THE PENNSYLVANIA STATE UNIVERSITY
MILLENNIUM SCHOLARS PROGRAM

DEPARTMENT OF ASTRONOMY AND ASTROPHYSICS

IDENTIFYING EMISSION LINE GALAXIES IN THE FENIKS PILOT SURVEY VIA SPECTRAL ENERGY DISTRIBUTION FITTING WITH BAGPIPES

VICTORIA BOEHM
SPRING 2022

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Astronomy and Astrophysics
and in Physics

Reviewed and approved* by the following:

Danilo Marchesini
Professor of Physics and Astronomy
Thesis Supervisor

Joel Leja
Assistant Professor of Astronomy and Astrophysics
Reviewer

* Signatures are on file in the Millennium Scholars Program office.

ABSTRACT

Detailed spectroscopic studies of today's massive galaxies indicate that they formed most of their stars very rapidly when the universe was very young, but then abruptly stopped their star-formation activity. Several important questions on the star-forming phase and quenching mechanisms of today's massive galaxies remain unanswered. Studying intensely star-forming galaxies in the distant universe provides a path to furthering our understanding. In this work, we present the preliminary analysis of data obtained by the FLAMINGOS-2 Extragalactic Near-Infrared *K*-band Split (FENIKS) pilot survey, which aims to identify star-forming emission line galaxies (ELGs) in the distant universe by improving spectral resolution in the *K*-band wavelengths. We utilize the BAGPIPES Python package to perform spectral energy distribution fitting of a selection of galaxies observed in the FENIKS pilot survey. The fitted models allow us to estimate emission line equivalent widths and identify distant ELGs. The stellar population properties (e.g., stellar mass, star-formation rate, dust attenuation) of the identified ELGs are characterized, with the end goal of illuminating the star-forming and quenching mechanisms of distant galaxies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

It has been a very long road to get here. I am eternally grateful for the support of my family, in whom I found the strength to keep going no matter how challenging the road became. I might not have stayed with it if you hadn't been there to keep me moving forward, and I owe you all beyond measure.

I am also grateful for the support of my academic family both here at Penn State and at other universities. I would like to thank my thesis advisor, Dr. Danilo Marchesini; your patience and understanding as I scrambled to catch up to this exciting new field of galaxy evolution was greatly appreciated, and I hope some day we may meet in person, when times improve. I also want to thank my research mentor at Penn State, Dr. Eric Hudson; you remain one of my staunchest advocates and helped transform a nervous first-year into a confident researcher. Lastly, I wish to thank Dr. Robert Harmon; I never forgot the words you told me on July 9th of 2019, and I will never again let doubt stop me from trying.

On a more technical note, I would also like to thank the other students in the Marchesini Lab Group, as well as Dr. Adam Carnall. Insightful discussions with these individuals yielded both progress and clarity in this research. This research makes use of the Tufts High Performance Compute Cluster (https://it.tufts.edu/high-performance-computing), and I thank the IT staff who maintain this valuable resource. This research was funded in part by the National Science Foundation (NSF) through grant NSF-2009442; the findings presented herein are the work, opinions, and conclusions of the author and do not necessarily reflect the views of the NSF.

I would not have made it this far without the help of the Millennium Scholars Program. Apart from years of financial and academic support that allowed me to pursue research from the moment I stepped on campus, the support of advisors Georjanne Rosa and Katelyn Jones helped me find my bearings at a time in my life when I wasn’t sure who I was. I thank you both for having the patience to listen.

Lastly, I want to thank my Finest friends in Cohort 6, to whom I owe my fondest memories of my time at Penn State. You are all incredible people – I cannot wait to see what you do next.

Chapter 1

Introduction

1.1 A Quick Introduction to the Field of Galaxy Evolution

The universe today is populated by a diverse array of galaxies. A galaxy is broadly defined as a large collection of stars orbiting a mutual center of mass and held together by the force of gravity exerted by a halo of dark matter. Such a broad definition necessarily encompasses a wide variety of objects. Galaxies span an enormous range in size, with the smallest dwarf galaxies hosting just a few thousand stars while the awe-inspiringly massive elliptical galaxies host trillions. Galaxy morphology is equally diverse; the elliptical galaxies have spherical or ellipsoidal shapes, while the spiral galaxies are shaped like flattened disks bearing bulging centers and elaborate spiral arms studded with star-forming nebulae. There are also the irregular galaxies, whose bizarre and difficult-to-classify shapes result from perturbations like gravitational interactions with other galaxies or accretion of material from the intergalactic medium (IGM). It is believed that most galaxies host a supermassive black hole at their centers. Galaxies can also be classified by their current growth rate; galaxies that are still growing by forming new stars are appropriately referred to as star-forming galaxies, while galaxies that have effectively stopped forming new stars are called quenched or quiescent galaxies.

Of course, it was not always this way. In the earliest days of the universe, 13.8 billion years ago, the universe was a mostly homogeneous mixture of hot plasma. One of the important outstanding questions in astronomy is the question of how the diverse population of galaxies we see today evolved out of this primordial plasma. The cosmic microwave background radiation (CMBR), a pattern of microwave light that spans the entire visible sky and the oldest light that can be observed, has given us some clues as to how the process began; irregularities in the CMBR suggest that rather than being perfectly uniform, some parts of the primordial plasma were denser than others, and these dense regions formed the seeds from which galaxy centers grew. But while the process of forming galaxy centers and dark matter halos is a straightforward picture of mass and gravitational attraction, the physics of forming the stars in the galaxy is much more complex. A complete picture of a galaxy's star forming processes must describe how the gas and dust in a nebula collapses to form stars, how heating from newly-formed stars disperses nebulae and inhibits further star formation, how gravitational interactions with other galaxies can promote bursts of intense star formation or disturb star-forming nebulae, and how the supermassive black hole that resides at most galaxy centers can shut down star formation by blowing all of the galaxy's gas away. Star formation is an intricate dance of many internal processes, external processes, and feedback loops, and the complete story is not presently within our grasp.

Research in the field of galaxy evolution seeks to answer the question of how the early primordial plasma of the universe, as revealed to us by the CMBR, evolved to form today's diverse and fascinating population of galaxies. Because galaxy evolution is a slow process spanning billions of years, it is impossible to simply wait and watch the process unfold. However, the speed of light is not infinite, but is instead about 9.5 trillion kilometers per year in vacuum; this distance travelled by light in one year is known as a light-year. If one looks at galaxies that are billions of light-years away, one will see the galaxies as they were billions of years ago. Examining many galaxies at varying distances from us offers glimpses of galaxies at different ages in the universe's history. Rather than observe one galaxy evolve over cosmic time, we instead observe snapshots of galaxies in different stages of their evolution at different times in cosmic history. By connecting these snapshots and interpreting them with physical theories, we attempt to piece together the complete story of how today's massive and diverse galaxies grew from the primordial plasma of the early universe.

Throughout this thesis, we adopt the astronomical convention of describing galaxy distances and ages with the redshift z. Because the universe is expanding, the light that we observe coming from distant galaxies is redshifted, or stretched out to longer wavelengths. The further away the galaxy is, the more redshifted it will be, corresponding to a higher z. Because of the finite speed of light, a greater distance also indicates a greater lookback time and thus a younger universe. Our current time and position in space are denoted by $z = 0$, so at $z = 0$, the universe is 13.8 billion years (Gyr) old. If we adopt the standard ΛCDM model of the universe, then $z = 1$ corresponds to a universe that is 6.1 Gyr old, while $z = 2$ means the universe is about 3.5 Gyr old. $z = 3$ and $z = 4$ respectively denote ages of about 2.3 Gyr and 1.5 Gyr. The most distant light we can see is the CMBR at $z \sim 1100$, which corresponds to a universe that is just 380,000 years old.

1.2 Outstanding Questions in Galaxy Evolution and the Motivation for this Project

Over the past decade, several deep-sky galaxy surveys such as NMBS (Marchesini et al. 2010), selections from the fields surveyed by COSMOS/UltraVISTA (Muzzin et al. 2013), and ZFOURGE (Straatman et al. 2016) have pointed to the existence of a large population of massive galaxies at redshifts $3.0 \leq z \leq 4.0$, corresponding to universe ages of roughly 1.5 Gyr $\leq t \leq$ 2.3 Gyr. This suggests that in the early universe, galaxies formed their stellar populations extremely rapidly and efficiently. Detailed spectroscopic follow-up studies in this redshift range also suggest that some of these massive galaxies had already quenched, or ceased rapid star formation, as early as redshifts $z \sim 3.5$ (Forrest et al. 2020; Saracco et al. 2020). The quenching mechanisms of these massive, quiescent, $z > 3.0$ galaxies must have been very efficient; in fact, no currently-proposed quenching mechanism is rapid enough to account for the existence of these galaxies. The observations so far indicate that in the early universe, galaxy star-forming and quenching mechanisms operated far faster than astronomers previously believed.

Understanding the discrepancies between theoretical predictions and actual observations has several important consequences for the field. A model's validity can be weighed in part by how accurate its predictions are. If our current models of galaxy evolution do not accurately predict what we observe, then this indicates that our understanding of galaxy evolution is incomplete or flawed. Galaxy evolution is a broad term that encompasses many different physical processes and astrophysical disciplines. To paint an accurate picture of how galaxies evolve requires understanding how galaxies accrete gas from the IGM, how gravitational interactions and collisions with other galaxies affect the galaxy's own internal structure, how feedback from both its stars and its central supermassive black hole can change the course of a galaxy's growth, how dust is formed within a galaxy and how it can be lost or destroyed, and more. It's quite conceivable that there are gaps in our knowledge of at least one of those topics, and finding those gaps would open up new questions for both physicists and astronomers to explore.

To understand the discrepancies between the theories and the observations, a robust picture of what really happened in the universe at redshifts $z > 3$ is needed. One path towards building this picture is to conduct a population study of both quenched and intensely star-forming galaxies at $z > 3$, with the end goal of identifying the respective quenching and star-forming mechanisms of these galaxies. This was one of the many goals of the FLAMINGOS-2 Extragalactic Near-Infrared

K-band Split (FENIKS) pilot survey (Esdaile et al. 2021), an international collaboration between American, Canadian, European, and Australian universities utilizing the FLAMINGOS-2 instrument onboard the Gemini South Telescope, located in the Chilean Andes and owned by the National Science Foundation's NOIRLab. The FENIKS pilot survey was carried out from 2017 to 2019 and served as a small-scale proof of concept for the newly developed K-split filters ($K_{\rm blue}$ and $K_{\rm red}$) which were developed to improve spectral resolution in the K-band wavelengths. By improving spectral resolution in the K-band, researchers would be able to more robustly identify both quenched galaxies and star-forming emission line galaxies (ELGs) at redshifts $z > 3$. This is a redshift regime not yet effectively targeted by other surveys such as NMBS and ZFOURGE; the data collected by the FENIKS survey offers a wealth of new knowledge about $z > 3$ galaxies that will be analyzed for years to come.

To constrain the scope of this project, we choose to focus here on the active star-forming ELGs, and leave the topic of the quiescent galaxies to future studies. ELGs are distinct from other galaxies because their spectra contain strong emission line features that may originate from nebular gases being excited by newly-formed stars. Our long-term goal for this research is to determine how galaxies in the early universe grew so quickly and why they suddenly stopped; by examining ELGs in detail, we can work towards an understanding of the first half of this question. We combine the newly-obtained FENIKS data, which effectively targets ELGs at $2.0 < z < 6.0$, with data from previous missions like NMBS and ZFOURGE, which employed medium band filters in the J and H wavelength bands to effectively target ELGs at $0.0 < z < 4.0$. The combined dataset represents a higher-resolution and more complete sample of galaxies in the redshift range $0.0 < z < 6.0$. We analyze this dataset to identify ELGs that we will later target in more detailed follow-up spectroscopic studies. These follow-up studies will be used to infer the star-forming mechanisms of massive early ELGs and thus help to answer the first part of our overarching research question.

1.3 Outline for this Thesis

In this thesis, we present the preliminary analysis of a dataset combining previous medium band galaxy surveys with the newly-obtained data from the FENIKS pilot survey, with the goal of identifying star-forming ELGs for future spectroscopic follow-up studies. Chapter 1 has presented the motivations for this analysis. Chapter 2 will provide the needed theoretical background, and Chapter 3 will detail the data being analyzed as well as the analysis methods being applied. Our analysis utilizes the Python programming language, in particular a recently-developed Python package called Bayesian Analysis of Galaxies for Physical Inference and Parameter EStimation (BAGPIPES) (Carnall et al. 2018). With this Python package, we are able to infer the emission line strengths of galaxies imaged by the FENIKS pilot survey, which allows us to identify star-forming ELGs. Chapter 4 presents the analysis itself, concluding with the total number of ELGs detected by the FENIKS pilot survey. Chapter 5 concludes with discussion of the results and some preliminary analysis of the stellar population properties of the identified ELGs, as well as discussion of plans for more detailed characterization of the identified ELGs with the end goal of illuminating the star-forming and quenching mechanisms of these distant galaxies.

Chapter 2

Background

2.1 Characterizing Galaxies through Spectral Energy Distribution Fitting

Discerning star-forming ELGs from quenched galaxies is not a trivial task. At cosmic distances, the individual stars and nebulae of galaxies cannot be resolved, and images of ELGs may look superficially similar to quenched galaxies. Additionally, galaxy evolution is a slow process that cannot be observed on human timescales; re-imaging the galaxies even decades apart would not reveal significant changes that could separate ELGs from quiescent galaxies. Because by-eye distinctions cannot be made for high-z galaxies, alternative approaches must be used. A popular approach to this problem is spectral energy distribution (SED) modelling (Conroy 2013). The SED of a galaxy describes how luminous the galaxy is as a function of the output wavelength of light. Many of the galaxy's interesting physical properties, such as its star formation rate (SFR), star formation history (SFH), stellar mass, dustiness, and age have a direct impact on the shape of the SED. Recovering these physical properties from the SED is the mission of stellar population synthesis (SPS), the technique from which many SED modelling programs are developed.

The idea behind SPS is conceptually straightforward. We consider galaxies to be collections of different stellar populations, each of which have their own total masses, ages, initial mass functions, and metallicities. Each stellar population, being made of different types of stars, will generate a different spectrum; younger populations will be more luminous and emit more light in blue, violet, and ultraviolet (UV) wavelengths, while older populations will be dimmer and emit more of their light in red and infrared (IR) wavelengths. SED modelling combines stellar population spectra derived from SPS theories with spectra from nebular gases and the interstellar medium (ISM); excited nebular gases contribute light to the SED at multiple discrete wavelengths as well as continuously across certain wavelength bands, while the dusty ISM absorbs stellar UV photons and re-emits them as IR photons. The final SED is thus the sum of the SPS-derived stellar spectra, the excited nebular gas emission, and the ISM-reddened stellar light. The advantage of SED modelling is that because we know the stellar populations input into the model, we immediately know many of the interesting physical properties of the galaxy modelled by the SED.

The technique of SED modelling to characterize real galaxies works as follows. A galaxy is observed, either via spectroscopy over a continuous band of wavelengths, or via photometry in multiple wavelength bands. SED modelling programs are used to generate hypothetical SEDs based on different combinations of SPS-derived stellar populations mixed with varying levels of nebular emission and ISM reddening. The hypothetical SEDs are compared to the observations of the real spectrum/photometry to see how well the hypothetical picture matches reality. Due to limitations in the precision of the data as well as degeneracies in the modelled SED, we cannot ever obtain a perfectly fit SED that we can claim with full confidence is accurate to reality. However, we can recover a range of SEDs, and thus a range of physical property estimates, that could describe the observed galaxy. Therefore, by fitting SEDs to observed galaxy spectra or photometry, we can estimate the galaxy's physical properties: e.g., how old the galaxy is, how much mass it contains, and whether or not the galaxy is actively forming stars at an appreciable rate. The reliability of these estimates depends on part in the quality of the SED modelling program used, and in part on the quality of the input data.

2.2 Targeting ELGs and the Need for Better Spectral Resolution

The SEDs of intensely star-forming galaxies are distinguished by the presence of strong emission lines, which originate from nebular gases being excited by newly-formed stars. The subsequent de-excitation of the nebular gas emits light at discrete wavelengths, which gives emission lines very thin spectral profiles. Thus, a star-forming ELG could be identified by examination of its modelled SED for the presence of emission lines.

Ideally, as ELGs are best identified by their SEDs, one would collect high-resolution spectra of galaxies by observing them through a spectrograph, thus yielding the best-constrained model SEDs to describe them. In practice, this is not feasible; a spectrograph necessarily must disperse the incoming light and reduce the quality of the already-faint signal, lowering the signal-to-noise ratio (SNR) of the observations to the point where robust characterization of the galaxies would become impossible.

In lieu of obtaining spectra, astronomers instead obtain photometric observations of the galaxies in multiple filters, discretely sampling the spectrum at multiple wavelength bands, with each band having its own bandwidth. A model SED is then fitted to the photometric data points, with the robustness and accuracy of the model SED dependent in part on the widths of the filters used in the observation. In selecting filter widths, one must deal with competing tradeoffs. The use of broadband filters which cover large spectral ranges allows for more light to be collected, improving the SNR of the observations. However, broadband filters have low spectral resolution, making fine spectral features such as emission lines difficult to detect. On the other hand, one could use medium band filters, which sample a smaller range of wavelengths and thus collect less light (thereby risking lower SNR or else demanding longer exposure times), but are better at resolving emission lines, enabling more robust identification of ELGs.

2.3 Dividing the *K*-band with Medium Band Filters

Because of the observing challenges associated with using medium band filters, the choice of which wavelength bands should receive the medium band treatment is an important one. After all, it would be inefficient to conduct medium band surveys in wavelength bands that have no emission lines or sharp continuum features to resolve in the first place. If one is looking for ELGs at redshifts $z > 3$, then one should select bands that contain emission lines at those redshifts. This goal is part of what motivated the design of the FENIKS pilot survey, which aimed to improve spectral resolution in the *K*-band wavelengths. The conventional broadband filter for the *K*-band, the K_s filter, is centered on $\lambda_{cen} = 2.2\ \mu$m and covers the range from 1.9 μm to 2.5 μm for a total width of about $\Delta\lambda = 0.6\ \mu$m. The FENIKS *K*-split filters break the broadband K_s filter into two medium band filters, K_{blue} ($\lambda_{cen} = 2.06\ \mu$m, $\Delta\lambda = 0.25\ \mu$m) and K_{red} ($\lambda_{cen} = 2.31\ \mu$m, $\Delta\lambda = 0.27\ \mu$m). These wavelength bands were chosen in part because at redshifts $2 < z < 5$, these bands sample several important emission lines, such as the [SII] doublet, the [NII] doublet, the Balmer Hα and Hβ lines, and the [OIII] doublet. The proximity and relative narrowness of the K_{blue} and K_{red} bands allows these thin emission lines to be sampled by one filter without simultaneously being sampled by the other; the resultant "boosting" of one filter's flux and not the other's is a robust indication of an emission line.

2.4 A Brief Diversion to Address Active Galactic Nuclei (AGN)

We divert for a moment to address a caveat brought up in Chapter 1.2. All star-forming galaxies that have strong emission lines due to ionized nebular gases are ELGs; however, not all ELGs are star-forming galaxies because emission line features can have other physical origins. If, for example, the galaxy possesses an active galactic nucleus (AGN; i.e., an actively accreting supermassive black hole), photoionization of gases within the supermassive black hole's accretion disk will generate a spectrum of strong emission lines that can be difficult to distinguish from lines generated by star formation in the absence of spatially-resolved spectroscopic data. If one has such spectroscopic data, one can measure diagnostic emission line ratios which can distinguish between AGN and star formation processes (Kewley et al. 2019). However, we use only photometric observations here and do not have access to such diagnostic line ratios. Thus, if any AGN were observed by FENIKS, our SED modelling program will falsely flag them as star-forming ELGs.

As our primary goal is to study star-forming ELGs and not AGN, this may seem like a major flaw in our study. However, there are two reasons why AGN needn't concern us here. One reason is that AGN make up less than about 30% of the galactic population, so the majority of our detections will not be AGN falsely flagged as star-forming ELGs. The other reason is that mistakenly finding an AGN is not necessarily a bad thing. Though star-forming ELGs are our objects of interest for this particular study, identifying AGN for future studies is relevant to our broader goal of determining the star-forming and quenching mechanisms of early galaxies, as AGN are believed to be one mechanism by which the star formation activity may quench and a star-forming galaxy becomes quiescent. Thus, permitting some AGN to slip into our sample set is an acceptable result of our study, even if it is not necessarily intended. In the future, follow-up spectroscopic studies of our identified ELGs will allow us to robustly identify the AGN that made it into our ELG sample set.

We have presented the theory behind SED modelling and the choice behind the FENIKS survey's *K*-split filters. In the next chapter, we will provide the details of the FENIKS pilot survey, including the size of the pilot dataset, what fields were targeted, and what data from other surveys were combined with our results to yield our final analysis dataset. We will also discuss the SED modelling program that we chose to use, and what modifications were made to this program to allow us to quantify the number of ELGs observed by FENIKS.

Chapter 3

Data Collection and Analytical Methods

3.1 The FLAMINGOS-2 Extragalactic Near-IR *K*-Band Split Pilot Survey

To test the effectiveness of the new *K*-split filters, a small-scale pilot survey was carried out from 2017 to 2019. While the full FENIKS survey intends to cover 0.5 deg^2 of sky in three deep legacy fields (COSMOS, GOODS – South, and UDS), the pilot survey covered a much smaller field of just 0.02 deg^2 of sky, with survey areas taken from the Chandra Deep Field Survey (CDFS) field as well as from the COSMOS 352 and COSMOS 544 fields. Despite a factor of 25 reduction in total survey area, the pilot survey still imaged 9,593 galaxies at redshifts $0 \leq z \leq 6$. In Table 1, we present the coordinates of the targeted fields as well as the total exposure time achieved for the K_{blue} and K_{red} filters in each field.

Field Name (right ascension [°], declination [°])	**Filter**	**Total Exposure Time [hours]**
CDFS (53.082, -27.809)	K_{blue} K_{red}	1.9 2.8
COSMOS 352 (150.090, 1.703)	K_{blue} K_{red}	3.2 3.7
COSMOS 544 (150.442, 2.557)	K_{blue} K_{red}	6.6 3.0

Table 1: Summary of observations in CDFS, COSMOS 352, and COSMOS 544 fields. Adapted from Esdaile et al. (2021), with permission.

The new *K*-split observations from the pilot survey were combined with two previously-compiled photometric catalogs, these being the UltraVISTA DR3 (Muzzin et al. 2013) and ZFOURGE (Straatman et al. 2016) catalogs. The UltraVISTA catalog covers galaxies in the COSMOS 352 and 544 fields, while the ZFOURGE catalog covers galaxies from the CDFS field. By combining the FENIKS pilot photometry with photometry from these catalogs, we were able to obtain a final photometric dataset of 50 filters for the COSMOS fields, and 37 filters for the CDFS fields. Figures 1 and 2 present the filter curves for all of the filters used by UltraVISTA and ZFOURGE, respectively. In Figure 3, we show the photometric data of galaxy 16181 from the UltraVISTA catalog as an example of what our data typically looks like, with the new K_{blue} and K_{red} data points highlighted.

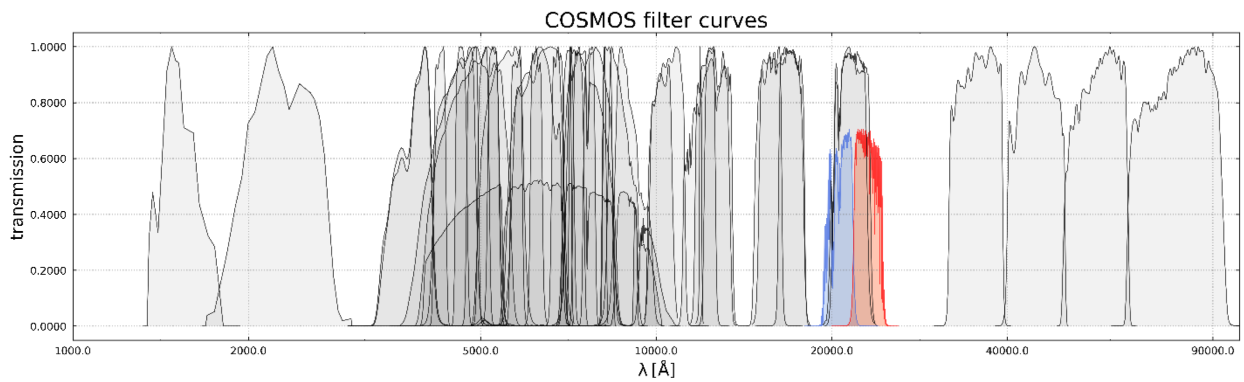


Figure 1: The transmission curves for the filters used on the COSMOS fields, which were covered by the UltraVISTA survey. The K_{blue} and K_{red} filters are highlighted in blue and red, respectively. The majority of our filters cover the visible and near-IR bands, with some coverage into both the UV and mid-IR.

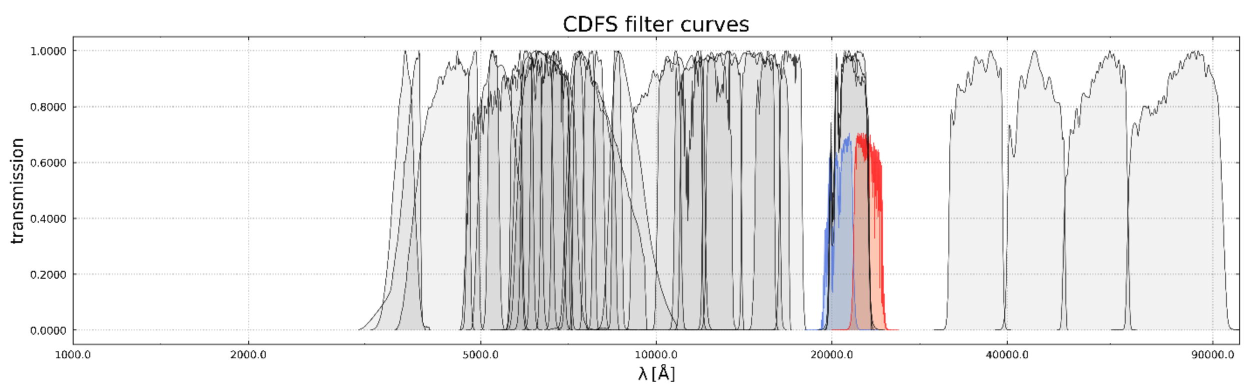


Figure 2: The curves for the filters used on the CDFS field, which was surveyed by ZFOURGE. The K_{blue} and K_{red} filters are highlighted in blue and red, respectively. Compared to the COSMOS fields, these galaxies lack UV coverage, but have a similar coverage level for the visible, near-IR, and mid-IR regimes.

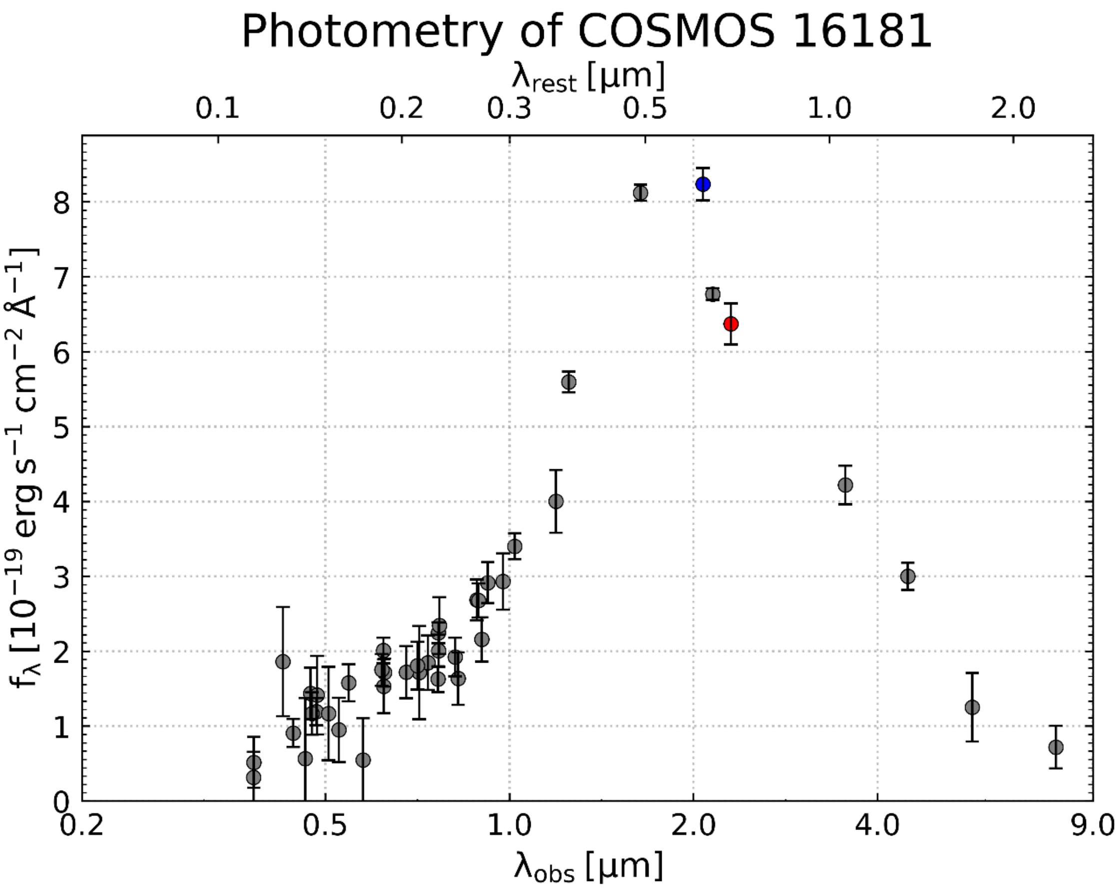


Figure 3: Galaxy 16181 from the UltraVISTA catalog. The observed, redshifted wavelength in microns is plotted on the lower x-axis, while the upper x-axis shows the rest-frame wavelength (the wavelength at which the light was originally emitted before it was cosmically redshifted). The flux density has been scaled by a factor of 10^{19}. The blue point is K_{blue} and the red point is K_{red}; between them, in grey, is the original K_s point.

If a filter is "missing", it means a significant observing error, systematic or otherwise, occurred that rendered the data point from that filter completely unusable for analysis. In the ZFOURGE catalog, we found the majority of our missing data points were from filters F105W (45 data points missing), F140W (55), F814W (21), and IA598 (9). Considering that we analyzed 3,198 ZFOURGE galaxies (and thus more than 90,000 data points overall), these numbers are remarkably small. In the 4,082 UltraVISTA galaxies that we analyzed, we found the majority of our missing data points were from the ultraviolet filters, with 1,337 FUV and 1,320 NUV data points missing, and approximately 1,145 missing data points each for the u, g, r, I, and z filters from the Canada-France-Hawaii Telescope (CFHT).

3.2 Bayesian Analysis of Galaxies for Physical Inference and Parameter EStimation

Because of the sheer size of our dataset, it was necessary to use a computer script to process and analyze our data. Our script was written in Python and utilizes BAGPIPES (Carnall et al. 2018) for SED modelling purposes. BAGPIPES is a Bayesian fitting program that uses a nested sampling algorithm (Skilling 2006) to explore semi-random combinations of physical parameter estimates, using these parameters to generate model SEDs that it measures against the data. The parameters it explores are not totally random; the user must specify upper and lower bounds for each parameter, and for components such as the SFH and dust attenuation, a specific form must be chosen. For dust attenuation, the user must choose to use one of the following dust attenuation curves: the Calzetti et al. (2000) curve, the Cardelli et al. (1989) curve, the Charlot & Fall (2000) curve, or the Salim et al. (2018) curve. Users also must specify an SFH model. BAGPIPES allows for the implementation of piecewise or non-parametric models, but it also comes with a set of parametric models that parametrize the SFR as a function of the time t. Some of the parametric models offered by BAGPIPES include the delta-function burst model in which all stars form at once, the constant model which treats SFR as constant from some time t_0 to the present, the exponential model which treats SFR as a decaying exponential beginning at some starting time t_0, and the delayed exponentially declining model, which gives SFR as a function of time according to

$$SFR(t) \propto te^{-\frac{t}{\tau}} \qquad (1)$$

where the parameter τ is a decay timescale. We adopt the delayed model throughout this thesis and will limit our discussions of SFH and related parameters to delayed models.

BAGPIPES generates SED models within the user-specified constraints by interpolating between grids of pre-made SPS models covering a range of stellar ages and metallicities. The SPS models used by BAGPIPES are Bruzual & Charlot (2003) models, generated using the Kroupa & Boily (2002) initial mass function. Nebular gas emission, which may be generated by the ionization of gas by newly-formed stars and thus correlated to the galaxy's SFR, is added to the stellar spectra using the CLOUDY photoionization code presented in Ferland et al. (2017). Emission from the ISM is handled by the Draine & Li (2007) model, while attenuation from the IGM is handled by the Inoue et al. (2014) model.

The generated model SEDs are compared to the input photometric data using the log-likelihood function

$$\ln L = -\frac{1}{2}\left(\sum_i \ln(2\pi\sigma_i^2) + \sum_i \frac{\left(f_i - f_i(\boldsymbol{\theta})\right)^2}{\sigma_i^2}\right) \qquad (2)$$

where L is the likelihood, f_i is the ith input photometric data point with corresponding uncertainty σ_i, and $f_i(\boldsymbol{\theta})$ is the ith photometric point predicted by the model that uses the parameter set $\boldsymbol{\theta}$. Models that fit the photometric input more closely have larger log-likelihood values. BAGPIPES will always accept a model that increases the log-likelihood. Though it may seem counterintuitive to do so, BAGPIPES does not automatically reject every parameter set that decreases the log-likelihood. Instead, based on the log-likelihood value, BAGPIPES will accept some parameter sets even if they produce less accurate solutions; in degenerate, multidimensional parameter spaces, this ensures that the algorithm does not become trapped on local maxima and is able to conduct a more thorough exploration of the parameter space.

After many thousands of iterations, BAGPIPES returns the posterior distributions of the parameters. These are series of histograms showing how frequently it returned to each parameter value. The most frequently sampled values are those that produced the highest log-likelihood, and thus these values produce better-fitting models than those that were sampled less often and had low log-likelihoods. Each sampled set of parameters is used to generate a SED; typically, upwards of 2,000 different SEDs are generated and added into the SED posterior distribution. Each SED also generates an expected set of photometric observations, thus giving us access to a photometry posterior distribution as well. An example of a range of SEDs taken from the posterior, as well as the corresponding parameter posteriors, is shown in Figures 4 and 5. In Figure 6, we show an example of a SFH posterior distribution using the delayed model detailed earlier.

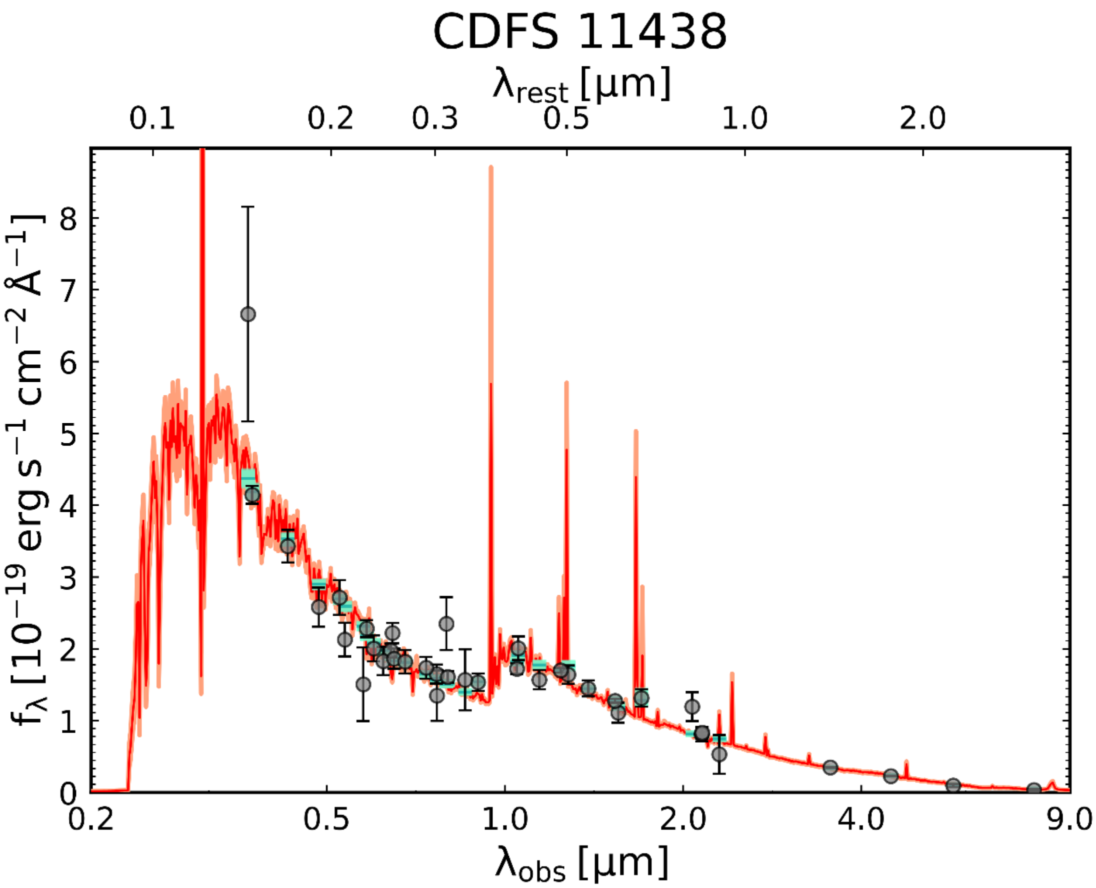


Figure 4: The fitted SED posterior for galaxy 11438 from the CDFS field, taken from the ZFOURGE catalog and observed at a redshift of $z \sim 1.54$. The dark red line marks the median SED, while the pale red region outlines the 16th to 84th percentiles of SEDs explored by BAGPIPES while fitting. The grey points with error bars represent the photometry, while the cyan boxes represent the photometry posterior. Overlap between the grey points and cyan boxes indicates agreement between the model and the observations. The presence of sharp emission lines and a high flux density at rest-frame wavelengths < 0.2 μm suggest that this galaxy has not quenched yet.

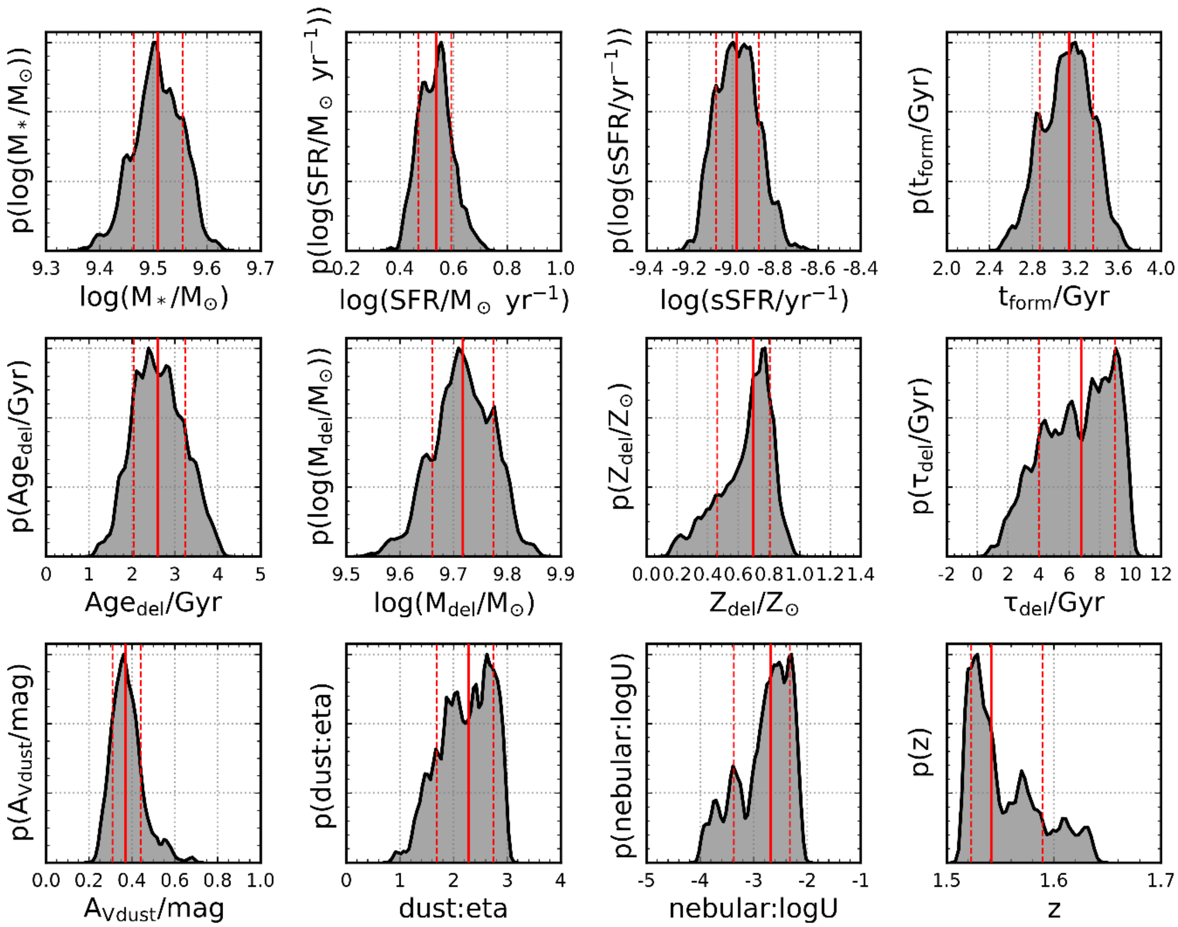


Figure 5: The posterior distributions for galaxy 11438. The solid red line in each plot represents the median value, while the dashed red lines mark the 16th and 84th percentile values. From left to right, top to bottom, the parameters are: stellar mass, average SFR over the last 100 Myr, specific SFR (the SFR divided by the stellar mass), average age of the universe when most of the galaxy's stars formed, time since star formation switched on, stellar mass formed by the delayed SFH model (including stellar remnants and material returned to the ISM), metallicity of the stars, timescale of the delayed SFH model, attenuation due to dust, extra dust factor for nebulae, base-10 logarithm of the ionization parameter, and redshift.

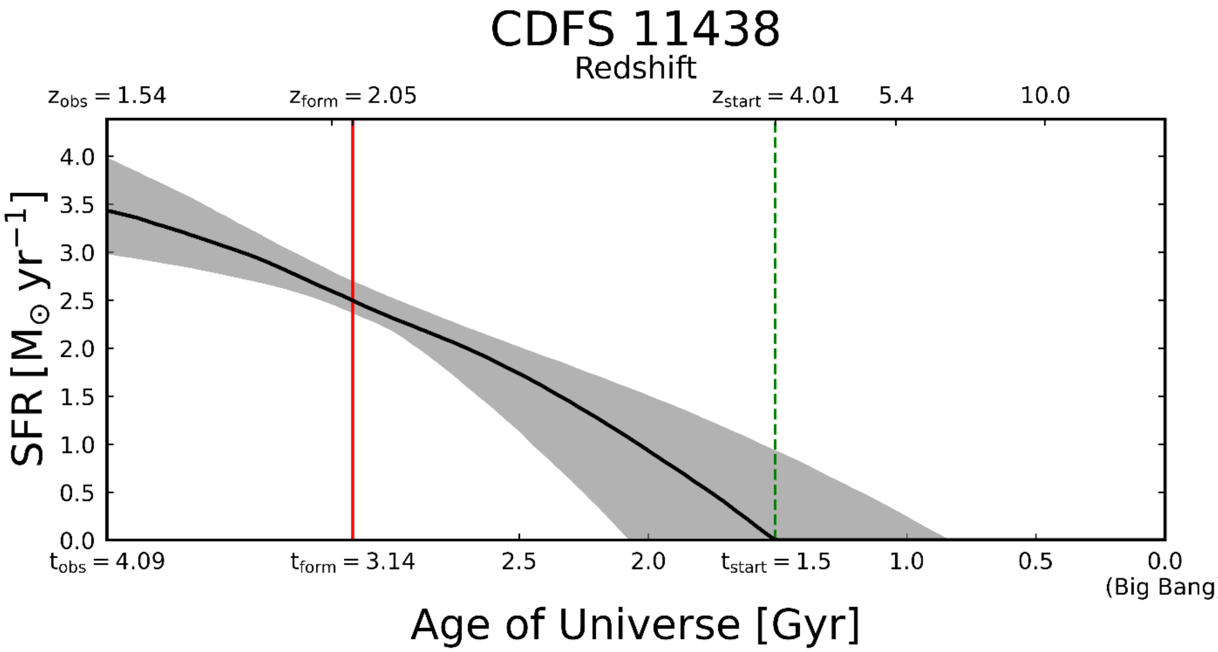


Figure 6: The SFH posterior distribution for galaxy 11438, using the delayed exponential model. The solid black line is the median SFH model, with the grey shaded region representing 16th through 84th percentile models. The lower x axis marks the age of the universe, with the upper x axis denoting the corresponding redshift z. The solid red line marks the location of t_{form}, the average age of the universe when most of the galaxy's stars formed. The left edge of the plot marks t_{obs}, the age of the universe at the lookback time when the galaxy was observed. The dashed green line marks the location of $t_{start} = t_{obs}$ - Age$_{del}$, which is the median age of the universe at which star formation in galaxy 11438 first started. The difference between t_{obs} and t_{form} is the mass-weighted age, or the average age of stars in the galaxy weighted by their mass. Stars in galaxy 11438 are thus typically about 0.95 Gyr old. The median model shown in solid black has a decay timescale of 6.55 Gyr and first switched on 2.59 Gyr before the time at which it was observed.

Figure 5 represents the core of the BAGPIPES fitting procedure, and it is worth taking a brief diversion to discuss it in depth. Each histogram presented in this figure is the actual range of parameter values explored by BAGPIPES while model fitting was being performed for galaxy 11438 from the CDFS field. It is from these ranges of parameters that the SED posterior distribution is developed. These parameter ranges are also the sought-after estimates of the galaxy's physical properties that we can use to infer the galaxy's star-forming rate and history. We now address each parameter from top to bottom, left to right, to explore their physical significance. The first parameter is the stellar mass, and it describes how much mass is contained in the galaxy that is in the form of main-sequence and evolved stars. Stellar mass does not include the mass of dust and gas that in the ISM, regardless of whether or not this gas and dust is primordial or was ejected from stars. Stellar mass also does not include the mass locked away in the form of stellar remnants such as white dwarfs, neutron stars, and black holes. The second parameter is the SFR, and it describes the average rate over the past 100 Myr at which new stellar mass was being formed. This parameter is of key importance to our study; its value will be large in ELGs and considerably smaller in quenched galaxies. The third parameter, the specific SFR, is the SFR divided by the stellar mass of the galaxy. The fourth parameter, t_{form}, is the average age of the universe when the galaxy's stars formed, calculated as

$$t_{form} = \frac{\int_0^{t_{obs}} t\, SFR(t)dt}{\int_0^{t_{obs}} SFR(t)dt} \qquad (3)$$

It can be thought of as an average "birth year" for the stars in the galaxy; referencing Figures 5 and 6, $t_{\mathrm{form}} = 3.14$ Gyr indicates that on average, the galaxy's stars were formed when the universe was 3.14 Gyr old. The fifth parameter, $\mathrm{Age}_{\mathrm{del}}$, is the time since star formation first began in the observed galaxy (the subscript "del" refers to the delayed model of SFH); again referencing Figures 5 and 6, $\mathrm{Age}_{\mathrm{del}} = 2.59$ Gyr means that the galaxy's first stars formed 2.59 Gyr before the epoch of observation. The sixth parameter is the total mass formed by the delayed star formation model. Unlike the stellar mass parameter, this parameter also includes the mass locked away in stellar remnants, as well as the mass of gas and dust that was once contained within stars but was ejected and returned to the ISM by processes like stellar winds and supernovae. The seventh parameter is Z, not to be confused with the redshift z. Z is the metallicity of the stars being formed by the galaxy, where metallicity refers to the fractional abundance of elements heavier than hydrogen and helium (such elements are known as metals in astronomy). In the early universe (and thus for older galaxies), Z tends to be smaller than it is today because metals had not yet had time to form. The eighth parameter, τ_{del}, is the timescale of the delayed SFH model and tells us whether star formation in the galaxy occurred in a burst (small τ_{del}) or occurred more gradually (large τ_{del}). The ninth parameter, $(A_V)_{dust}$, is the attenuation in V-band magnitudes due to dust, and represents the intensity of dust obscuration in the galaxy. The tenth parameter, written above as dust:eta and formally written as η, is a multiplicative factor that accounts for the excess dust within nebulae. The young stars forming within nebulae can be heavily enshrouded in dust, causing the light from these stars to be more strongly attenuated. When generating the model spectrum for galaxies modelled or fitted with BAGPIPES, the attenuation factor $10^{-0.4(\eta-1)(A_V)_{dust}}$ is multiplied into the spectrum of stars in nebulae; as η increases, this factor shrinks and thus the intensity of the spectrum emitted by stars in nebulae is reduced, thus describing an increase in dust attenuation. The attenuation factor is also applied to the emission line fluxes, so a higher value of η will diminish the strength of the modelled emission lines. The eleventh parameter, written as nebular:logU and formally written as log(U), is the base-10 logarithm of the CLOUDY ionization parameter, $U = \phi(H)/cn(H)$, where $\phi(H)$ is the flux of hydrogen-ionizing photons, c is the speed of light, and $n(H)$ is the density of hydrogen atoms. log(U) is also key to our project, as more negative values of log(U) indicate galaxies with a higher flux of ionizing radiation; thus, log(U) is partly responsible for setting the shape of the nebular emission spectrum. The twelfth parameter, z, is the redshift, which is only a free parameter for galaxies with redshifts that were determined photometrically and not spectroscopically (more on this in Chapter 4).

For the purposes of identifying ELGs, we were interested in finding ways to characterize the strengths of the galaxies' emission lines. To this end, we added new features into the BAGPIPES code. Specifically, we made it possible to break each model SED into its three major components: the stellar continuum spectrum, the emission line spectrum, and the dust emission spectrum. To quantify the impact that emission lines were having on the shape of the spectrum, we extracted predicted photometric points from two versions of the model SED, one which contained all three components and one which had no emission line component. We thus generated two sets of posteriors for each galaxy's SED and photometry; one posterior was made from models that included the emission line spectrum ("lined"), and the other posterior was made from models that excluded the emission line spectrum ("lineless"). Two examples of these lineless SEDs in comparison to the lined SEDs are shown in Figure 7.

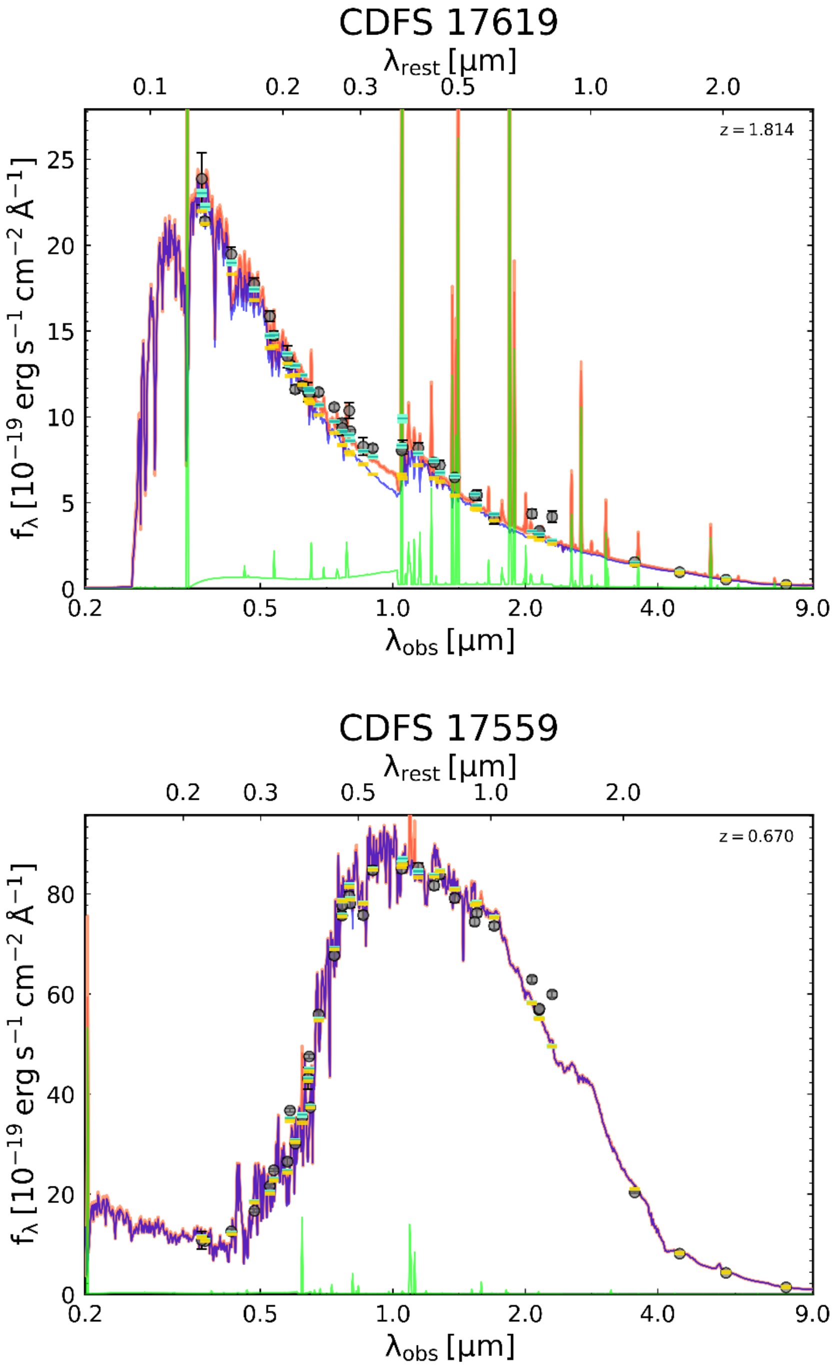


Figure 7: The split SEDs of galaxies 17619 (top) and 17559 (bottom), both from the ZFOURGE catalog and with redshifts labelled in the upper right. In each plot, the red lines represent the complete SEDs while the blue lines represent the median lineless SEDs. The green line represents the difference between these two SEDs – in effect, it is the contribution to the spectrum by emission lines alone. Cyan boxes represent the complete photometry posterior, while golden bars indicate the median photometry of the median lineless SED. Galaxy 17619 clearly has a strong emission line component, stark discrepancies between the red and blue, and significant differences between the cyan and golden boxes; it is more than likely a star-forming ELG. Galaxy 17559 has very few visible emission lines, a tight matchup between the red and blue SEDs, and minimal differences between the cyan and golden boxes; it is likely to be a quiescent galaxy.

Having access to these two kinds of photometric models allowed us to design plots that compare the change in magnitude between lined and lineless models as a function of galaxy redshift and wavelength of observation; the development, analysis, and implications of these “delta-magnitude” or Δm plots will be detailed in Chapter 4.

We have outlined the FENIKS pilot dataset and the program that we used to conduct our analysis. In the next chapter, we will discuss the analysis itself, starting from our choice of constraints for our SED models, and then discussing the development and interpretation of the Δm plots, concluding with the total number of ELGs detected by this analysis.

Chapter 4

Analysis and Results

4.1 Model Fitting for a Select Sample of 14 FENIKS Galaxies

We began our analysis by selecting a set of 14 galaxies from the complete dataset to use as a testbed for model selection and code development. These 14 galaxies were selected because they contained minimal missing filters and errors that were, for the most part, smaller than 10%; thus, any difficulties in fitting could be attributed to flaws in the code and not in the data itself.

For fitting, we used a delayed SFH model with the parameter τ_{del} as the exponential decay timescale. We permitted the galaxy's age to vary between 100 Myr and 15 Gyr – that is, between recent cosmic history and a bit more than the age of the universe. While BAGPIPES was free to sample any age within these limits, the program rejects ages that would predate the Big Bang; for example, a galaxy at redshift $z = 0$ could not have an age of more than 13.8 Gyr, and a galaxy at redshift $z \sim 3$ could not have an age of more than 2.3 Gyr. If the algorithm samples a model galaxy that is older than the universe, the likelihood value is immediately set to 0 and the model gets rejected. The timescale τ_{del} was constrained to be between 10 Myr and 10 Gyr, where small values of τ_{del} describe rapid bursts of star formation in which stars are formed almost all at once, and large values of τ_{del} describe steady star formation histories in which the galaxy forms stars at a nearly constant rate throughout its life. The total mass formed by the SFH was constrained to be between 1 and 10^{15} solar masses. The metallicity of stars in the galaxy was constrained to be between 0.04 and 1.5 times the solar metallicity. We used a Calzetti et al. (2000) dust attenuation model, with a dust attenuation coefficient constrained to be between $A_V = 0$ (no dust attenuation) and $A_V = 5$ (strong dust attenuation). We also constrained the value of η to lie between 0 (no dust around young stars) and 3 (significant dust enshrouding young stars). The logarithm of the ionization parameter, log(U), was constrained to lie between -4.0 (low ratio of ionizing photons to hydrogen atoms) and -2.1 (high ratio of ionizing photons to hydrogen atoms). Finally, the redshift of the galaxy was handled on a case-by-case basis; for those galaxies that had a spectroscopically-observed redshift z_{spec}, we fixed $z = z_{spec}$. Otherwise, we set the redshift to lie between $z_{phot} - \sigma$ and $z_{phot} + \sigma$, where z_{phot} is the photometrically-determined redshift and σ is the standard deviation on z_{phot}, both derived using the EAZY photometric redshift fitting program (Brammer et al. 2008). This choice was motivated by the fact that z_{spec} tends to be a very well-constrained measurement, while z_{phot} is subject to comparatively larger uncertainties.

Using these constraints, we performed SED fitting for 14 galaxies, and developed our SED splitting routine using these SEDs. Examples of the fitted SEDs and split SEDs we developed from these 14 galaxies were presented in Chapter 3. With the basic SED modelling infrastructure in place, we turned our attention to the development of the Δm plots.

4.2 Extending the Fitting Algorithm to a Larger Sample

The initial set of 14 galaxies served us well in developing the initial SED modelling and SED splitting routines, but a far larger sample would be needed in developing the delta-magnitude plots. To extend our analysis, we selected additional galaxies based on a set of criteria designed to exclude from the analysis galaxies with questionable or less robust measurements. Because the

CDFS and COSMOS catalogs were conducted by two different teams using different instruments, filters, and catalog formats, it was necessary to develop two different criteria.

For the CDFS fields, we restricted our attention to galaxies with AB magnitudes in the *K*-band $15 \leq m_K \leq 26.5$ to exclude galaxies that were unrealistically bright (most likely caused by cosmic ray contamination) and galaxies that were too dim for the measurements to be considered reliable. We required each galaxy to have star = 0 and nearstar = 0; this means that each galaxy is not likely to actually be a star that was imaged by mistake, nor is it too close on the sky to a bright star, which could cause the photometry to be contaminated. We required use = 1 and use_nosnr = 1; this is a flag raised by the catalog when the galaxy passes a set of criteria established by the ZFOURGE survey (see Straatman et al. (2016) for details). We constrained redshift to lie in the range $0.2 < z < 6.0$. We also required the number of Source-Extractor (SE) flags $SE_{flag} \leq 3$ and $SE_{flag} \neq 1$, where SE is a catalog-building program that was used to construct the multi-wavelength photometric catalogs from the survey images. Finally, we required the EAZY redshift fit quality $\chi_p < 1{,}000$, which means the fit to the galaxy's photometric redshift was not catastrophically bad. This results in selecting 3,198 galaxies from the CDFS field.

For the COSMOS 352 and COSMOS 544 fields, we restricted our attention to galaxies with AB magnitudes in the *K*-band $15 \leq m_K \leq 25$, again excluding galaxies that were contaminated by cosmic rays or too dim to use. The upper bound was adjusted to reflect the change in survey depth between the UltraVISTA and ZFOURGE surveys. We required the galaxies to have 0 star and contamination flags, which means that the galaxies are neither stars nor are they contaminated by starlight or cosmic rays. We required each galaxy to have less than 3 missing filters and to lie in the redshift range $0.2 < z < 6.0$. Finally, we required that the number of SE flags (denoted in the COSMOS catalogs as K_{flag}) were restricted to $K_{flag} \leq 3$ and $K_{flag} \neq 1$, and that the EAZY redshift fit reached a chi-squared value $\chi_p < 1{,}000$ to ensure it was not catastrophically bad. This results in selecting 2,139 galaxies from the COSMOS 544 field, and 1,943 galaxies from the COSMOS 352 field.

The final criteria-selected sample covers 2,139 galaxies from the COSMOS 544 field, 1,943 galaxies from the COSMOS 352 field, and 3,198 galaxies from the CDFS field, for a final sample size of 7,280 galaxies out of the 9,593 imaged in the complete pilot dataset. Our sample set of 7,280 galaxies spans a redshift range from $z = 0.2002$ to $z = 5.9128$, and a stellar mass range from $10^{6.6}$ to $10^{11.7}$ solar masses.

4.3 Delta-Magnitude and its Relation to the Equivalent Width of Emission Lines

With our extended sample in place, we could begin the development of the Δm plots. As mentioned before, we are able to create two different versions of the SED and photometry posteriors for each galaxy, with one version including the emission line component and one excluding it. The photometry posteriors describe how much flux should have been observed through each of our filters based on the underlying SED model used to generate the photometry. If emission lines are present, then the photometry posterior for the lined model p_{lined} will be brighter than the photometry posterior for the lineless model p_{lineless}. From Hogg et al. (2002), the change in flux between the two models can be converted into a change in magnitude Δm in the X filter via

$$\Delta m = 2.5 \log\left(\frac{p_{lined}}{p_{lineless}}\right) = 2.5 \log\left(\frac{\int \frac{d\nu}{\nu} f_{lined,obs}(\nu) X(\nu)}{\int \frac{d\nu}{\nu} f_{lineless,obs}(\nu) X(\nu)}\right) \qquad (4)$$

where $f_{\mathrm{lined,\ obs}}(\nu)$ is the observed (i.e., redshifted) spectrum of the galaxy with the emission lines included, $f_{\mathrm{lineless,\ obs}}(\nu)$ is the observed spectrum with emission lines excluded, and $X(\nu)$ is the transmission curve for the X filter. This term is defined to be positive when emission lines are present. Δm can be used as a proxy for the rest-frame equivalent width W_{RS} of the emission line. The equivalent width of an absorption or emission line in the rest frame of the absorber or emitter is defined by

$$W_{RS} = \int \left(1 - \frac{f_{lined}(\lambda)}{f_{lineless}(\lambda)}\right) d\lambda \qquad (5)$$

and unlike the spectra used in calculating the observed change in magnitude, the spectra here are in the rest frame and not redshifted. For absorption lines, W_{RS} describes the width of a perfect absorption line (with $f_{lined} = 0$, or 100% absorption) that removes the same amount of flux from the stellar continuum spectrum as the actual absorption line (which is an imperfect absorber with $0 < f_{lined} < f_{lineless}$). Emission lines have $f_{lined} > f_{lineless}$; in these cases, we take the absolute value of the equivalent width to report W_{RS} as a positive number. The equivalent width is a metric of the emission line strength; a stronger emission line adds more flux to the observed spectrum and increases the magnitude of W_{RS}.

Increasing the ratio $f_{lined}/f_{lineless}$ has the effect of increasing both W_{RS} and Δm, which allows us to use Δm as an effective proxy for W_{RS}. Galaxies with high Δm values are thus expected to have emission lines with large rest-frame equivalent widths. However, the precise relationship between Δm and the rest-frame equivalent width of the emission line is filled with nuances. The observed flux coming from a galaxy is affected by many factors, including the intrinsic strength of the emission line (controlled in part by the present SFR and sSFR), the shape of the nebular emission spectrum (controlled by the nebular parameter log(U)), and attenuation due to dust (controlled by the coefficient A_V). In addition, the flux f collected by a filter spans a bandwidth much larger than the emission line itself; thus, the Δm value reported for a specific emission line of equivalent width W_{RS} may be contaminated by other features contained within the filter's bandwidth, such as other emission lines, nearby absorption features, and the shape of the continuum around the targeted emission line.

To explore the complex relationships between Δm observed in a chosen filter, the rest-frame equivalent width W_{RS} of an emission line targeted by that filter, and the four modelling parameters (log(U), A_V, SFR, and sSFR), we modelled 935 galaxies using BAGPIPES and generated simulated photometric observations for each. All of these modelled galaxies employed a delayed SFH with a total mass formed of 10^{10} solar masses and an age Age_{del} = 1 Gyr. The models covered decay timescales τ_{del} ranging from 50 Myr to 1.0 Gyr (which, in tandem with the fixed mass formed and age, covered SFRs ranging from 0.0 to 13.9 $M_{\odot}$/yr), dust attenuation coefficients ranging from 0.0 to 5.0, and log(U) parameter values ranging from -4.0 to -2.0. We modelled these observations twice, once for galaxies at a redshift z = 2.5 as observed through the K_{red} filter, and again for galaxies at a redshift z = 2.1 as observed through the K_{blue} filter. These redshifts were chosen because both filters at these redshifts will be centered very nearly on λ = 6,562.81 Å, which is the wavelength of the Hα emission line. From the simulated photometric observations as well as built-in methods from BAGPIPES, we extracted both Δm in the chosen filter and W_{RS} of the Hα line in the rest frame for each modelled galaxy. We plotted the modelled rest-frame Hα equivalent widths W_{RS} against the modelled Δm and the results are shown below in Figure 8.

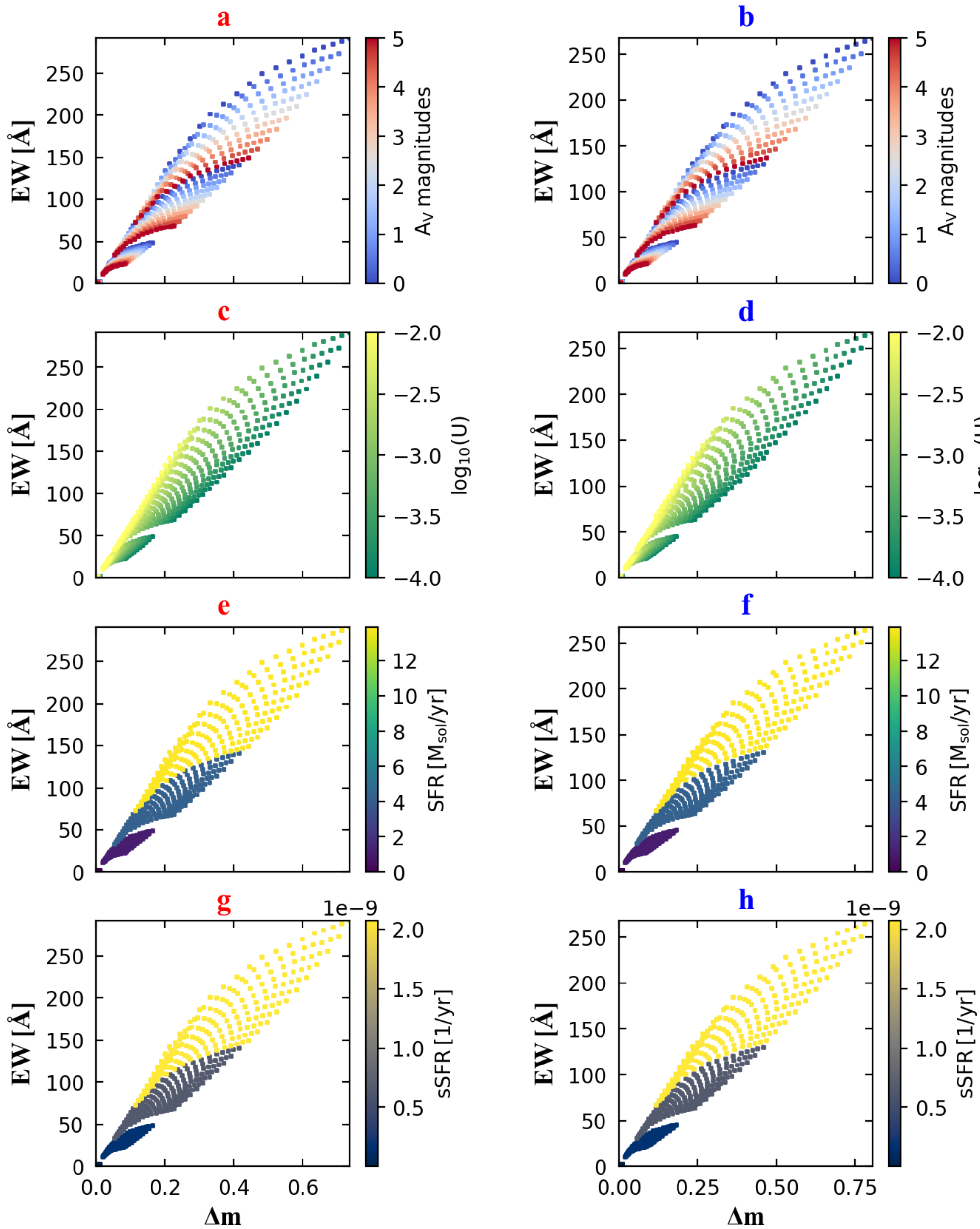


Figure 8: The simulated W_{RS} vs. Δm plots for the K_{red} filter (left column) and for the K_{blue} filter (right column). Each plot is color-coded to highlight the relationship between a particular parameter and W_{RS}. The color-coding in 8a and 8b represents the dust attenuation coefficient A_V, with cooler colors indicating less dust. 8c and 8d are color-coded to the value of log(U), with darker tones indicating a higher flux of ionizing radiation. 8e and 8f are color-coded to the value of the SFR, with brighter colors indicating higher SFR. 8g and 8h are similarly coded to the value of the sSFR. The plot titles are color-coded to the filter used in each plot.

Examination of these plots reveals a great deal about the relationship between the observed Δm and the implied rest-frame equivalent width of the Hα emission line. For fixed W_{RS}, Figures 8a and 8b reveal that increasing the level of dust attenuation will cause the observed Δm to decrease; in other words, in the presence of more dust, a stronger emission line is needed to produce the same Δm that would be observed from a weaker emission line in a less dusty environment. Figures 8c and 8d shows that for a fixed log(U) value, Δm increases linearly with W_{RS}. More negative values of log(U) decrease the slope of the graph, meaning that higher Δm are observed for weaker emission lines if the magnitude of log(U) is high. However, it should be noted that this dependency is entirely a result of the fact that the filters are wider than the emission line itself; additional simulations conducted with hypothetical filters of much narrower width showed that this dependency disappears if the filter's bandwidth spans only the emission line itself and does not include other spectral features (see Appendix A). Thus, the dependency of Δm on log(U) is not related to emission line strength but is instead the result of the filter integrating parts of the spectrum other than the emission line. Finally, Figures 8e, 8f, 8h, and 8g plot galaxies with varying levels of star formation, and show that galaxies with low SFRs and sSFRs generally have lower Δm values and weaker emission lines with lower equivalent widths. Most importantly, for the range of log(U), SFR, sSFR, and A_V values covered by the FENIKS galaxies, these simulations suggest that an observed Δm value of about 0.30 corresponds to a rest-frame equivalent width of at least 100 Å or more. From these plots, we can thus define a threshold Δm value above which a galaxy can be classified as an ELG; for this study, we take that threshold to be 0.30. We note that the exact choice of threshold is arbitrary, and that adopting higher thresholds should be used to select more extreme ELGs. The appropriate threshold is also filter-dependent; while a cutoff of 0.30 might target at least 100 Å in medium width IR filters, a different threshold will be needed for other bands.

4.4 Delta-Magnitude Plots and Separating ELGs from non-ELGs

For all of the galaxies in our extended sample, we calculated the Δm values and produced a plot of Δm versus galaxy redshift z for the medium band IR filters used in the observations. This choice was motivated by the fact that at different redshifts, each filter will be observing a different rest-frame wavelength and will thus be sampling different emission lines. Equivalently, this means the same emission is sampled by different filters at different redshifts. In Figures 9, 10, and 11, we present three Δm plots, one for the K_{blue} filter, one for the K_{red} filter, and one for the K_s filter that they split. In Figures 12, 13, 14, 15, and 16, we present additional Δm plots for the other medium band IR filters used in our catalog, these being the J_1, J_2, J_3, H_1, and H_2 filters which split the broad J- and H-band filters.

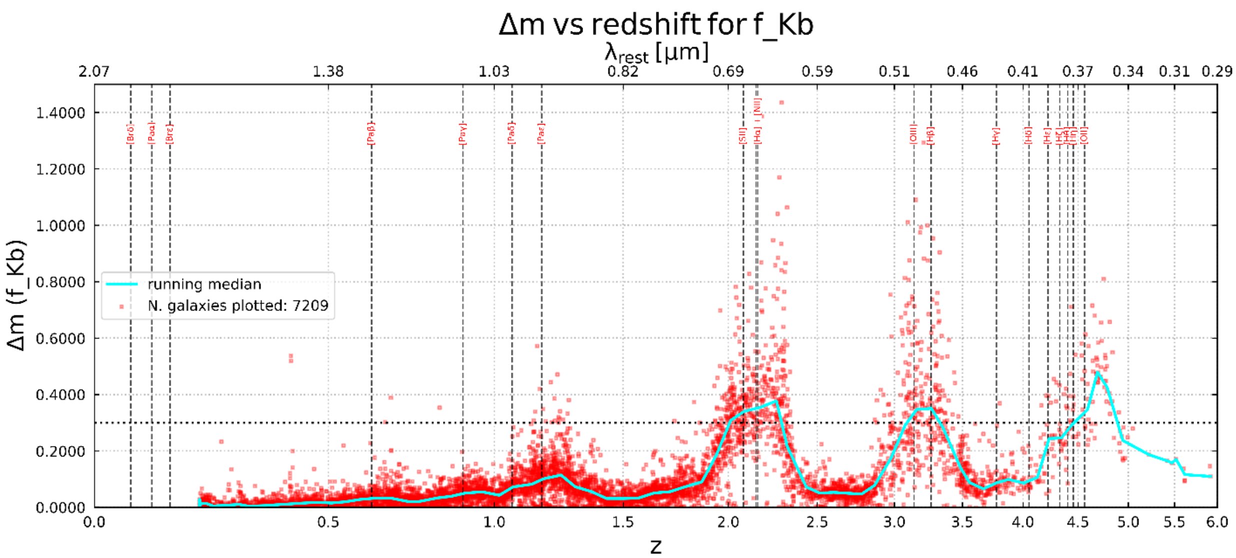


Figure 9: The Δm plot for the K_{blue} filter. Each dashed vertical line indicates an emission line, while the horizontal dashed line indicates the threshold Δm value (0.30) above which a galaxy is flagged as a potential ELG. The cyan line is the running median, which highlights trends in the data.

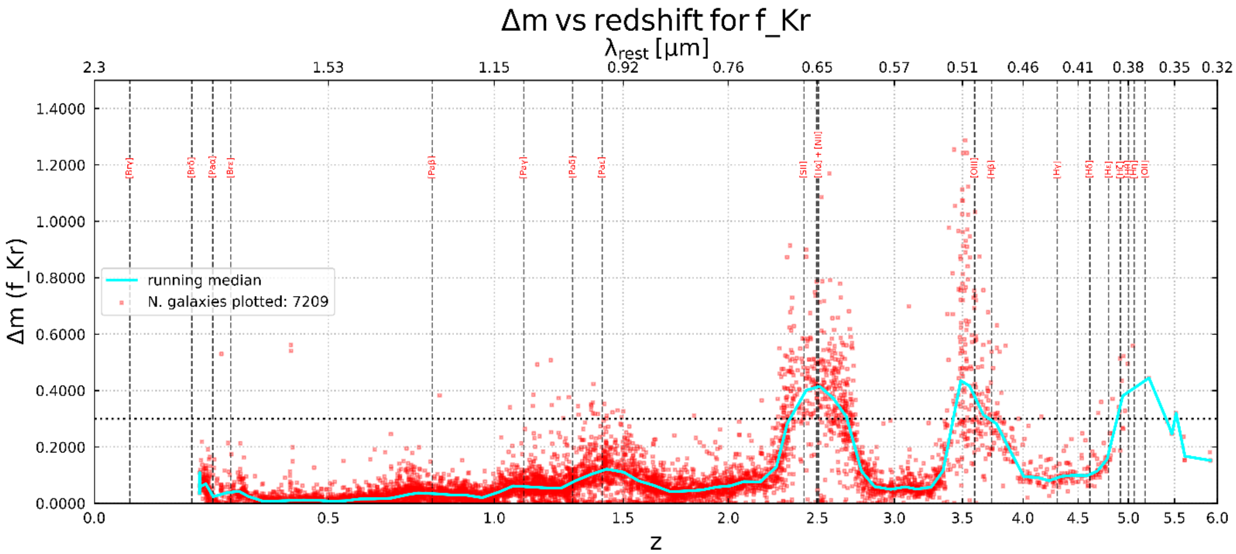


Figure 10: The Δ*m* plot for the *K*red filter. Each dashed vertical line indicates an emission line, while the horizontal dashed line indicates the threshold Δ*m* (0.30) value above which a galaxy is flagged as a potential ELG. The cyan line is the running median, which highlights trends in the data.

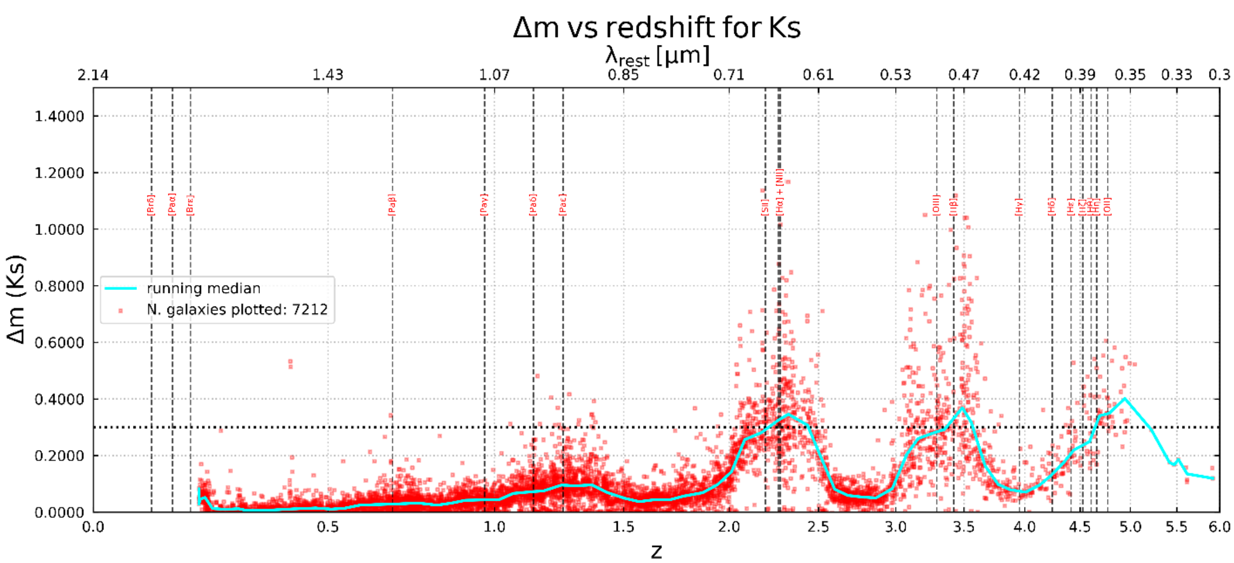


Figure 11: The Δ*m* plot for the *K*s filter. Each dashed vertical line indicates an emission line, while the horizontal dashed line indicates the threshold Δ*m* (0.30) value above which a galaxy is flagged as a potential ELG. The cyan line is the running median, which highlights trends in the data.

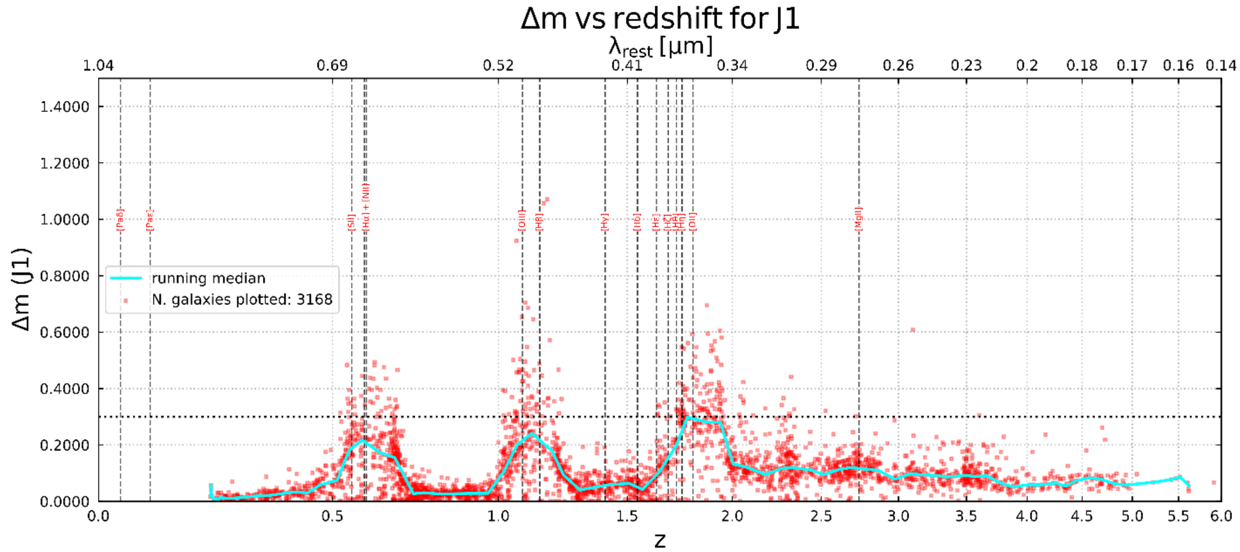


Figure 12: The Δ*m* plot for the *J*₁ filter. Each dashed vertical line indicates an emission line, while the horizontal dashed line indicates the threshold Δ*m* (0.30) value above which a galaxy is flagged as a potential ELG. The cyan line is the running median, which highlights trends in the data.

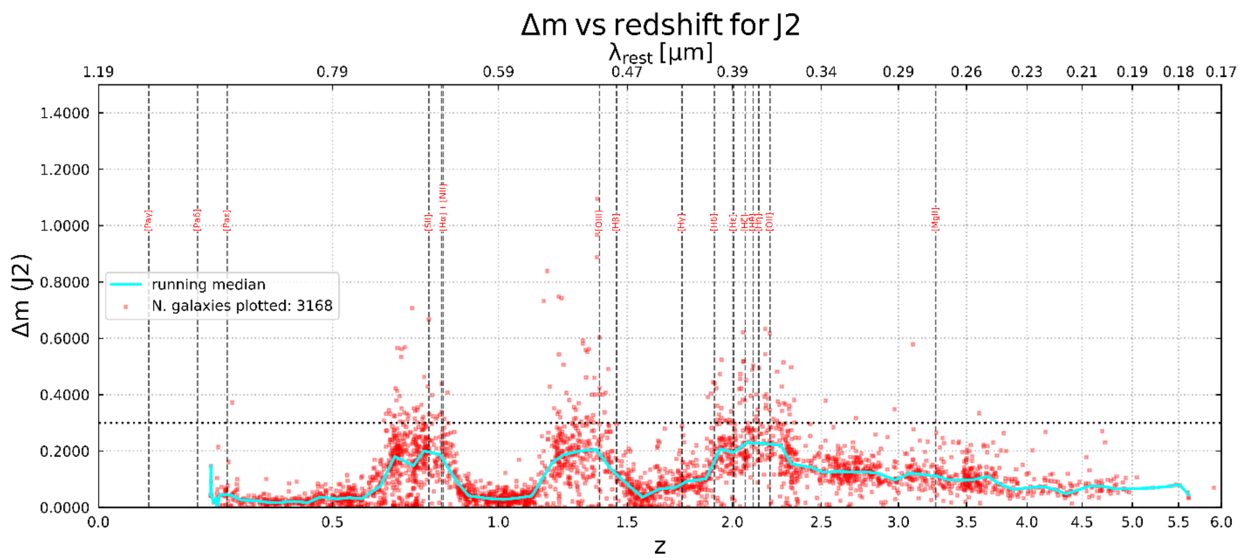


Figure 13: The Δ*m* plot for the *J*₂ filter. Each dashed vertical line indicates an emission line, while the horizontal dashed line indicates the threshold Δ*m* (0.30) value above which a galaxy is flagged as a potential ELG. The cyan line is the running median, which highlights trends in the data.

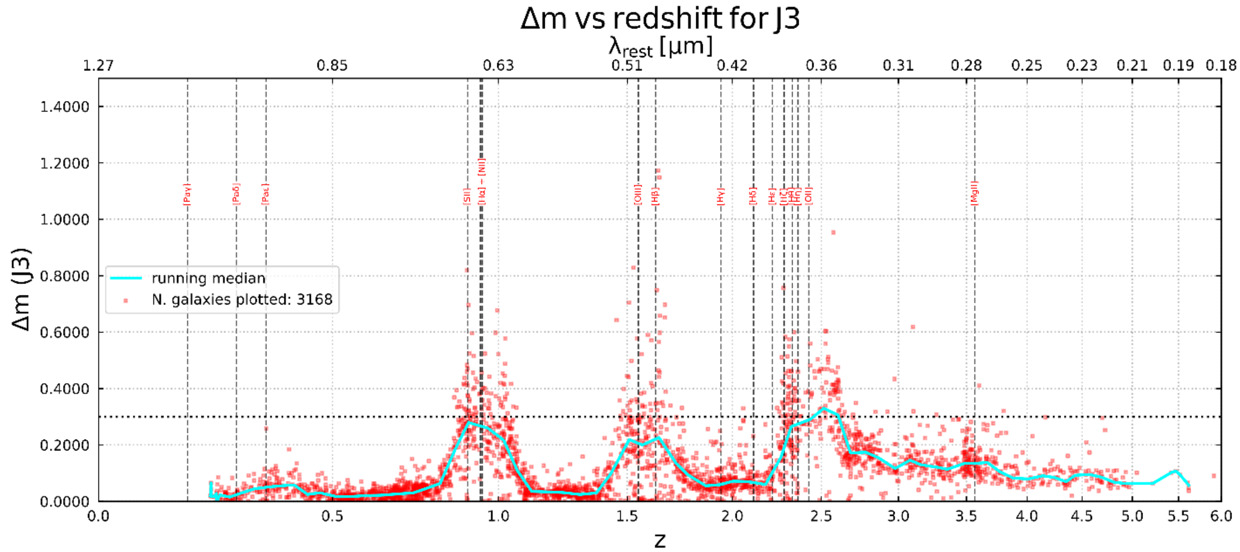


Figure 14: The Δ*m* plot for the J_3 filter. Each dashed vertical line indicates an emission line, while the horizontal dashed line indicates the threshold Δ*m* (0.30) value above which a galaxy is flagged as a potential ELG. The cyan line is the running median, which highlights trends in the data.

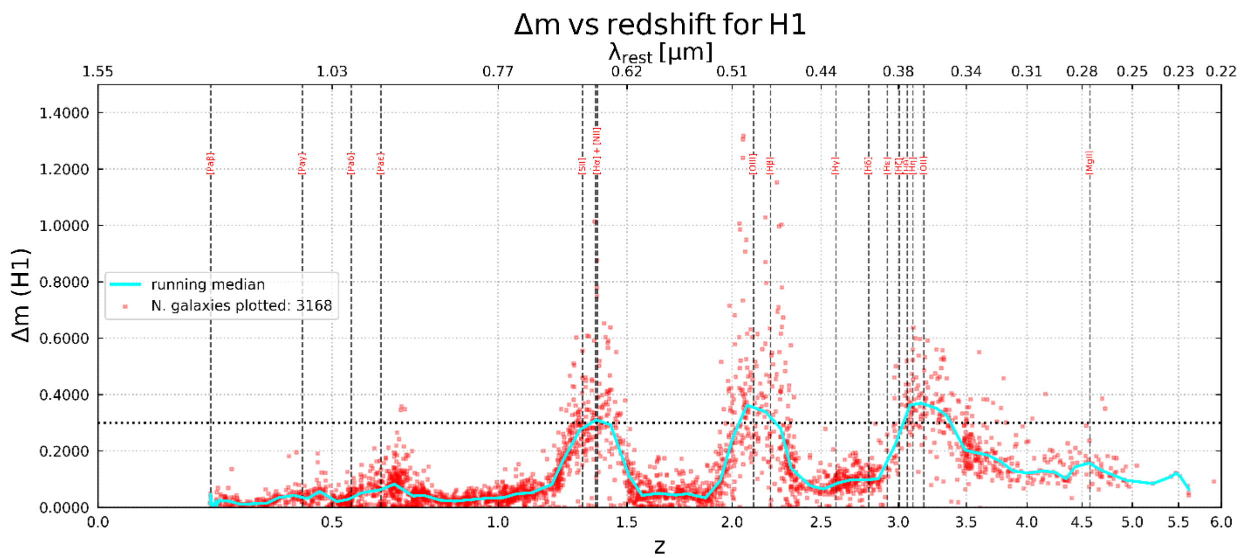


Figure 15: The Δ*m* plot for the H_1 filter. Each dashed vertical line indicates an emission line, while the horizontal dashed line indicates the threshold Δ*m* (0.30) value above which a galaxy is flagged as a potential ELG. The cyan line is the running median, which highlights trends in the data.

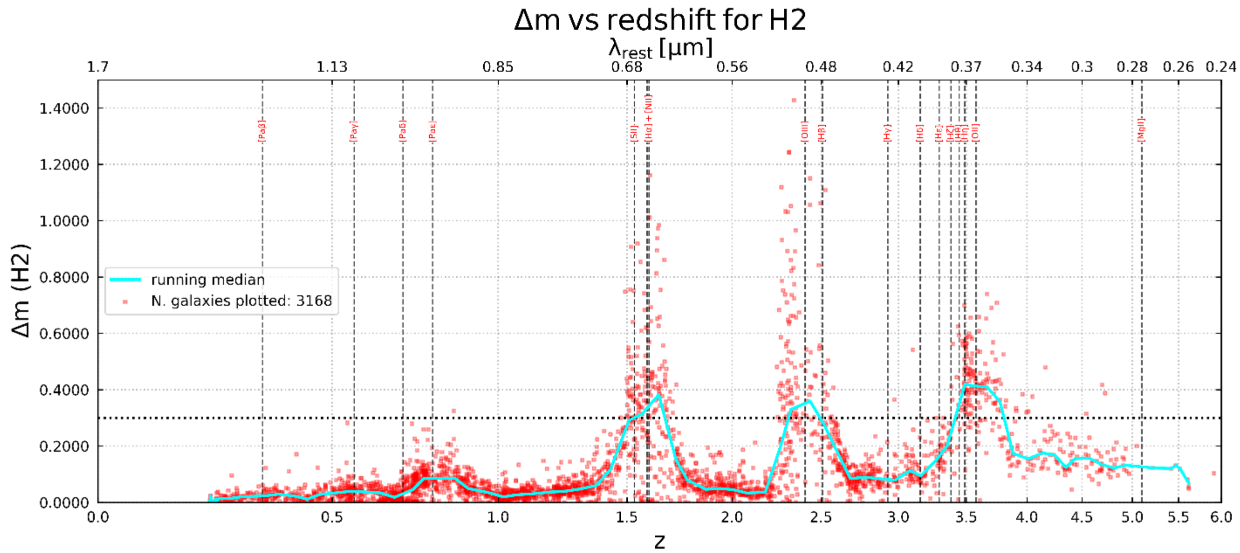


Figure 16: The Δ*m* plot for the *H*2 filter. Each dashed vertical line indicates an emission line, while the horizontal dashed line indicates the threshold Δ*m* (0.30) value above which a galaxy is flagged as a potential ELG. The cyan line is the running median, which highlights trends in the data.

There are several features of interest to note in these plots. For one, all plots show distinct peaks in the curve corresponding to the wavelengths of emission lines such as the Hydrogen Paschen and Balmer series lines, as well as the [SII], [NII], and [OIII] doublets. However, the peaks for the medium band *K*-split filters are both taller and narrower than the peaks for the broadband K_s filter, demonstrating the improved spectral resolution of the medium filters.

The redshift-dependent boosting effect can also be observed in the *K*-band plots by examining the peaks corresponding to the overlapping Hα line and [NII] doublet. Notice that in the K_{blue} delta-magnitude plot, the Hα+[NII] peak lies roughly between redshifts 1.90 and 2.40, while the same peak falls roughly between redshifts 2.25 and 2.75 for K_{red}. This shows that the Hα+[NII] lines, which originate from rest-frame wavelengths of 6,562.81 Å, 6,548.05 Å, and 6,583.45 Å, are redshifted to the K_{blue} filter ($1.9 < \lambda_{obs} < 2.2\ \mu$m) at redshifts of about $1.90 < z < 2.40$, and are redshifted further into the K_{red} filter ($2.1 < \lambda_{obs} < 2.4\ \mu$m) at redshifts of about $2.25 < z < 2.75$. The boosting effect for Hα+[NII] can thus be observed through the *K*-split filters for galaxies at redshifts between 1.90 and 2.75, with K_{blue} boosting observable at $1.90 < z < 2.40$, and K_{red} boosting observable at $2.25 < z < 2.75$.

The most important feature to note about these graphs is that even at the redshift regimes where the filters sample emission lines, there are still galaxies with Δm values close to 0. Such galaxies do not have significant emission lines present, which indicates that they are likely quenched, quiescent galaxies that are not forming stars. On the other hand, the galaxies corresponding to the peaks of the Δm plots are likely to have strong emission lines present and thus be active star-formers or AGN. By examining each galaxy's Δm values as a function of filter and redshift (which together are a proxy for wavelength), and selecting only those with high Δm values, it should be possible to automatically identify ELGs from the dataset.

4.5 The Number of ELGs Found by FENIKS

We now bring together all of the discussion above to yield the final result. Our exploration of the relationship between Δm and W_{RS} suggests that we can select galaxies with $W > 100$ Å in the rest frame by setting the threshold Δm value to be 0.30. We apply this threshold to the K_{blue}, K_{red}, J_1, J_2, J_3, H_1, and H_2 Δm plots shown above. Tallying up the galaxies flagged as ELGs by each Δm plot, we find that the total number of galaxies in the FENIKS pilot survey dataset with emission line equivalent widths in excess of 100 Å is 1,648, representing ~23% of the sample of galaxies modelled with BAGPIPES. Of these detections, 595 were detected by K_{blue}, 500 by K_{red}, 216 by J_1, 158 by J_2, 261 by J_3, 331 by H_1, and 384 by H_2. Note that the same galaxy can be detected by multiple filters, hence these numbers do not sum to 1,648. Of the galaxies detected via these methods, 589 were detected through only one filter, 524 were detected by two filters, 287 by three, 101 by four, 87 by five, 25 by six, and 35 by all seven medium band IR filters.

Chapter 5

Discussion and Conclusions

5.1 Discussion of Results

Our final analysis concluded that 1,648 galaxies out of the 7,280 galaxies that we analyzed were ELGs. These results are, of course, highly dependent on the Δm threshold that we chose, and we could get quite different results by adjusting Δm. If we wanted to select for only the most active and extreme of the ELGs, we could set Δm to 0.8 or 0.9, pulling out galaxies with equivalent widths exceeding 300 Å. Alternatively, we are free to drop the threshold Δm and examine galaxies forming stars at much more moderate rates. The key takeaway from this analysis is not the exact number of ELGs so much as it is the development of a program that allows us to select ELGs at the desired level of star formation intensity; such a tool will be invaluable when the need for detailed spectroscopic follow-up studies arises, as we can use this tool to quickly generate a list of viable targets.

In selecting these targets, we may need to turn our attention to an as-yet underutilized part of our program, this being the parameter posterior distributions such as the ones shown in Figure 5. While exceeding a Δm threshold is a good start towards declaring a galaxy to be an ELG, cross-referencing this information with the reported SFR and sSFR values in its parameter posteriors would be an effective way to double down on this claim. In addition, recall that the Δm and W_{RS} relationship, as presented in Chapter 4.3, is more nuanced than it may first appear. A galaxy flagged as a potential ELG by an initial Δm check could be better scrutinized by examination of its estimated SFR and sSFR, which should both be appreciably high if the galaxy truly is an ELG.

We present the beginnings of this work below. In Figure 17, we present the distributions of the stellar mass, A_V, SFR, sSFR, and mass-weighted age of all galaxies in our dataset as well as of those specifically flagged as ELGs. In Figures 18, 19, and 20, we present three color-coded plots of the modelled galaxies' stellar masses against redshift z, with color-coding used to describe the dust attenuation coefficient A_V, the SFR, and the sSFR. There are two plots in each of these figures; one contains all of the galaxies analyzed while the other presents only those flagged as ELGs.

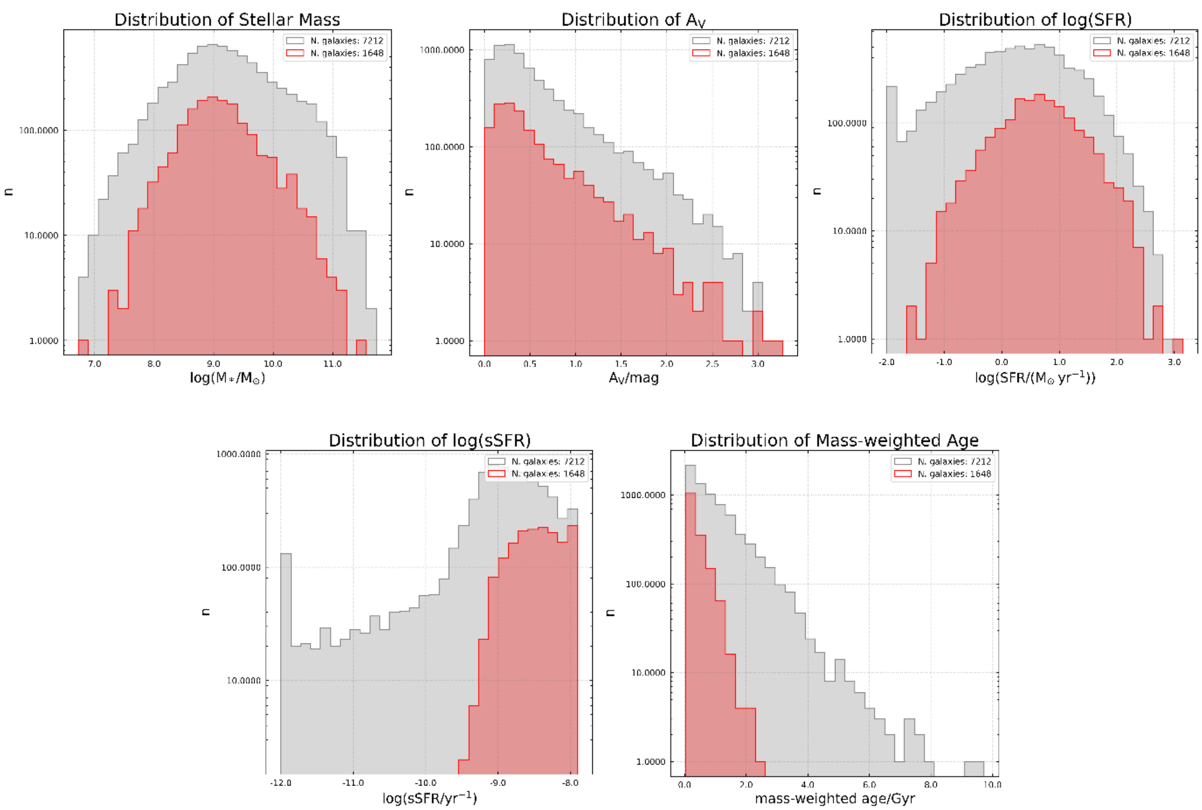
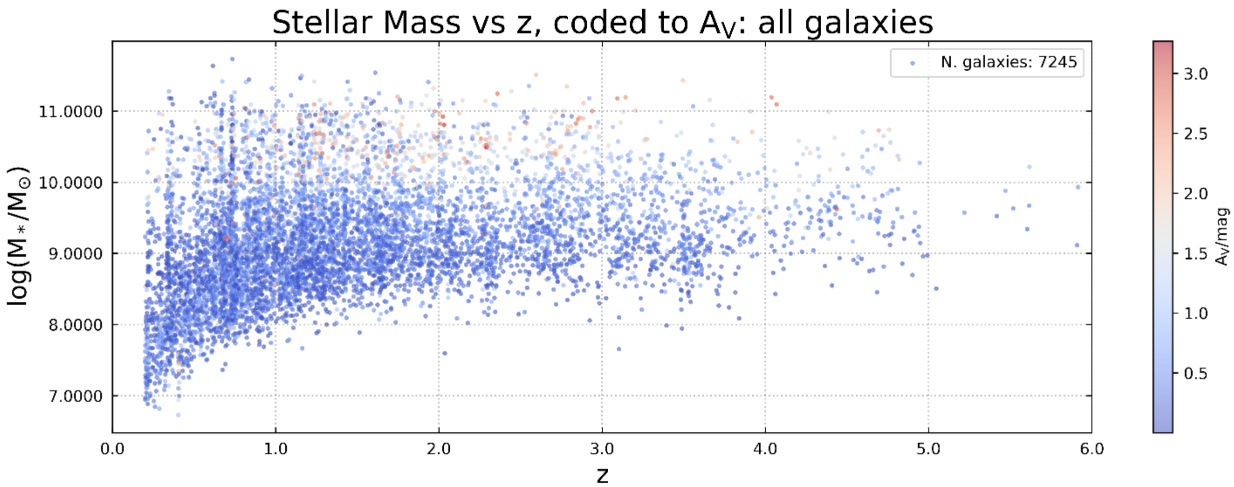


Figure 17: The distributions of (left to right, top to bottom) stellar mass, A_V, SFR, sSFR, and mass-weighted age. The grey histogram plots all galaxies analyzed for which a model was successfully fitted and a parameter output file generated. The red histogram includes only those galaxies flagged as ELGs. For clarity, galaxies that had values of log(sSFR/yr^{-1}) $\leq$ -12 were binned. Galaxies with log(SFR/(M$_{\odot}$ yr^{-1})) $\leq$ -2 were also binned for clarity.



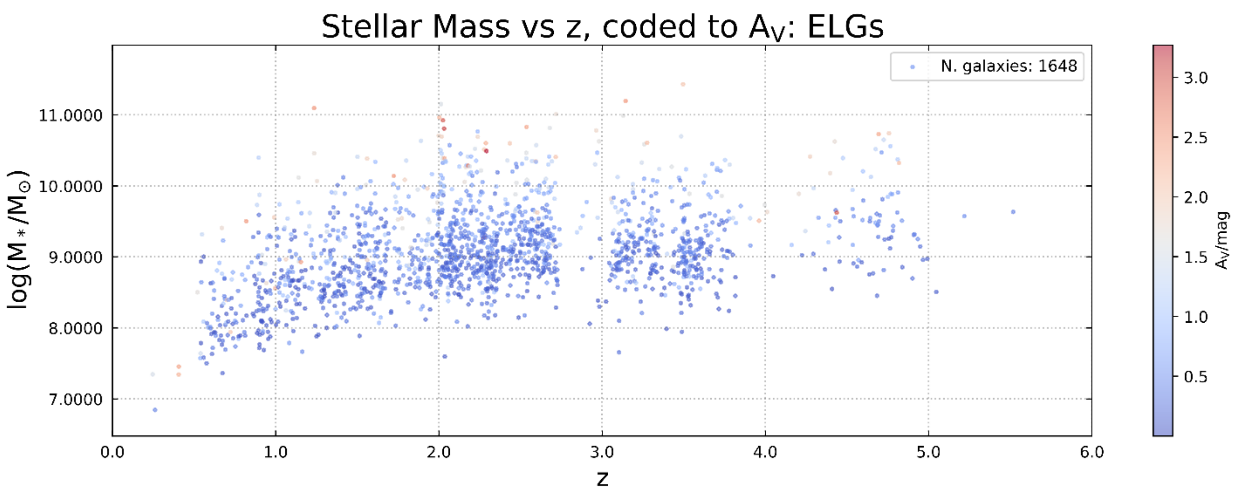


Figure 18: The plots of stellar mass versus redshift, color-coded to A_V. The upper plot includes all galaxies, while the lower plot shows only those flagged as ELGs.

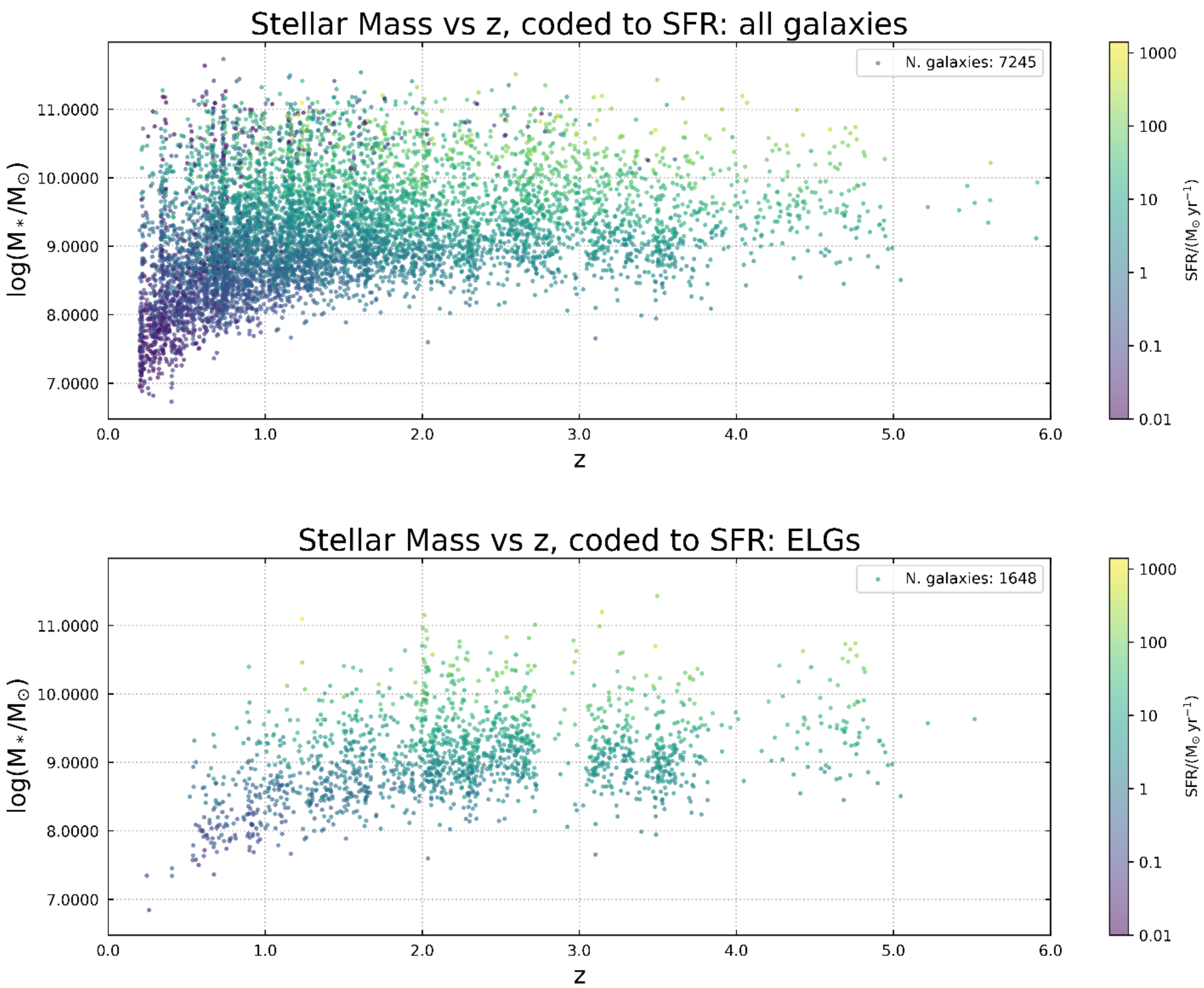


Figure 19: The plots of stellar mass versus redshift, color-coded to SFR. The upper plot includes all galaxies, while the lower plot shows only those flagged as ELGs. The SFRs of the characterized galaxies span six orders of magnitude; as a result, the color-coding on these plots is logarithmic.

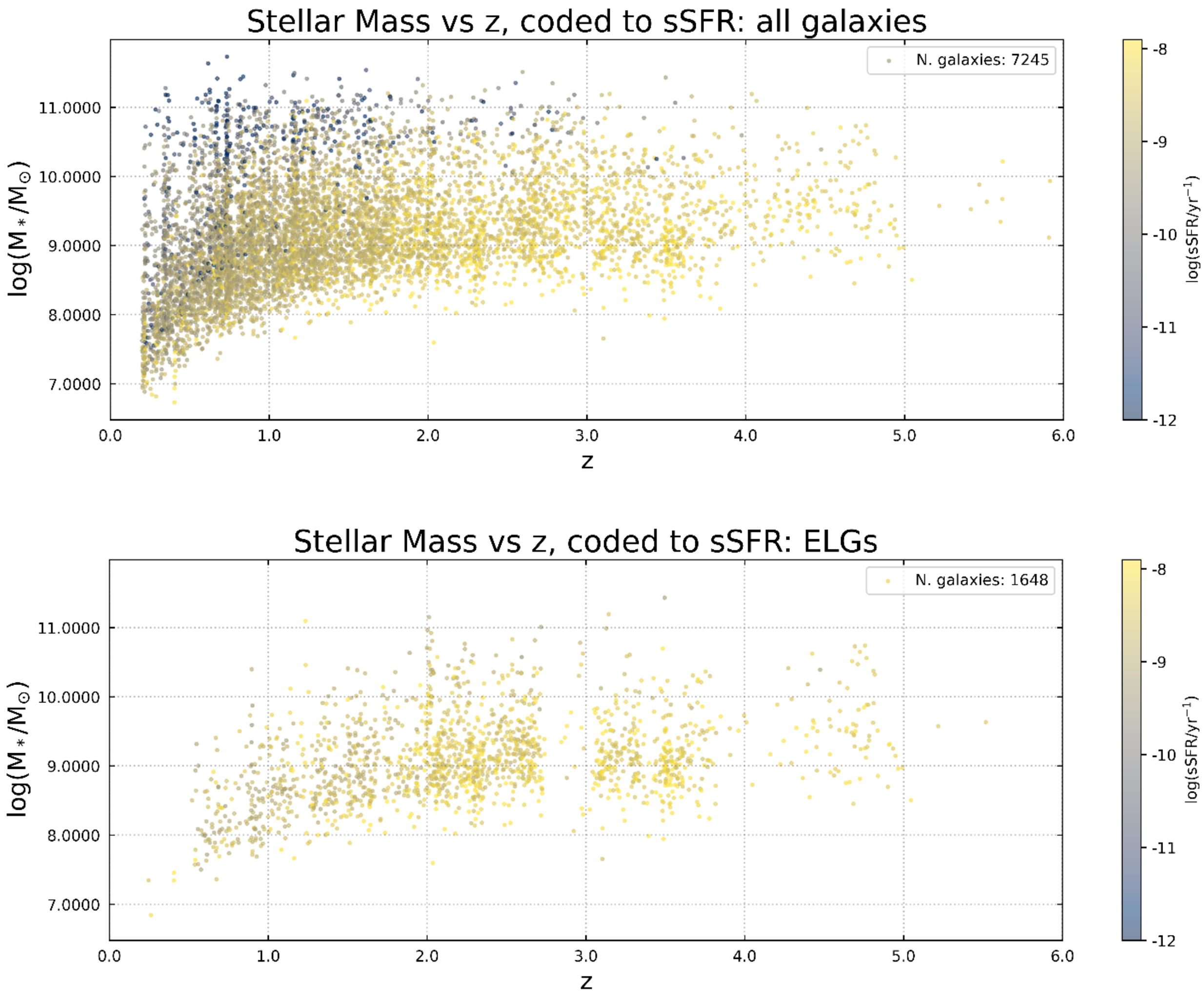


Figure 20: The plots of stellar mass versus redshift, color-coded to sSFR. The upper plot includes all galaxies, while the lower plot shows only those flagged as ELGs. As with SFR, we present the sSFR in logarithmic scaling because the data spans more than five orders of magnitude. For clarity, we set the log(sSFR/yr^{-1}) of all galaxies with log(sSFR/yr^{-1}) < -12 to -12.

Examining Figure 17 shows certain trends in the ELGs that are not reflected in the complete galaxy sample. The stellar mass shows similar trends in both sets, but the galaxies flagged as ELGs generally have log(sSFR/yr^{-1}) > -9, while non-ELG galaxies can have substantially lower values of log(sSFR/yr^{-1}). Similarly, the fraction of galaxies in the dataset flagged as ELGs goes up as SFR increases; as star-forming galaxies tend to also be ELGs, this trend agrees well with our expectations. Note also that most ELGs have mass-weighted ages of less than about 2 Gyr, meaning that these galaxies formed most of their stars more recently than non-ELGs. This, too, agrees with expectations, as star-forming galaxies will have many more young stars than older, quenched galaxies which long ago exhausted their gas supplies or otherwise found a way to quench star formation.

Examining Figures 18, 19, and 20 reveals some of the limitations of our methods. In particular, the plot color-coded to SFR shows that there are some galaxies that BAGPIPES assigns appreciably high SFR values to that are not flagged as ELGs. Notice that the potential ELGs that the algorithm failed to detect are clustered in specific redshift bands; for example, between $z = 2.7$ and $z = 3.0$, there is a cluster of galaxies with SFRs exceeding 10 $M_{\odot}$/yr that are nonetheless absent from the plot of galaxies flagged as ELGs. Such redshift regimes indicate regions where emission lines have not been redshifted into any of the seven medium band IR filters analyzed here; we could likely recover these galaxies as ELGs if we extended our analysis to include filters that target emission lines at these redshifts. Alternatively, some of these galaxies may be highly active star-formers and yet not be ELGs. This is the case for galaxies heavily obscured by dust, as obscuration from dust can suppress the intensity of the emission lines relative to the stellar continuum, reducing the value of Δm below our cutoff threshold even though the galaxy is forming many new stars.

The accuracy of our results is constrained in part by the reliability of our data. Our analysis required each analyzed galaxy to have SNR > 5 as part of the selection criteria, which means that each photometric data point carried a fractional uncertainty of 20% or less. Our SED posteriors tended to reflect this uncertainty, with the range of generated spectra (the percentiles as shown in Figure 4) being broadest around data points that carried the highest uncertainties. In choosing targets for follow-up studies, we could set stricter constraints on the SNR. However, the SNR of a galaxy in one survey does not necessarily set its SNR in future missions, and the SNR of a galaxy covered in a deep-sky survey is certainly going to be lesser than its SNR in a dedicated follow-up. Thus, constraining SNR >> 5 is most likely unnecessary and would only discard valuable targets.

In this analysis, we restricted our analysis of the Δm plots to the medium band filters, which are K_{red}, K_{blue}, J_1, J_2, J_3, H_1, and H_2. We only used the other filters in our dataset for SED fitting purposes. In the near future, it could be worthwhile to broaden our Δm plot analysis to include some of these filters, extending the reach of our analysis into different redshift ranges not yet well-sampled by our methods. However, using broadband filters does introduce a selection bias favoring ELGs with more intense emission lines, compared to the use of medium band filters which are better able to detect less extreme ELGs.

5.2 Conclusions and Future Directions

In this thesis, we presented the preliminary analysis of a dataset combining previous medium band galaxy surveys with the newly-obtained data from the FENIKS pilot survey. Our analysis suggests that out of the 7,280 galaxies we analyzed, 1,648 galaxies are ELGs with emission line equivalent widths in excess of 100 Å. The code developed for this project represents a useful tool that can be used to select galaxies from the fields surveyed by FENIKS at desired levels of star formation activity; the galaxies thus selected can be targeted in detailed spectroscopic follow-up studies designed to more rigorously constrain the properties of each galaxy. In the short-term, several underutilized features of our program will be expanded upon to make the selection process more rigorous and efficient, allowing us to confidently select the best targets for detailed spectroscopic studies as well as further validate our methods. In the long-term, follow-up missions informed by this analysis will yield insight into the star formation mechanisms of these early, fastgrowing galaxies, bringing us closer to determining how the universe's early galaxies grew and quenched so quickly.

BIBLIOGRAPHY

Conroy, C. 2013, Annual Review of Astronomy and Astrophysics, 51, 393. doi:10.1146/annurev-astro-082812-141017

Esdaile, J., et al. 2021, The Astrophysical Journal, 162, 225. doi:10.3847/1538-3881/ac2148

Carnall, A. C., McLure, R. J., Dunlop, J. S., et al. 2018, Monthly Notices of the Royal Astronomical Society, 480, 4379. doi:10.1093/mnras/sty2169

Bruzual, G. & Charlot, S. 2003, Monthly Notices of the Royal Astronomical Society, 344, 1000. doi:10.1046/j.1365-8711.2003.06897.x

Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, Revista Mexicana de Astronomía y Astrofísica, 53, 385.

Draine, B. T. & Li, A. 2007, The Astrophysical Journal, 657, 810. doi:10.1086/511055

Inoue, A. K., Shimizu, I., Iwata, I., et al. 2014, Monthly Notices of the Royal Astronomical Society, 442, 1805. doi:10.1093/mnras/stu936

Kroupa, P. & Boily, C. M. 2002, Monthly Notices of the Royal Astronomical Society, 336, 1188. doi:10.1046/j.1365-8711.2002.05848.x

Skilling, J. 2006, Bayesian Analysis, 1, 833.

Marchesini, D., Whitaker, K. E., Brammer, G., et al. 2010, The Astrophysical Journal, 725, 1277. doi:10.1088/0004-637X/725/1/1277

Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, The Astrophysical Journal Supplement, 206, 8. doi:10.1088/0067-0049/206/1/8

Straatman, C. M. S., Spitler, L. R., Quadri, R. F., et al. 2016, The Astrophysical Journal, 830, 51. doi:10.3847/0004-637X/830/1/51

Forrest, B., Annunziatella, M., Wilson, G., et al. 2020, The Astrophysical Journal Letters, 890, L1. doi:10.3847/2041-8213/ab5b9f

Saracco, P., Marchesini, D., La Barbera, F., et al. 2020, The Astrophysical Journal, 905, 40. doi:10.3847/1538-4357/abc7c4

Brammer, G., van Dokkum, Pieter G., & Coppi, P. 2008, The Astrophysical Journal, 686, 2. doi:10.1086/591786

Kewley, Lisa J., Nicholls, David C., & Sutherland, Ralph S. 2019, Annual Review of Astronomy and Astrophysics, 57, 511-570. doi:0.1146/annurev-astro-081817-051832

Nell, B., Dalcanton, Julianne J., Conroy, C., and Johnson, Benjamin D. 2017, The Astrophysical Journal, 840, 1. doi:10.3847/1538-4357/aa6c66

Hogg, David W. et al. 2002, arXiv, astro-ph/0210394

Appendix A. Simulations of EW Dependency on log(U)

In Chapter 4.3, we claimed that the effects of log(U) on the W_{RS}-Δm relationships were entirely a result of the filters being wider than the emission line itself, thus causing Δm to be calculated from parts of the galactic spectrum that included features other than the emission line. We now present the additional simulations conducted which led us to draw this conclusion. We generated a hypothetical thin filter centered on 22,965 A and targeting galaxies at a redshift of z = 2.5; at this redshift, the Hα line is centered at 22,969 A and is thus acutely targeted by the hypothetical thin filter. Figure 1A below shows the hypothetical thin filter curve in grey, plotted over the actual K_{red} filter curve used in our study.

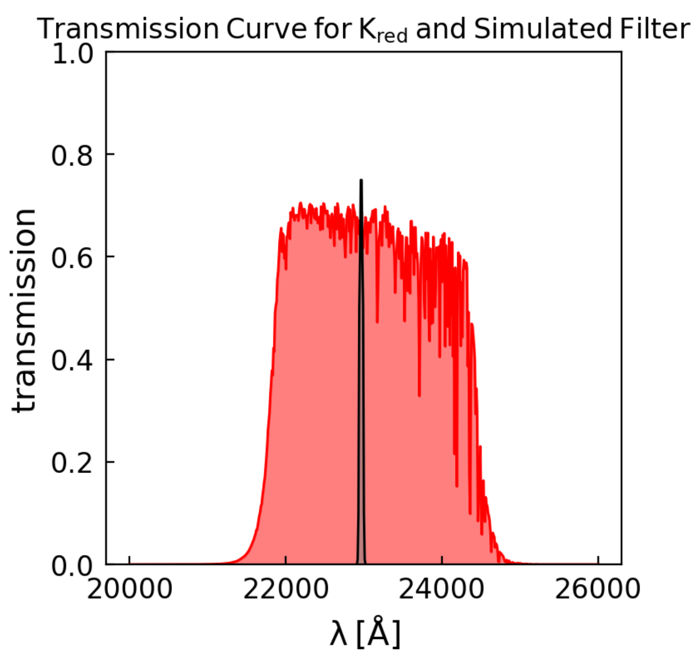


Figure 1A: The hypothetical filter curve (grey) against the actual K_{red} curve (red). The hypothetical thin filter's width is orders of magnitude thinner than the K_{red} curve and is designed to sample only the flux of the Hα emission line at a redshift z = 2.5.

We run the same set of simulations against this hypothetical figure as we did in Chapter 4.3 against our real filters. Figure 2A shows the resultant rest-frame Hα equivalent width vs. Δ*m* plots generated using the hypothetical thin filter.

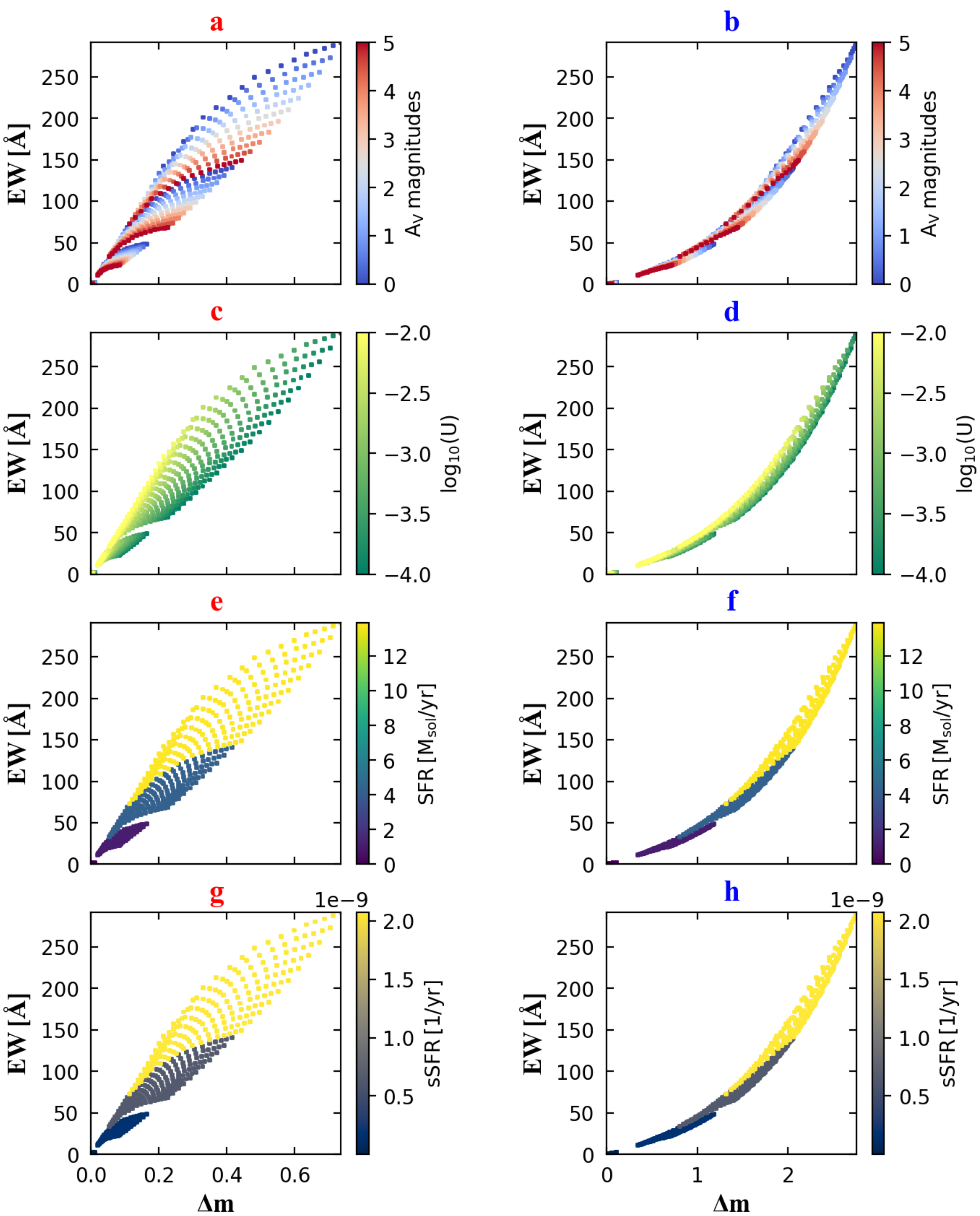


Figure 21: The simulated W_{RS} vs. Δ*m* plots for the K_{red} filter (left column) and for the hypothetical thin filter (right column). The color-coding in these plots is identical to that of Figure 7.

While most of the relationships between W_{RS} and Δm are preserved even in the thin filter case (W_{RS} increases with SFR and sSFR, dust reduces the observed Δm when W_{RS} is fixed), notice that the log(U) dependency has almost completely collapsed. There is still some dispersion of the W_{RS}-Δm curve as a result of the thin filter's finite width, but compared to the much broader K_{red} curve, the effects of log(U) on the W_{RS}-Δm relationship are much diminished. We conclude from these simulations that W_{RS} is largely insensitive to the value of log(U), a result that agrees well with other studies in nebular emission modelling (e.g., Nell et al. 2017).

Appendix B. Remarks on the Strength of the *K*-split Filters

The FENIKS pilot survey and the full FENIKS survey both aim to improve spectral resolution in the *K* band by introducing the *K*-split filters, K_{blue} and K_{red}. We briefly remark here on the effectiveness of these filters in improving spectral resolution by presenting Figures 1B and 2B below. These figures show the change in magnitude Δm detected in each of the medium filters for each galaxy versus the value of Δm measured in the conventional broadband K_s filter.

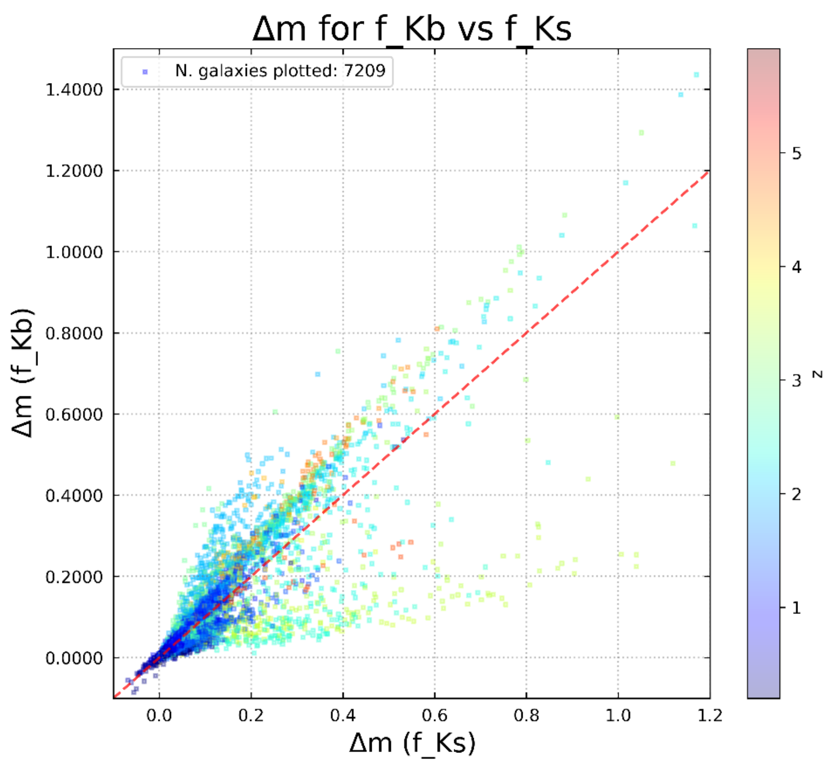


Figure 1B: The plot of Δm in K_{blue} versus Δm in K_s, color-coded to redshift. The red dashed line marks the line of equal performance between K_{blue} and K_s; above this line lie points where the medium filter outperformed the broad filter. In the redshift bands where both filters target emission lines (e.g., $z \sim 2.0$), we expect higher Δm values to be recovered in the medium filter.

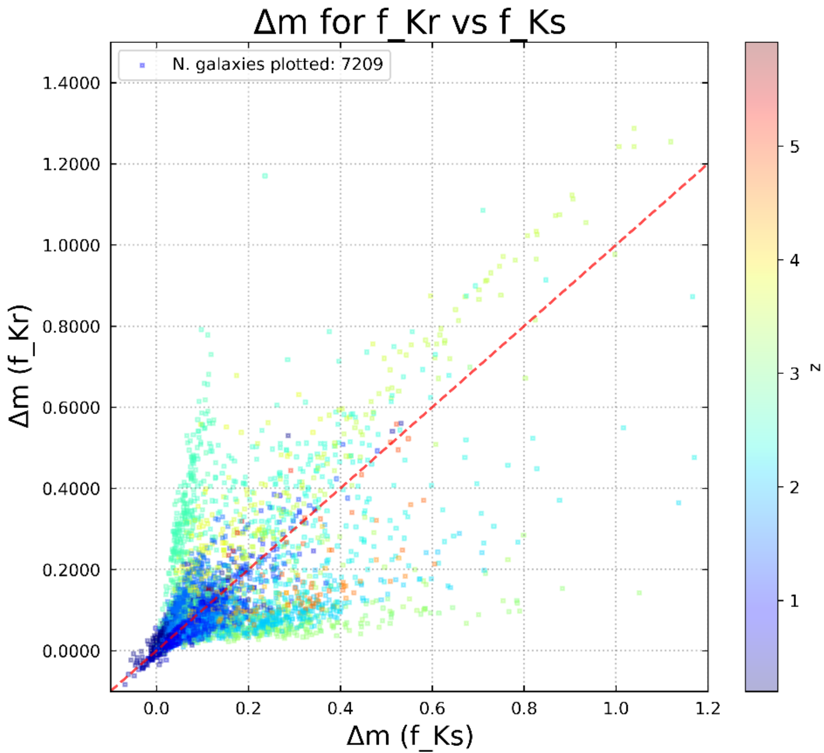


Figure 2B: The plot of Δ*m* in K_{red} versus Δ*m* in K_s, color-coded to redshift. The red dashed line marks the line of equal performance between K_{red} and K_s; above this line lie points where the medium filter outperformed the broad filter. In the redshift bands where both filters target emission lines (e.g., $z \sim 2.5$), we expect higher Δ*m* values to be recovered in the medium filter.

In redshift ranges where neither K_s nor the *K*-split filters target emission lines, we expect the filters to perform similarly, not because the filters have similar resolving power, but because there are few to no spectral features present to resolve. This is reflected in the central line of data points (not the dashed red line) present in both Figures 2A and 2B; this line of scatter points, which has a slope of just over 1, shows the broadband and medium filter performing comparably, with the medium filter just slightly outperforming the broadband filter. On the other hand, in redshift ranges where both the broadband filter and one of the *K*-split filters target emission lines, we expect the *K*-split filter to recover higher Δ*m* values than the conventional broadband filter. This is reflected in both figures as a sharp spur jutting upwards to the left of the central line. These spurs show that in the appropriate redshift bands, the *K*-split filters outperform the conventional filter by

+0.2 Δm or more. Because the *K*-split filters are narrower, there will be redshift ranges where the K_s filter still targets emission lines that have been redshifted out of the range covered by the medium filter. In these regions, we expect the broadband filter to outperform the medium filter. This is reflected in each plot as a sharp spur jutting rightwards, located just beneath the central line. Notice that the color-coding on each plot indicates redshift. In the regions where K_{red} outperforms K_s, K_s tends to outperform K_{blue}, and vice versa; that is, the lower spur in one plot corresponds to the same redshift range as the upper spur in the other plot. For example, the upper spur for K_{blue} (where it outperforms the broadband filter) targets redshifts $z \sim 2.0$, the same approximate redshift band that makes up the lower spur for K_{red} (where it is outperformed by the broadband filter). This is because the emission lines visible in K_{blue} at $z \sim 2.0$ are also visible in K_s at that same redshift, but they have not yet been redshifted into the z range at which they can be targeted by K_{red}.

ACADEMIC VITA

Academic Vita of Victoria Boehm
vab5187@psu.edu

Education
Major(s) and Minor(s): B.S. Astronomy and Astrophysics, B.S. Physics, Minor Mathematics
Honors: Dean's List (FA 2019 – present)

Thesis Title: Identifying Emission Line Galaxies in the FENIKS Pilot Survey via Spectral Energy Distribution Fitting with BAGPIPES
Thesis Supervisor: Danilo Marchesini

Work Experience
September 2018 – Present
Undergraduate research student
Conducted research on the electronic properties of strained graphene and BSCCO using scanning tunnelling microscopy.
The Pennsylvania State University, University Park, PA 16802
Eric Hudson

January 2022 – May 2022
Tutor for PHYS 211/212/212H
Hosted office hours for freshman-level physics courses at Penn State.
Assisted students with understanding classical physics concepts such as kinematics, dynamics, momentum, energy, electrostatics, circuits, and magnetism.
The Pennsylvania State University, University Park, PA 16802
Adidi Etim-Hunting

September 2021 – December 2021
Tutor for MATH 140H
Hosted office hours for freshman-level math courses at Penn State.
Assisted students with developing the foundation in algebra, trigonometry, and pre-calculus needed to succeed in college calculus courses.
The Pennsylvania State University, University Park, PA 16802
Adidi Etim-Hunting

June 2021 - Present
Summer researcher
Conducted research on the evolution of star-forming galaxies at redshifts 0 to 6 as observed by the FENIKS pilot survey.
Tufts University, Medford, MA 02155 (remote)

Danilo Marchesini

August 2020 – May 2021
Learning assistant for ASTRO 291/292
Hosted office hours for sophomore-level astrophysics courses at Penn State. Reinforced students' understanding of course concepts.
Assisted students with approaching and solving homework problems and exams.
The Pennsylvania State University, University Park, PA 16802
Donald Schneider

May 2019 – July 2019
Summer researcher
Conducted research at the undergraduate level on the distribution and size of starspots on the surface of the rotating variable star LO Pegasi.
Ohio Wesleyan University, Delaware, OH 43015
Robert Harmon

Awards:

Penn State Millennium Scholars Program Scholarship: FA 2019 – present

Professional Memberships:

American Astronomical Society: 2019-present

Sigma Pi Sigma Honor Society, Third Chapter: 2021-present

Presentations:

V. A. Boehm, A. Rishi, and R. Harmon. "Imaging Starspots on LO Pegasi, 2014-2019". Eberly College of Science Undergraduate Research Exhibition. Sep. 2019

V. A. Boehm and D. Marchesini. "Identifying Emission Line Galaxies from the FENIKS Pilot Survey via Spectral Energy Distribution Fitting with BAGPIPEs". Eberly College of Science Undergraduate Research Exhibition. Oct. 2021

V. A. Boehm, A. Rishi, and R. Harmon. “Imaging Starspots on LO Pegasi, 2014-2019”. In: *American Astronomical Society Meeting Abstracts #235*. Vol. 235. American Astronomical Society Meeting Abstracts. Jan. 2020, p. 273.04.

Community Service Involvement:

January 2019 – January 2021
Millennium Society general body member
Participated in STEM outreach events such as ENVISION 2020 and Haunted U, with the goal of engaging local State College residents in science demonstrations and careers in STEM.
The Pennsylvania State University, University Park, PA 16802
Millennium Society