure 6.9:** A distribution of the  $\beta$  values found using equations 8 (in blue) and 9 (in red). The median of the distributions shown as the dotted lines are found at values that were expected.  $\text{H}\alpha$  shows a sample of galaxies at  $z \approx 5$  and  $\text{O}[\text{III}]$  at  $z \approx 7$ . For galaxies at  $z \approx 7$ , the distribution is more leftward.

With reference to figure 2.7, we will chose two filters to collect flux values for as shown in the schematic diagram (figure 6.8). For a galactic sample at  $z \approx 5$  where the observed emission is  $\text{H}\alpha$ , the Lyman break is observed in a filter just before F090W. Hence, the  $\beta_{\text{UV}}$  values are found from the flux extrapolated in the 090W and 150W filters. Using the magnitudes of the galaxies in these filters and the mean wavelengths from table 2.1,

$$\beta_{\text{UV}}(\text{H}\alpha) = \frac{m_{090\text{W}} - m_{150\text{W}}}{2.5 \log\left(\frac{\lambda_{\text{eff}, 150\text{W}}}{\lambda_{\text{eff}, 090\text{W}}}\right)} - 2 \quad (8)$$

And for a sample at  $z \approx 7$ , the filter used for this are 115W and 200W filters, giving the following relation,

$$\beta_{\text{UV}}(\text{O}[\text{III}]) = \frac{m_{115\text{W}} - m_{200\text{W}}}{2.5 \log\left(\frac{\lambda_{\text{eff}, 200\text{W}}}{\lambda_{\text{eff}, 115\text{W}}}\right)} - 2 \quad (9)$$

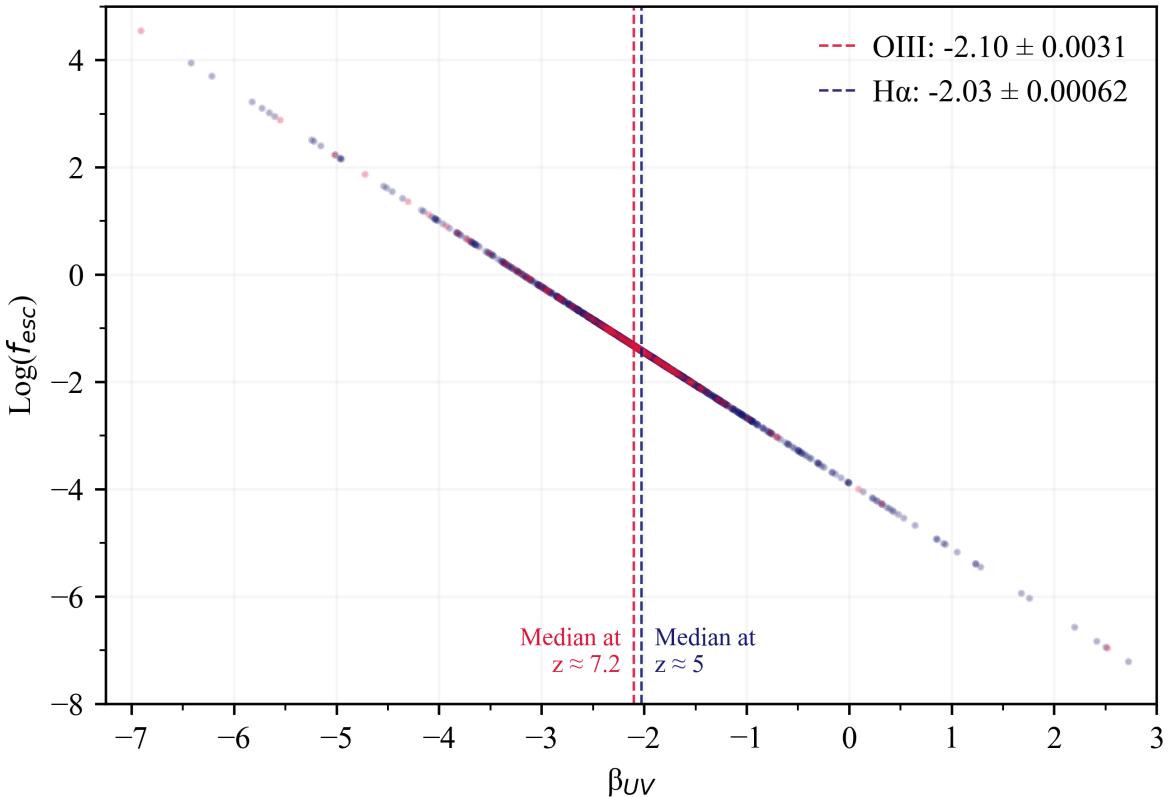
Although the sample of galaxies observed at  $z \approx 7$  is lower than the sample at  $z \approx 5$ , figure 6.9 shows that the distribution corresponding to the former, peaks at  $\beta_{\text{UV}} = -2.1$ . In a “dust-free” scenario, we had expected a gaussian distribution that was centred at  $-2$ . The

$z \approx 7$  sample is redder, possibly due to this sample consisting of more dust than the sample at  $z \approx 5$ .

### 6.3.1 Photon Escape Fraction

The estimate on the escape fraction is evaluated using the relationships derived and used in Chisholm et al. 2022,

$$f_{\text{esc}} = 1.3 \times 10^{-4} \times 10^{-1.22\beta_{\text{UV}}} \quad (10)$$

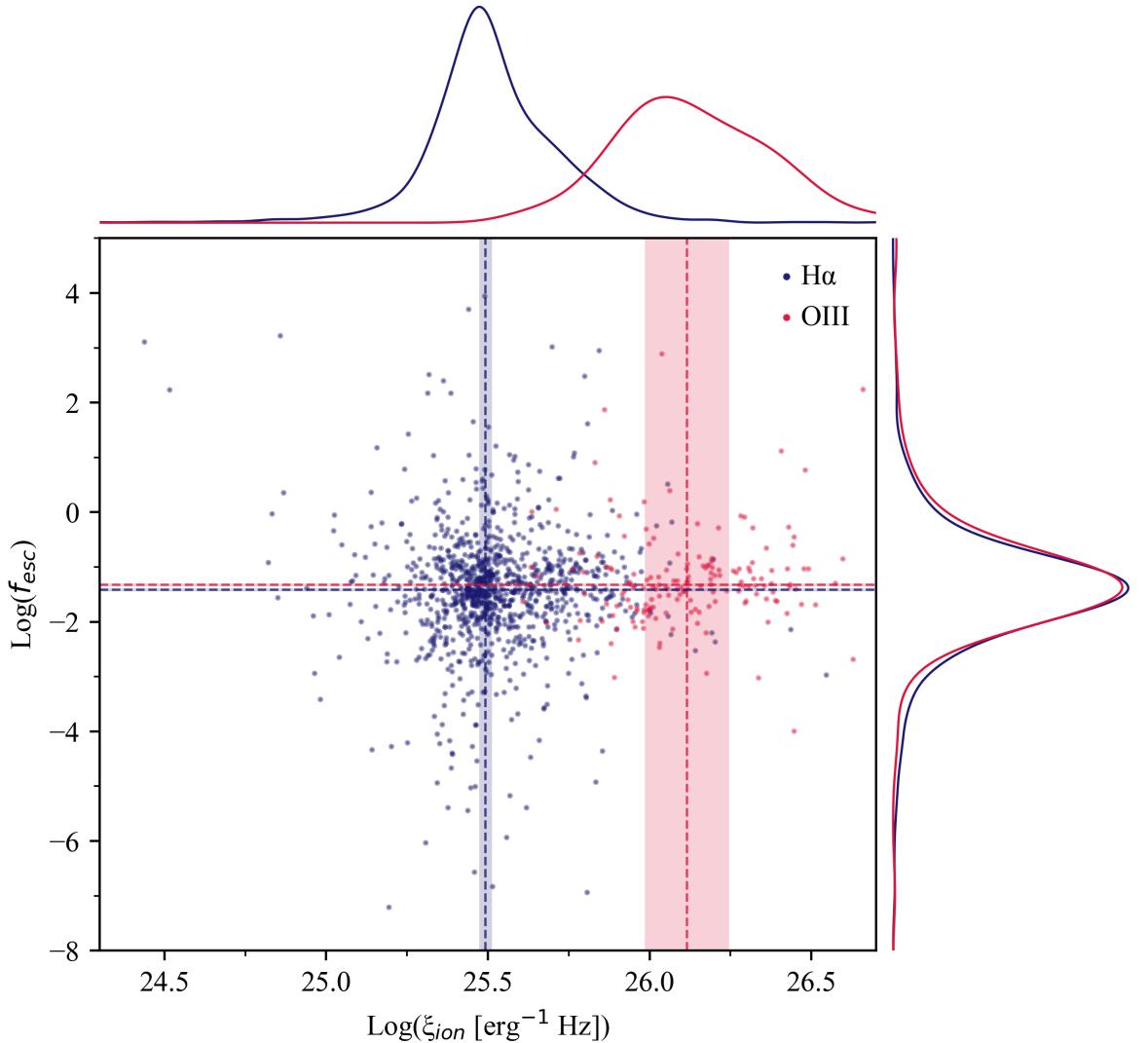


**Figure 6.10:** Plotted on the  $y$ -axis are the logarithmic values that were evaluated with equation 10. The plot shows the behaviour between photon escape fraction,  $f_{\text{esc}}$  and the slope of the UV spectrum,  $\beta_{\text{UV}}$  that are modelled from EAzY. The  $\beta_{\text{UV}}$  values are extrapolated as described by figure 6.8.

The values found in figure 6.10, using equation 10 are shown against the same range of  $\beta_{\text{UV}}$  values in figure 6.9. There is clear logarithmic relation that describes how the escape fraction changes with increasing  $\beta_{\text{UV}}$ . A higher number of photons are allowed to escape as the spectral slope becomes more negative—when the galaxy's spectrum is redder in comparison. There are hardly any blue galaxies at  $z \approx 7$  but these would likely have an very low escape fraction based on the established trend that would not vary if the galaxy's spectrum

were to get bluer. The magnitudes of the values in figure 6.10 compare well with literature values in Chisholm et al. 2022.

## 6.4 Understanding Photon Ionisation



**Figure 6.11:** The axes on this plot are constrained to the same ranges that we have previously seen figures 6.7 and 6.10. Also, as in previous figures, the  $\text{H}\alpha$  corresponds to a galactic sample at  $z \approx 5$  while  $\text{O[III]}$  being at  $z \approx 7$ . On the top and right axes, a smooth-line histogram are plotted to better visualise the properties of these samples better.

In figure 6.11 we have gathered the findings for the ionised photon escape fraction,  $f_{\text{esc}}$  and the photon ionisation production efficiency,  $\xi_{\text{ion}}$ . The scatter for the  $\xi_{\text{ion}}$  of galaxies seen at  $z \approx 7$  is much broader than at  $z \approx 5$ . This may be due to there being fewer sources detected as we arrive at higher redshifts. There is a larger error on the  $\xi_{\text{ion}}$  for  $\text{O[III]}$ , as well due to this reason. The  $f_{\text{esc}}$  values do not vary greatly for both of the samples. This suggests

that despite having a different rate of photon ionisation, the number of photons that are allowed to escape during ionisation does not change.

Most of the sources are clustered closer to the median values found for each of the samples. These are,

H $\alpha$ : (at  $z \approx 5$ )

- $f_{esc} = 3.88 \times 10^{-2} \pm 6.94 \times 10^{-7}$
- $\xi_{ion} = 25.49 \pm 0.02 \text{ erg}^{-1} \text{ Hz}$

O[III]: (at  $z \approx 7$ )

- $f_{esc} = 4.74 \times 10^{-2} \pm 6.89 \times 10^{-7}$
- $\xi_{ion} = 26.12 \pm 0.13 \text{ erg}^{-1} \text{ Hz}$

These samples show roughly the same  $\xi_{ion}$  as the deep rest-frame far-ultraviolet (FUV) spectroscopy (Stark, Walth, et al. 2015; Mainali et al. 2018; Tang et al. 2019; Hutchison et al. 2019) and rest frame optical observations (Bouwens, Smit, et al. 2016; De Barros et al. 2019; Endsley et al. 2021) at a range between of (24.5–26.7)  $\text{erg}^{-1} \text{ Hz}$ . These findings compare well with the literature range of  $\log(\xi_{ion})$  (25.3–26)  $\text{erg}^{-1} \text{ Hz}$ .

## 7 Conclusion

### 7.1 Summary

The sources found at high redshifts after fitting with EAzY were indeed galaxies at high redshifts, verified by the spectroscopic sample in figure 6.2. Majority of our galactic sample have  $dz \leq 0.15$  with minimal outliers and lie on a linear line that goes through 0. The sample does have some catastrophic outliers which were expected from the photometric redshift fitting.

The distribution of EWs in the sample are skewed gaussians centred at  $\approx 800 \text{ \AA}$ . This indicates that there is a population of extreme emission line galaxies that are undergoing reionisation at  $z \approx 5$  and  $z \approx 7$ . The number of galaxies that emit H $\alpha$  are higher telling us that these galaxies could be undergoing reionisation.

The distribution of the UV spectral slopes,  $\beta_{UV}$  show fairly smooth gaussians centred at  $-2$ , as expected. The distributions have been evaluated for “dust-free” galaxies. Majority of the galaxies at  $z \approx 7$  have redder UV stellar continuum. There are a higher number of galaxies at  $\beta \approx -2$  for the sample at  $z \approx 7$ . This could likely imply that galaxies in this sample has more dust than in the sample at  $z \approx 5$ .

The ionised photon escape fraction,  $f_{\text{esc}}$  is correlated with  $\beta_{\text{UV}}$ , as is the photon ionisation production efficiency,  $\xi_{\text{ion}}$  with EW. As the spectra of the galaxy becomes bluer, the lower the photon escape fraction. The more intense the emission, the higher the number of ionised photons observed. Galaxies at higher redshifts have higher photon ionisation production efficiency.

## 7.2 Areas for Improvement

Some of the aspects that could have been looked at given more time would be:

- To understand the properties that lacked sources at higher redshifts, literature source values could have been used to support the analysis and complete the picture at higher redshifts.
- The ZPs set for EAzY could have been amended to accommodate the program to better fit sources that were seen to have high spectroscopic redshifts than low photometric redshift estimates. EAzY could have also underestimated the redshifts of these sources due to the kinds of templates set for the program. Exploring a different set of templates could have a meaningful impact on the estimates.
- The specific star formation rate (sSFR) could have been examined using the EW of H $\alpha$  to understand a different aspect of these sample of galaxies during the epoch of cosmic reionisation.

## References

- Astropy Collaboration, A. M. Price-Whelan, B. M. Sipcz, et al. “The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package”. In: The Astronomical Journal 156.3, 123 (Sept. 2018), p. 123. DOI: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f). arXiv: [1801.02634 \[astro-ph.IM\]](https://arxiv.org/abs/1801.02634).
- Astropy Collaboration, A. M. Price-Whelan, P. L. Lim, et al. “The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package”. In: The Astrophysical Journal 935.2, 167 (Aug. 2022), p. 167. DOI: [10.3847/1538-4357/ac7c74](https://doi.org/10.3847/1538-4357/ac7c74). arXiv: [2206.14220 \[astro-ph.IM\]](https://arxiv.org/abs/2206.14220).
- Astropy Collaboration, T. P. Robitaille, et al. “Astropy: A community Python package for astronomy”. In: Astronomy & Astrophysics 558, A33 (Oct. 2013), A33. DOI: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068). arXiv: [1307.6212 \[astro-ph.IM\]](https://arxiv.org/abs/1307.6212).
- Barbary, K. “SEP: Source Extractor as a library”. In: *Journal of Open Source Software* 1.6 (2016), p. 58. DOI: [10.21105/joss.00058](https://doi.org/10.21105/joss.00058). URL: <https://doi.org/10.21105/joss.00058>.
- Beckwith, S. V. W. et al. “The Hubble Ultra Deep Field”. In: The Astronomical Journal 132.5 (Nov. 2006), pp. 1729–1755. DOI: [10.1086/507302](https://doi.org/10.1086/507302). arXiv: [astro-ph/0607632 \[astro-ph\]](https://arxiv.org/abs/astro-ph/0607632).
- Bertin, E. and S. Arnouts. “SExtractor: Software for source extraction.” In: *Astronomy & Astrophysics Supplement Series* 117 (June 1996), pp. 393–404. DOI: [10.1051/aas:1996164](https://doi.org/10.1051/aas:1996164).
- Bouwens, R. J., G. D. Illingworth, et al. “A candidate redshift  $z \sim 10$  galaxy and rapid changes in that population at an age of 500Myr”. In: Nature 469.7331 (Jan. 2011), pp. 504–507. DOI: [10.1038/nature09717](https://doi.org/10.1038/nature09717). arXiv: [0912.4263 \[astro-ph.CO\]](https://arxiv.org/abs/0912.4263).
- Bouwens, R. J., R. Smit, et al. “The Lyman-Continuum Photon Production Efficiency  $\xi_{ion}$  of  $z \sim 4\text{-}5$  Galaxies from IRAC-based H $\alpha$  Measurements: Implications for the Escape Fraction and Cosmic Reionization”. In: The Astrophysical Journal 831.2, 176 (Nov. 2016), p. 176. DOI: [10.3847/0004-637X/831/2/176](https://doi.org/10.3847/0004-637X/831/2/176). arXiv: [1511.08504 \[astro-ph.GA\]](https://arxiv.org/abs/1511.08504).
- Brammer, G. B., P. G. van Dokkum, and P. Coppi. “EAZY: A Fast, Public Photometric Redshift Code”. In: The Astrophysical Journal 686.2 (Oct. 2008), pp. 1503–1513. DOI: [10.1086/591786](https://doi.org/10.1086/591786). arXiv: [0807.1533 \[astro-ph\]](https://arxiv.org/abs/0807.1533).
- Brammer, Gabriel. *EAZY Photo-Z GitHub Repository*. <https://github.com/gbrammer/eazy-photoz/?tab=readme-ov-file>. [Accessed on 20-03-2024]. 2015.
- Calzetti, D., A. L. Kinney, and T. Storchi-Bergmann. “Dust Extinction of the Stellar Continua in Starburst Galaxies: The Ultraviolet and Optical Extinction Law”. In: The Astrophysical Journal 429 (July 1994), p. 582. DOI: [10.1086/174346](https://doi.org/10.1086/174346).
- Centre for Astrophysics and Supercomputing. *Lookback Time*. URL: <https://astronomy.swin.edu.au/cosmos/L/Lookback+Time>.
- Chisholm, J. et al. “The far-ultraviolet continuum slope as a Lyman Continuum escape estimator at high redshift”. In: Monthly Notices of the Royal Astronomical Society 517.4 (Dec. 2022), pp. 5104–5120. DOI: [10.1093/mnras/stac2874](https://doi.org/10.1093/mnras/stac2874). arXiv: [2207.05771 \[astro-ph.GA\]](https://arxiv.org/abs/2207.05771).
- De Barros, S. et al. “The GREATS H $\alpha$ +[Oiii] luminosity function and galaxy properties at  $z = 8$ : walking the way of JWST”. In: *Monthly Notices of the Royal Astronomical Society* 489.2 (Apr. 2019), pp. 2355–2366. ISSN: 0035-8711. DOI: [10.1093/mnras/stz940](https://doi.org/10.1093/mnras/stz940).

- eprint: <https://academic.oup.com/mnras/article-pdf/489/2/2355/29592286/stz940.pdf>. URL: <https://doi.org/10.1093/mnras/stz940>.
- Doppler, C. *Beitrage zur fixsternenkunde*. 1846.
- Draper, P. *Gaia Data Release 1 (Gaia DR1)*. <http://star-www.dur.ac.uk/~pdraper/gaia/gaia.html>. [Accessed on 15-03-2024]. 2016.
- East Asian Observatory. *Starlink Software*. <https://starlink.eao.hawaii.edu/starlink/2023A>. [Accessed on 15-03-2024]. 2023.
- Eisenstein, D. J. et al. “Overview of the JWST Advanced Deep Extragalactic Survey (JADES)”. In: *arXiv e-prints*, arXiv:2306.02465 (June 2023), arXiv:2306.02465. DOI: [10.48550/arXiv.2306.02465](https://doi.org/10.48550/arXiv.2306.02465). arXiv: [2306.02465 \[astro-ph.GA\]](https://arxiv.org/abs/2306.02465).
- Endsley, R. et al. “The [O III]+H  $\beta$  equivalent width distribution at z > 7: implications for the contribution of galaxies to reionization”. In: *Monthly Notices of the Royal Astronomical Society* 500.4 (Jan. 2021), pp. 5229–5248. DOI: [10.1093/mnras/staa3370](https://doi.org/10.1093/mnras/staa3370). arXiv: [2005.02402 \[astro-ph.GA\]](https://arxiv.org/abs/2005.02402).
- ESA. *ALMA Observes Most Distant Oxygen Ever*. [Accessed: 30-01-2024]. European Space Agency, 2016. URL: <https://www.eso.org/public/news/eso1620/>.
- . *Webb: Galaxy mergers solve early Universe mystery*. [Accessed on 27-01-2024]. European Space Agency, 2024. URL: [https://www.esa.int/Science\\_Exploration/Space\\_Science/Webb/Webb\\_galaxy\\_mergers\\_solve\\_early\\_Universe\\_mystery](https://www.esa.int/Science_Exploration/Space_Science/Webb/Webb_galaxy_mergers_solve_early_Universe_mystery).
- Förster Schreiber, N. M. and S. Wuyts. “Star-Forming Galaxies at Cosmic Noon”. In: *Annual Review of Astronomy & Astrophysics* 58 (Aug. 2020), pp. 661–725. DOI: [10.1146/annurev-astro-032620-021910](https://doi.org/10.1146/annurev-astro-032620-021910). arXiv: [2010.10171 \[astro-ph.GA\]](https://arxiv.org/abs/2010.10171).
- Hutchison, T. A. et al. “Near-infrared Spectroscopy of Galaxies During Reionization: Measuring C III] in a Galaxy at z = 7.5”. In: *The Astrophysical Journal* 879.2, 70 (July 2019), p. 70. DOI: [10.3847/1538-4357/ab22a2](https://doi.org/10.3847/1538-4357/ab22a2). arXiv: [1905.08812 \[astro-ph.GA\]](https://arxiv.org/abs/1905.08812).
- Kennicutt Robert C., J. “The Integrated Spectra of Nearby Galaxies: General Properties and Emission-Line Spectra”. In: *The Astrophysical Journal* 388 (Apr. 1992), p. 310. DOI: [10.1086/171154](https://doi.org/10.1086/171154).
- Mainali, R. et al. “Spectroscopic constraints on UV metal line emission at z > 69: the nature of Ly emitting galaxies in the reionization era”. In: *Monthly Notices of the Royal Astronomical Society* 479.1 (June 2018), pp. 1180–1193. ISSN: 0035-8711. DOI: [10.1093/mnras/sty1640](https://doi.org/10.1093/mnras/sty1640). eprint: <https://academic.oup.com/mnras/article-pdf/479/1/1180/25129193/sty1640.pdf>. URL: <https://doi.org/10.1093/mnras/sty1640>.
- McElwain, M. W. et al. “The James Webb Space Telescope Mission: Optical Telescope Element Design, Development, and Performance”. In: *Publications of the Astronomical Society of the Pacific* 135.1047, 058001 (May 2023), p. 058001. DOI: [10.1088/1538-3873/acada0](https://doi.org/10.1088/1538-3873/acada0). arXiv: [2301.01779 \[astro-ph.IM\]](https://arxiv.org/abs/2301.01779).
- Naidu, R. P. et al. “Rapid Reionization by the Oligarchs: The Case for Massive, UV-bright, Star-forming Galaxies with High Escape Fractions”. In: *The Astrophysical Journal* 892.2, 109 (Apr. 2020), p. 109. DOI: [10.3847/1538-4357/ab7cc9](https://doi.org/10.3847/1538-4357/ab7cc9). arXiv: [1907.13130 \[astro-ph.GA\]](https://arxiv.org/abs/1907.13130).
- NASA James Webb Space Telescope. *Image of Planetary Rings*. <https://webbtelescope.org/contents/media/images/01GKT0RRJBP5ZMJRCQNPT8SXP?tag=planetary%20rings>. [Accessed on 15-03-2024].
- Oesch, P. et al. *FRESCO: The First Reionization Epoch Spectroscopic CComplete Survey*. JWST Proposal. Cycle 1, ID. #1895. Mar. 2021.

- Planck Collaboration et al. “Planck 2018 results. VI. Cosmological parameters”. In: *Astronomy & Astrophysics* 641, A6 (Sept. 2020), A6. DOI: [10.1051/0004-6361/201833910](https://doi.org/10.1051/0004-6361/201833910). arXiv: [1807.06209 \[astro-ph.CO\]](https://arxiv.org/abs/1807.06209).

Python Software Foundation. *Python multiprocessing Documentation*. [Accessed: 28-02-2024].

PyTorch Developers. *PyTorch GitHub Repository*. [Accessed: 28-02-2024].

Rieke, M. J. et al. “JADES Initial Data Release for the Hubble Ultra Deep Field: Revealing the Faint Infrared Sky with Deep JWST NIRCam Imaging”. In: *The Astrophysical Journal Supplement Series* 269.1, 16 (Nov. 2023), p. 16. DOI: [10.3847/1538-4365/acf44d](https://doi.org/10.3847/1538-4365/acf44d). arXiv: [2306.02466 \[astro-ph.GA\]](https://arxiv.org/abs/2306.02466).

Rigby, J. et al. “The Science Performance of JWST as Characterized in Commissioning”. In: *Publications of the Astronomical Society of the Pacific* 135.1046, 048001 (Apr. 2023), p. 048001. DOI: [10.1088/1538-3873/acb293](https://doi.org/10.1088/1538-3873/acb293). arXiv: [2207.05632 \[astro-ph.IM\]](https://arxiv.org/abs/2207.05632).

Robertson, B. E. et al. “Cosmic Reionization and Early Star-forming Galaxies: A Joint Analysis of New Constraints from Planck and the Hubble Space Telescope”. In: *The Astrophysical Journal Letters* 802.2, L19 (Apr. 2015), p. L19. DOI: [10.1088/2041-8205/802/2/L19](https://doi.org/10.1088/2041-8205/802/2/L19). arXiv: [1502.02024 \[astro-ph.CO\]](https://arxiv.org/abs/1502.02024).

Rodrigo, C. and E. Solano. “The SVO Filter Profile Service”. In: *XIV.0 Scientific Meeting (virtual) of the Spanish Astronomical Society*. July 2020, 182, p. 182.

Rodrigo, C., E. Solano, and A. Bayo. *SVO Filter Profile Service Version 1.0*. IVOA Working Draft 15 October 2012. Oct. 2012. DOI: [10.5479/ADS/bib/2012ivoa.rept.1015R](https://doi.org/10.5479/ADS/bib/2012ivoa.rept.1015R).

Space Telescope Science Institute. *JWST Near Infrared Camera*. <https://jwst-docs.stsci.edu/jwst-near-infrared-camera>. [Accessed on 14-03-2024]. 2019.

—. *JWST Near Infrared Spectrograph*. <https://jwst-docs.stsci.edu/jwst-near-infrared-spectrograph>. [Accessed on March 14-03-2024]. 2023.

Stark, D. P. “Galaxies in the First Billion Years After the Big Bang”. In: *Annual Review of Astronomy & Astrophysics* 54 (Sept. 2016), pp. 761–803. DOI: [10.1146/annurev-astro-081915-023417](https://doi.org/10.1146/annurev-astro-081915-023417).

Stark, D. P., R. S. Ellis, et al. “[Ly and Ciii] emission in z = 79 Galaxies: accelerated reionization around luminous star-forming systems?” In: *Monthly Notices of the Royal Astronomical Society* 464.1 (Sept. 2016), pp. 469–479. ISSN: 0035-8711. DOI: [10.1093/mnras/stw2233](https://doi.org/10.1093/mnras/stw2233). eprint: <https://academic.oup.com/mnras/article-pdf/464/1/469/18512610/stw2233.pdf>. URL: <https://doi.org/10.1093/mnras/stw2233>.

Stark, D. P., G. Walth, et al. “Spectroscopic detection of Civ 1548 in a galaxy at z = 7.045: implications for the ionizing spectra of reionization-era galaxies”. In: *Monthly Notices of the Royal Astronomical Society* 454.2 (Oct. 2015), pp. 1393–1403. ISSN: 0035-8711. DOI: [10.1093/mnras/stv1907](https://doi.org/10.1093/mnras/stv1907). eprint: <https://academic.oup.com/mnras/article-pdf/454/2/1393/9377644/stv1907.pdf>. URL: <https://doi.org/10.1093/mnras/stv1907>.

Sterling, L. S. and E. Y. Shapiro. *The Art of Prolog: Advanced Programming Techniques*. Springer, 1994. DOI: [10.1007/978-3-319-21957-8](https://doi.org/10.1007/978-3-319-21957-8). URL: <https://link.springer.com/book/10.1007/978-3-319-21957-8>.

Suzuki, T. L. et al. “[OIII] emission line as a tracer of star-forming galaxies at high redshifts: comparison between H and [OIII] emitters at z=2.23 in HiZELS”. In: *Monthly Notices of the Royal Astronomical Society* 462.1 (June 2016), pp. 181–189. ISSN: 0035-8711. DOI: [10.1093/mnras/stw1655](https://doi.org/10.1093/mnras/stw1655). eprint: <https://academic.oup.com/mnras/article-pdf/462/1/181/stw1655.pdf>. URL: <https://doi.org/10.1093/mnras/stw1655>.

- [pdf/462/1/181/18471309/stw1655.pdf](https://doi.org/10.1093/mnras/stw1655.pdf). URL: <https://doi.org/10.1093/mnras/stw1655>.
- SVO Filter Profile Service*. [Accessed on 10-03-2024]. June 2019. URL: <http://svo2.cab.inta-csic.es/theory/fps/index.php?mode=browse>.
- Tang, M. et al. “MMT/MMIRS spectroscopy of  $z = 1.3 - 2.4$  extreme [O III] emitters: implications for galaxies in the reionization era”. In: Monthly Notices of the Royal Astronomical Society 489.2 (Oct. 2019), pp. 2572–2594. DOI: [10.1093/mnras/stz2236](https://doi.org/10.1093/mnras/stz2236). arXiv: [1809.09637 \[astro-ph.GA\]](https://arxiv.org/abs/1809.09637).
- The Pandas Development Team. *pandas.DataFrame*. <https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.html>. [Accessed on 20-03-2024].
- Trebitsch, M. et al. “Reionization with star-forming galaxies: insights from the Low- $z$  Lyman Continuum Survey”. In: (Dec. 2022). DOI: [10.48550/arXiv.2212.06177](https://doi.org/10.48550/arXiv.2212.06177).
- United States Navy Department. *Principles and Applications of Underwater Sound*. [Accessed on 29-03-2024]. 1969, p. 194. URL: <https://archive.org/details/PrinciplesAndApplicationsOfUnderwaterSound>.
- University of Iowa Department of Physics and Astronomy. *Astronomical Redshift*. 2024. URL: <https://itu.physics.uiowa.edu/labs/advanced/astronomical-redshift>.