The Emission Ratio of the [O III] $\lambda\lambda$5007, 4959 Å Doublet with Redshift

Mayalen A. Laker

A senior thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Bachelor of Science

J. Ward Moody and Benjamin D. Boizelle, Advisors

Department of Physics and Astronomy

Brigham Young University

April 2022

ABSTRACT

The Emission Ratio of the [O III] $\lambda\lambda$5007, 4959 Å Doublet
with Redshift

Mayalen A. Laker
Department of Physics and Astronomy, BYU
Bachelor of Science

As a test of the immutability of forbidden-line emission processes with time, we measured the [O III] $\lambda\lambda$5007, 4959 emission lines of 12,241 galaxies from the Sloan Digital Sky Survey Data Release 8 to examine the [O III] $\lambda\lambda$5007, 4959 photon flux ratio as a function of redshift z. We fitted each line with a Gaussian curve and rejected any spectra not conforming to strict requirements of line symmetry, signal-to-noise levels, and continuum fit. We determined that a linear fit is not sufficient to fit the underlying continuum and that a higher-order polynomial or Legendre polynomial should be used for this purpose. We found the ratio of energy flux to be constant for galaxies having $0 < z < 0.467$ (corresponding to a look back time of approximately five billion years) with no significant evidence for a dependence with z.

Keywords: Cosmological evolution, Atomic physics, Emission line galaxies

ACKNOWLEDGMENTS

I am deeply grateful to my advisor J. Ward Moody for assistance and unwavering support throughout the research project. I would like to express my sincere gratitude to my fellow researcher, Conner Camacho, for his large contributions to the project, particularly with the code. I would also like to offer my special thanks to my new advisor, Benjamin D. Boizelle, for jumping into the project and offering his support. Lastly, I am very grateful to my excellent parents for the moral support they have provided through this journey.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The forbidden line doublet of doubly-ionized oxygen ([O III]), located at 5007 and 4959 Å, is among the most prominent emission features in extra-galactic nebulae. This doublet can be found in the spectra of star-forming galaxies, quasars, and Seyfert galaxies. The presence of these lines indicates a high temperature and a low density in the nebular gas, as explained in Section 1.1.

1.1 The Origin of the [O III] $\lambda\lambda$5007, 4959 Å Doublet

[O III] $\lambda\lambda$5007, 4959 Å doublet is formed when an electron in the $2s^2\ 2p^2\ {}^1D_2$ orbital, which has no permitted transitions away from it, drops to either the ${}^3P_1$ or ${}^3P_2$ state, emitting either a 5007 Å or 4959 Å photon, respectively. The probabilities of spontaneous emission of each transition are set by the properties of the atom, specifically the Einstein spontaneous emission coefficients. The value of this coefficient is 2.0×10^{-2} s^{-1} for the 5007 Å transition and 6.7×10^{-3} s^{-1} for the 4959 Å transition, leading to characteristic times $\tau = 1/A$ of $\tau = 50$ s and $\tau = 149$ s, respectively [11]. The details of these transitions can be seen in the O III energy level diagram shown in Figure 1.1. Due to large number statistics, the approximately three times greater probability of the 5007 Å transition means that it occurs approximately three times more often than the 4959 Å transition.

Thus, the integrated flux for the 5007 Å transition is expected be approximately three times greater than that of the 4959 Å transition. The ratio of the 5007 Å integrated flux to the 4959 Å integrated flux is what is hereafter referred to as the [O III] $\lambda\lambda$5007, 4959 Å flux ratio, or simply [O III] flux ratio or flux ratio.

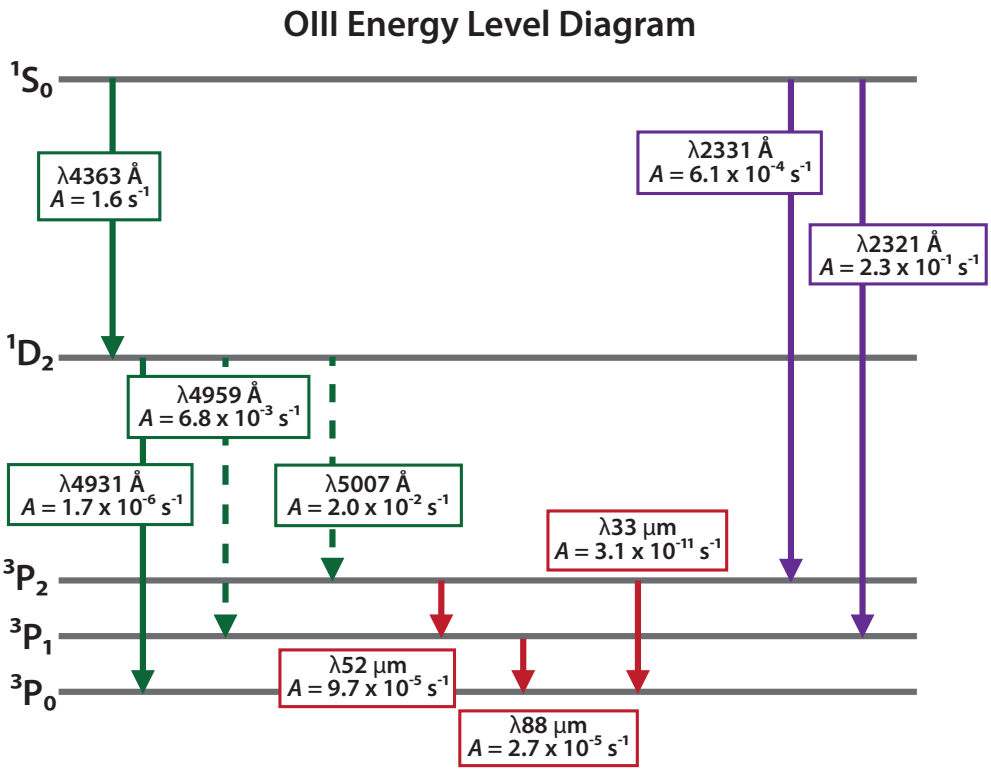

Figure 1.1 Energy level diagram for O III, showing the lowest energy states. Based on Figure 3.1 and other information in Osterbrock and Ferlund (2006) [11]. The green arrows represent optical transitions while the red and purple arrows represent transitions in the infrared and ultraviolet ranges, respectively. The two dashed arrows show the transitions of interest in this work.

Specific conditions must be in place for this emission doublet to form. In order for high enough concentrations of O III ions to be present to create strong emission lines, the temperature must be

in excess of about 10,000 K. Much hotter temperatures, however, will cause additional ionization and lead to a decrease in O III density as the O IV density increases. Additionally, overall density is a major factor in the doublet line strength. Collisions will de-excite the electrons out of the $^1D_2$ state if the density reaches the critical density, $n_e = 6.8 \times 10^5$ cm^{-3} at 10,000 K[11]. The doublet only emerges when the characteristic time between spontaneous emission processes is less than the average time between collisions.

As stated, the formation conditions affect the strength of the doublet lines, but the *ratio* should not vary from galaxy to galaxy or throughout time, as it is set by the Einstein A coefficient, which arises from the properties of the O III ion. As the O III ion is assumed (almost certainly correctly) to be identical wherever it appears, the Einstein A coefficients of its transitions should be identical as well. Thus the ratio is expected to be unchanging at all points in space and time.

Observational and theoretical studies have found several different values for the [O III] flux ratio, centering around about 2.98, which is the value obtained by taking the ratio of the Einstein A coefficients of the doublet lines. Measurement and instrument error in attaining extra-galactic spectra makes it difficult to find a precise value, and due to the extremely low densities required to form the forbidden emission doublet, the emission cannot be studied in an Earth-based laboratory. However, many previous studies have examined this value in depth, using many thousands of spectra in total, to find the agreed-upon value of approximately 2.98 [7] (shown in Table 1.1), which is the theoretical value as well [15].

1.2 The Accepted Cosmology

In the currently accepted cosmology, the fundamental constants are assumed to have been constant since shortly after the Big Bang. This is likely a very good assumption, but it is prudent to always be aware of assumptions made, and, when possible, to test them.

Table 1.1. Previous [O III] Flux Ratio Studies

Author(s)	Year	Ratio Value	Study Type	Samples
Rosa [14]	1985	3.03 ± 0.03	Obs.	
Leisy & Dennefeld [10]	1996	3.00 ± 0.08	Obs.	31
Galavis et al. [8]	1997	2.89	Theor.	-
Storey & Zeippen [15]	2000	2.98	Theor.	-
Bahcall et al. [2]	2004	2.99 ± 0.02	Obs.	165
Dietrich et al. [6]	2005	2.953 ± 0.014	Obs.	12
Dimitrijevic et al. [7]	2006	2.993 ± 0.014	Obs.	34

Previous tests have been done to show the consistency of physical constants such as the fine structure constant $\alpha = e^2/\hbar c$, which is related to the fine structure of atomic emission lines. Bahcall et al. (2004) [2] used the energy separation of the [O III] $\lambda\lambda$5007, 4959 Å doublet in 165 quasi-stellar objects (QSOs or quasars) with redshifts $0.16 < z < 0.80$ to constrain the evolution in the fine structure constant to be $\Delta\alpha/\alpha(0) = (0.7 \pm 1.4) \times 10^{-4}$, where $\alpha(0)$ is evaluated locally, or at $z = 0$. This analysis has been repeated several times with larger data sets to give $\Delta\alpha/\alpha(0) = (+2.4 \pm 2.5) \times 10^{-5}$ in 2010 with 1,568 QSO spectra with $z < 0.8$ [9], $\Delta\alpha/\alpha(0) = (-2.1 \pm 1.6) \times 10^{-5}$ in 2014 based on 2,347 QSOs with $0.02 < z < 0.74$ [13], and $\Delta\alpha/\alpha(0) = (0.9 \pm 1.8) \times 10^{-5}$ in 2015 with a sample of 13,175 QSOs of $z < 1$ [1]. Other studies that investigated the consistency of α include work done analyzing the Oklo natural fission reactor [12], using absorption lines from metals in QSOs [18], [19], and using strong gravitational lensing and Type Ia supernovae observations [5].

With this precedence, confirming that the [O III] doublet ratio was constant in the past is a

check that the constants that set the ratio, namely the Einstein spontaneous emission coefficients of the ion's transitions, have not varied. A change in these parameters would imply a change in the fundamental physics that sets their values, which would necessarily invite the possibility of the evolution of other physical constants through cosmic time. Though this seems odd, several theories, somewhat disfavored by the scientific community, state that this could be the case (see Uzan (2011) [17] for a thorough review). It therefore becomes necessary to provide evidence for or against these, as the far-reaching implications of the scientific cosmological model impacts nearly all parts of astrophysics.

1.3 Project Overview

This project aims to show through observations of star-forming galaxies that the [O III] $\lambda\lambda$5007, 4959 Å flux ratio and thus the corresponding Einstein A coefficients have been constant for at least the last approximately five billion years (to $z = 0.467$). This is accomplished by measuring the emission line fluxes of the [O III] doublet for over 12,000 star-forming galaxies and finding no correlation between flux ratio and redshift.

Compared to previous studies, we use a sample size two orders of magnitude larger than what has been done previously. While that means that we cannot give each spectrum as much of an individual treatment, the large sample size drastically reduces the random error in the study and tightens the formal error bars. We use star-forming galaxies instead of the active galactic nuclei (AGN) previous studies have used, which gives us the advantage of having fewer highly-ionized spectral lines to work around during the fitting process.

We did not expect a statistically significant non-zero trend in the [O III] flux ratio against increasing redshift. A significant non-zero trend would suggest that the ratio has evolved with time and that the Einstien spontaneous emission coefficients which determine the intrinsic flux ratio

have also evolved. If this were the case, the constant nature of physical constants in the favored cosmology would need to be re-examined. Conversely, the absence of a significant non-zero trend in the [O III] flux ratio against redshift would support the presumption that there has been no change in the Einstein spontaneous emission coefficients, and we would find no evidence against the favored cosmology.

Chapter 2

Methods

In order to use the [O III] $\lambda\lambda$5007, 4959 Å flux ratio as a test of the constant nature of the Einstein A coefficients and other physical constants, we needed to obtain and carefully analyze data showing strong [O III] emission. This chapter explains the methodology. Section 2.1 describes the data used for the study and Section 2.2 details the analysis we performed.

2.1 Data

Data for this project were obtained from the Sloan Digital Sky Survey (SDSS) Data Release (DR) 8 archive[1], taken with the 2.5 m telescope at Apache Point Observatory in Sunspot, New Mexico. All spectra were flux and wavelength calibrated by the SDSS or "MPA-JHU" pipeline, described by Stoughton et al. [16]. Each galaxy spectrum is tabulated with photon energy flux (hereafter flux) densities in units of 10^{-17} erg cm^{-2} s^{-1} Å^{-1} and spectral binning of about 1 Å. During the pipeline processing, unreliable or absent measurements were flagged for easy removal.

A list of star-forming galaxies, the population of interest, was obtained from the SDSS website[2].

[1] URL: *http://sdss3.org/dr8/*

[2] The online file, *galSpecLine-dr8.fits*, is available at *https://www.sdss.org/dr12/spectro/galaxy_mpajhu/*

Star-forming galaxies are desirable because they contain many hot, young stars that produce sufficiently energetic photons to ionize oxygen from O I to O III. While many AGNs meet this condition, they often produce higher-energy radiation that will excite other elements with spectral features in the [O III] doublet region, notably iron. Type I AGNs have the further disadvantage (for our purposes) that the Doppler-broadened line emission, especially Hβ, from the broad line region is "smeared" out to and beyond the [O III] doublet wavelengths, leading to difficulty fitting the emission lines and lower overall confidence in the [O III] flux ratio. Thus, we chose star-forming galaxies as the type of object whose spectra we would analyze.

We obtained the list of galaxies with identified emission lines from SDSS DR8. In all, it listed 22,389 spectra within the redshift range of $0 < z < 0.7$, where $z+1 = \lambda_{\text{obs}}/\lambda_{\text{emit}}$, including duplicated entries. The tabulated emission-line parameters, reasonably accurate for the studies of star formation, stellar mass, and oxygen abundances for which they were intended, were deprecated in favor of a reanalysis by a team from Wisconsin, Portsmouth, and Granada using data from DR12 [4]. A well-behaved example spectrum is shown in Figure 2.1.

A preliminary trend of the [O III] ratio with redshift was obtained by our group while working on a separate project[3], and it was determined that higher precision could be obtained in the calculated ratio. We determined through manual measurements using `IRAF`[4] that more accurate values for the λ4959 and λ5007 emission-line fluxes could be obtained, largely through more precisely locating the continuum around these specific lines. We therefore re-measured all line fluxes with Python code written specifically for the $\lambda\lambda$5007, 4959 lines.

[3]That project used the flux of the [O III] doublet compared to the flux of the [O III] 4363 Å emission line as a tracer of temperature [3].

[4]The Image Reduction and Analysis Facility (`IRAF`) is a well-known collection of software used commonly by astronomers to process optical telescope data.

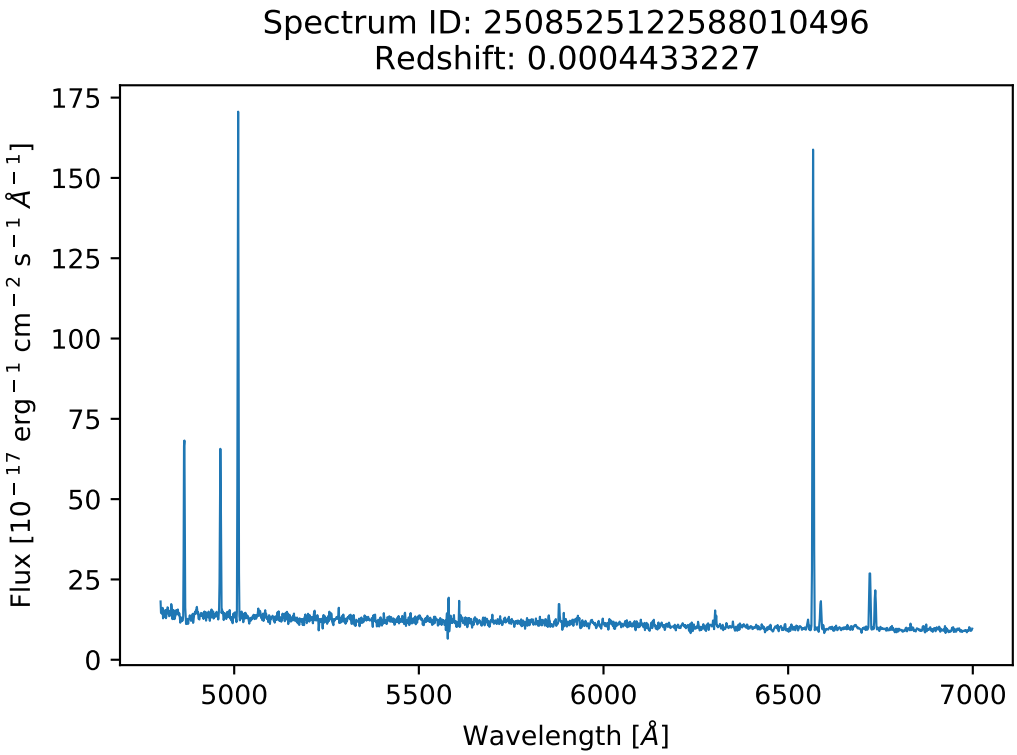

Figure 2.1 Example spectrum obtained from SDSS, which has been run through an improved calibration pipeline [16]. On the left are the Hβ 4861 Å and [O III] $\lambda\lambda$4959, 5007 Å emission lines. On the right, the [N II] $\lambda\lambda$6548, 6583 Å doublet, which surrounds the Hα 6563 Å line, and the [S II] $\lambda\lambda$6716, 6731 Å doublet are seen. The continuum appears to be approximately linear, with relatively small levels of noise throughout. This spectrum shows strong, high SNR [O III] doublet emission while the noise in the continuum is clearly visible.

2.2 Flux Measurement

The precise measurement of the emission line flux was the main challenge of this project. As indicated, we found that we could measure the emission lines fluxes more precisely ourselves with specialized code written for the task than using the values provided in the file discussed in the previous section. This was done by modeling each of the doublet emission lines separately from the continuum around that region. By creating our own code to describe those lines specifically, we were able to increase precision compared to the SDSS pipeline calculations.

2.2.1 Emission Line Modeling

To measure line flux, we wrote a fitting algorithm in Python using the `curve-fit` method inside the `Scipy optimize` package, which is based on non-linear least-squares fitting. The region of interest includes the 5007 and 4959 Å [O III] doublet lines, the 4861 Å Hβ line, and the continuum found on either side of this region. Because of its close proximity, a broad Hβ line can elevate the continuum between it and the 4959 Å [O III] line, making it unsuitable for use. Figure 2.2 shows a close-up of this region of interest for the spectrum plotted in Figure 2.1. Previous spectroscopic work in many sub-disciplines of astronomy have shown that emission lines can be modelled well by Gaussian, Lorentzian, Voigt, and Moffat profiles. Each of these profiles, as well as a sum of the Gaussian and Moffat profiles, were fit to the lines of each spectrum in the data set. We found that a Gaussian profile was the simplest function that gave robust solutions, in most cases, while avoiding difficulties from the model fit diverging when new parameters were added. We also found that the Gaussian profile gave us the smallest average sum-squared error,

$$\sigma_{\mathrm{SSE}} = \sum_i (\mathrm{data}_i - \mathrm{model}_i)^2, \tag{2.1}$$

of the fits. Because of this, and for computational ease, we used a Gaussian profile exclusively.

The code fit Gaussian profiles to both lines in the [O III] doublet. We estimated initial parameter guesses for the centroid and line width of the Hβ, λ4959, and λ5007 lines in each spectrum from the wavelength of their maximum value and their full width at half maximum (FWHM), respectively, and used them to seed the algorithm. When it converged, the algorithm returned a resultant amplitude, standard deviation, mean, and continuum offset.

The measured emission line flux was found to be highly dependent on the fit of the underlying continuum. A correct treatment of the continuum beneath the emission lines was critical to determining their ratio. Due to the width of the emission lines, we deemed it unwise to fit the continuum between them and had to rely on interpolation of data outside of the emission line

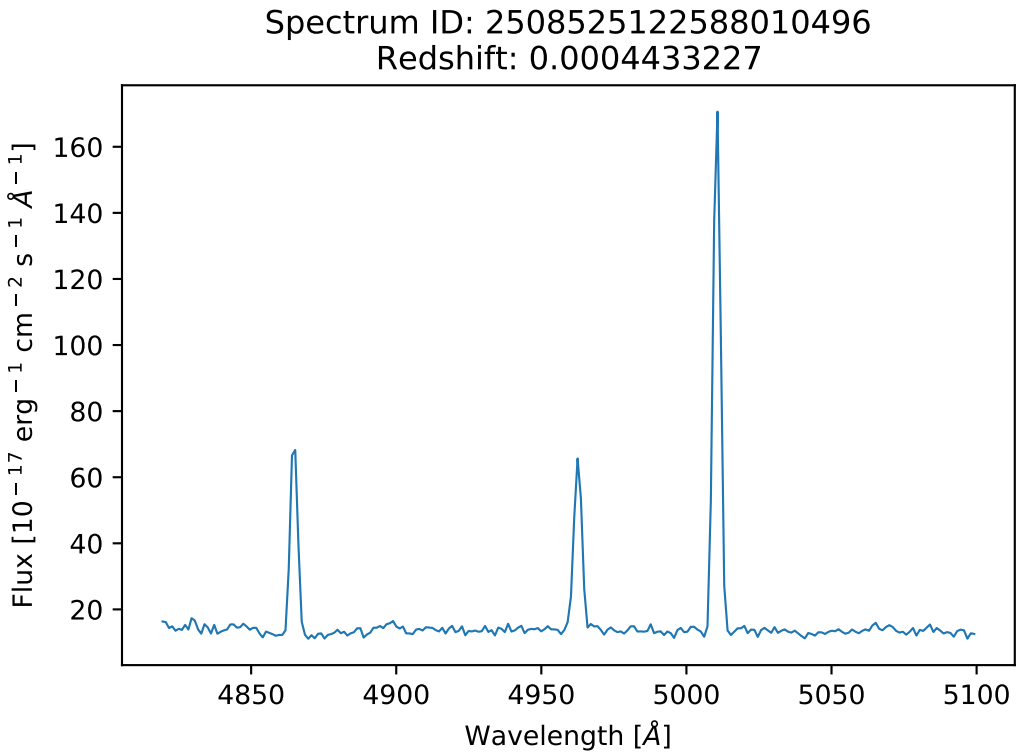

Figure 2.2 The same spectrum as Figure 2.1 plotted on a narrower wavelength range to highlight the Hβ λ4861 Å, [O III] λ4959 Å, and [O III] λ5007 Å emission lines. Once again, the noise in the continuum can be seen. Visually, the continuum appears to be very linear here.

region for our fitting. As a first attempt, we applied a linear estimation by fitting a line through the continuum 20 Å to 50 Å redward of λ5007 and 20 Å to 50 Å blueward of Hβ in the rest frame. We chose these particular wavelength ranges to avoid contamination from all but the broadest Hβ emission-line wings. The spectra with such broad Hβ emission likely came from the broad line region in Type I AGN. Spectra with Hβ emission broad enough to affect the fit were removed from the data set.

As is common with fitting routines, there were cases in which the data was very different from the model applied and the algorithm was not able to find a set best-fit of parameters to converge to. In these cases where the algorithm did not converge for one or more of the emission lines or the

continuum, the spectrum was simply removed from our data set. This accounted for only 1,391 of the 20,372 unique spectra in our original data set, or 6.8%.

2.2.2 Data Vetting

Each spectrum had to be thoroughly vetted for suitability to ensure that only appropriate data were used. Rejection criteria for improper fits were thus built into the fitting code itself. At the first stage of linear continuum fitting, we rejected spectra which had steeply-sloped continua out of concern that these could not be properly fitted around the wings of the emission lines. As will be discussed in the next two subsections, using only a linear continuum fit caused problems. However, even when a new Legendre polynomial continuum fitting routine was introduced, we rejected spectra if the fit returned one of the following:

1. A width of either the $\lambda 4959$ or $\lambda 5007$ line in excess of 30 Å. As previously indicated, this was to avoid broad emission in general since it tended to confuse continuum fitting,

2. An amplitude of the $\lambda 5007$ emission line less than 50×10^{-17} erg cm^{-2} s^{-1} Å^{-1} to avoid lines with low SNR, especially for the lower-amplitude $\lambda 4959$ line,

3. Nonphysical values from misfitting, such as negative line fluxes or flux values of the $\lambda 4959$ emission line which were greater than that of the $\lambda 5007$ emission line, or

4. An amplitude of either emission line that exceeded 1000×10^{-17} erg cm^{-2} s^{-1} Å^{-1} to conservatively guard against detector non-linearity.

Before the continuum fitting method was changed, each spectrum in which the fitting algorithm converged was viewed by a member of the group to assess the suitability of both the fits on the spectrum and the spectrum itself. We believed it to be unwise to give the computer algorithm an exclusive final say, so we reviewed each spectrum individually by eye to ensure that all fits were

reasonable. To do this, four plots per spectrum were generated. The first showed the full spectrum without any model profiles to assess the spectrum's overall suitability. The second showed the 188 Å-range of the spectrum centered around the [O III] lines with the linear continuum fit overlaid. The third and fourth plots showed the $\lambda 5007$ and $\lambda 4959$ emission lines, respectively, overlaid with their Gaussian fits and the local continuum fit.

To preserve objectivity, the calculated ratios, as well as the redshift, were hidden from the reviewer so that the visual analysis was based only on the appropriateness of the continuum and emission-line fits. A spectrum was removed from the data set if it met one or more of the following criteria:

1. The Gaussian fit on either emission line clearly did not match the shape of the data,
2. The data in either range did not contain an obvious emission line,
3. The continuum fit cut off the bottom of either emission line,
4. The continuum fit was below the bottom of either emission line,
5. A broad wing of the Hβ emission reached the $\lambda 4959$ emission line,
6. Both of the emission lines clearly showed double peaks,
7. The emission lines were broad enough to begin to blend, or
8. Any of the fitted lines were clearly unrealistic in any other way when compared to the data.

Figure 2.3 shows a view similar to that of what was shown to the reviewer for a spectrum deemed acceptable. For examples of rejected spectra, and the criterion violated, see Appendix A. This procedure reduced the number of spectra further to approximately 13,000.

During the visual analysis, several features were discovered that, while interesting and noteworthy, were not directly related to the [O III] doublet ratio analysis at hand. Their discussion is therefore left to Appendix B.

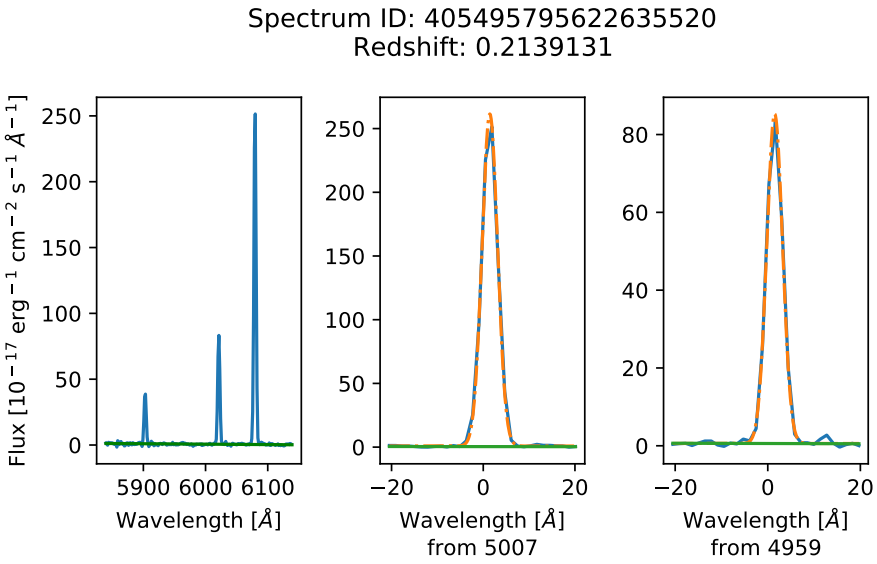

Figure 2.3 Example spectrum that was deemed appropriate during the visual review. The solid blue curves represent the data, the dashed orange curves represent the Gaussian model fits, and the solid green lines represent the linear continuum model fit. These plots are similar to those shown to the person doing the visual review, although these include the spectrum's redshift, which were hidden from the reviewer.

2.2.3 Analysis of Biases

Biases were a dominant concern because any bias that affected spectra differently at different redshift had the potential to invalidate the observed trend in the [O III] doublet flux ratio, whether it was zero or nonzero. Throughout this project, we investigated the effects of high noise levels and continuum inflections, as well as an apparent zero-point bias.

Noise Bias

Early on, we recognized that increasing noise would likely bias the measured [O III] ratio away from the intrinsic value. We generated model spectra to analyze the effect of a decreasing SNR on the [O III] ratio, whose emission lines were fit in the same manner as the data. Reasoning that more distant galaxies would tend to have a lower SNR value, we considered whether or not variations in

spectral SNR levels injected a systematic bias into the fitted line flux values.

Each model spectrum was constructed to have a wavelength range of 4500 Å to 5500 Å with a linear continuum and two Gaussians to represents the λ4959 and λ5007 emission lines. The linear continuum slope was randomly sampled from a normal distribution having a mean of -0.0005 and a standard deviation of 0.005. The y-intercept was randomly sampled from a uniform distribution of value from 5 to 100. These parameters, as well as those to be discussed next, were determined by reviewing the parameters most commonly fitted on the real spectra and trying to match that distribution. The amplitude of the λ4959 line was fixed at 200 while the amplitude of the λ5007 line was set 2.985 higher based on theoretical predictions. The centroid and standard deviation for each emission line pair were randomly sampled from uniform distributions, the centroid distribution having a range of 1 to 1.5 Å and the standard deviation distribution having a range of -4 to -1 Å.[5] The Gaussian emission lines were then added to the existing linear continuum. A copy of this ideal model spectrum was saved, and will hereafter be referred to as "Level 0". In all, 1,000 of these were generated.

To every Level 0 spectrum, eight different levels of noise was added to make a total of 9,000 model spectra. At each point on the model spectrum, a noise value was sampled from a standard normal distribution scaled to a value associated with each noise level and multiplied by 1% of the peak flux value in the model spectrum, and added into the spectrum. The values associated with noise Level 1 - Level 8 were 1, 2, 3, 5, 7, 10, 12, and 15, respectively. Because the height of the simulated emission lines were fixed, the signal and thus the SNR of each noise level was a function of the line width. The SNRs for the stronger λ5007 line averaged out to be approximately 100, 45, 30, 18, 13, 9, 7, and 6, respectively.

Finally, each of the 9,000 model spectra were run through our fitting routines. Of these, 6,942

[5]Because the Gaussain generating function squares the standard deviation, both positive and negative values are acceptable. Negative values of the standard deviation were found to help the code run more smoothly, so they were used over positive ones.

Table 2.1. Generated Spectra Noise Analysis

Noise Level	Scale	SNR	Count	Mean	Median	Std. Dev.	Std. Error
Level 0	0	>100	910	3.043	2.985	0.2298	7.6×10^{-3}
Level 1	1	100	908	3.054	3.025	0.3031	1.0×10^{-2}
Level 2	2	45	889	3.069	3.012	0.4285	1.4×10^{-2}
Level 3	3	30	869	3.167	3.032	0.7704	2.6×10^{-2}
Level 4	5	18	797	3.222	3.004	1.069	3.8×10^{-2}
Level 5	7	13	737	3.317	2.927	1.507	5.6×10^{-2}
Level 6	10	9	600	3.226	2.683	1.777	7.3×10^{-2}
Level 7	12	7	451	2.976	2.361	1.833	8.6×10^{-2}
Level 8	15	6	360	2.838	2.268	1.951	1.0×10^{-1}

could be properly fit and their flux ratios returned, with most fitting failures coming from the higher noise levels. Examples of generated spectra, Level 0 - Level 8, and their fits are shown in Figures 2.4 - 2.6. Table 2.2.3 shows statistics for the returned fits for each noise level. Figure 2.7 shows a histogram of the ratio distribution of all the simulated spectra that made it through the fitting routine, while Figures 2.8 - 2.10 show histograms of the ratio distribution broken down by noise level. From Figure 2.11, the median of the [O III] line ratio is preferred over the mean through Level 5 noise for our analysis, which corresponds to an SNR value of 13.

This trend in the [O III] line ratio comes from a slight asymmetry in the overall error spread that becomes very pronounced at higher noise levels. As the noise level increases, the spread extends further out on the right wing of the curves than on the left. This is because in forming a ratio, the spread on the lower value side of the distribution can only extend towards the value of zero but no

further. However the spread on the opposing side can extend indefinitely, having no numerical limit. As a result there is a natural statistical bias towards higher [O III] ratio values with higher noise levels.

Because the mean values possess a bias towards higher values with greater SNR, we deemed the median, not the mean, values for the estimated [O III] doublet ratio to be more appropriate. We also restricted our data set to include only spectra with a $\lambda 5007$ flux value of 50×10^{-17} erg cm^{-2} s^{-1} Å^{-1} or greater as previously noted. With this constraint, and somewhat to our surprise, we found that the average SNR was not statistically different in the 15 different redshift bins considered (see Table 3.1). Thus, although the SNR affects the calculated ratio, we did not need to apply a correction for it beyond using median values for each bin.

Continuum Color Score

Through careful observation and analysis, we determined that an independent linear continuum fit caused a bias with continuum concavity. The linear fit was unable to account for continuum slope inflections, and either overshot or undershot the bottom of one or both of the [O III] emission lines. More often, a portion of the blue wing of the $\lambda 4959$ Å was subtly sliced off, leading to a flux ratio increase.

We implemented a "continuum color score" to understand how the bias affected the data. This was an index which we used to estimate the continuum slope through the emission lines. This index was the mean of the above blueward continuum values minus the mean of the redward values, divided by their sum. Contrary to typical astronomical conventions, more positive values indicate a bluer continuum. We found a strong correlation with continuum color and ratio, confirming that a linear fit was inappropriate.

2.2.4 Reformulation of Flux Measurement

To obtain a better estimate of the continuum, we adopted a two-stage process. After fitting the emission lines with Gaussian profiles, we subtracted this fit from each spectrum. We then fit the 188 Å range from 70 Å blueward of λ4959 to 70 Å redward of λ5007 in the residual spectrum as measured in the rest frame. The data in this range was modelled with both a general third-order polynomial and a third-order Legendre polynomial, $P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$. The results of either fit were statistically identical, so we used Legendre polynomial fitting exclusively.

The fit to the continuum was then subtracted from the original spectrum and the λ5007 Å and λ4959 Å lines were again fitted with Gaussian profiles. These second fits were integrated over the emission lines and any small residual flux from the continuum offset was subtracted off. The ratio was then obtained by dividing the λ5007 flux counts by the λ4959 flux counts.

Zero-Point Bias

After inroducing Legendre-polynomial fitting procedure, we found a zero-point bias in our fitted results. As detailed in Section 3.1, our calculated [O III] flux ratio remains higher than theoretical predictions or past observational studies, so we assume it to be a bias from our fitting method. While we could not determine its precise cause, we anticipate that spectra at different redshifts are affected equally. Thus, our results in Section 3.2 concerning the variation of the flux ratio should remain valid.

2.3 Data Fitting Summary

To summarize the data fitting process, each spectrum was initially assumed to be a simple superposition of a a linear underlying continuum and Gaussian Hβ and [O III] λ4959 and λ5007 emission lines. Each spectrum was fitted as such, using Python code written specifically for the

task. Visual analysis of each fitted spectrum was used to reject inappropriately fitted spectra that were not removed by the rejection criteria built into the code.

We performed analysis of the effect of increasing noise on our calculated [O III] flux ratio to find that it was unstable with an [O III] λ5007 SNR of less than about 13. When we looked into the effect of continuum concavity on the calculated flux ratio, we saw that the ratio was correlated with our continuum color score. We thus determined that our original scheme of fitting a superimposed linear continuum onto the spectrum was inappropriate for our data.

We applied a new fitting method to the data where the continuum and the emission lines were fit independently. First, the emission lines were fit with Gaussian profiles, which were then subtracted from the rest of the spectrum. We fit what was left over with a third-order Legendre polynomial. The Legendre polynomial fit was then subtracted from the original data, and the emission lines were fit again to find their final parameters, which we then used to calculate the flux ratio. A zero-point bias exists, but we believe that it does not change our conclusions.

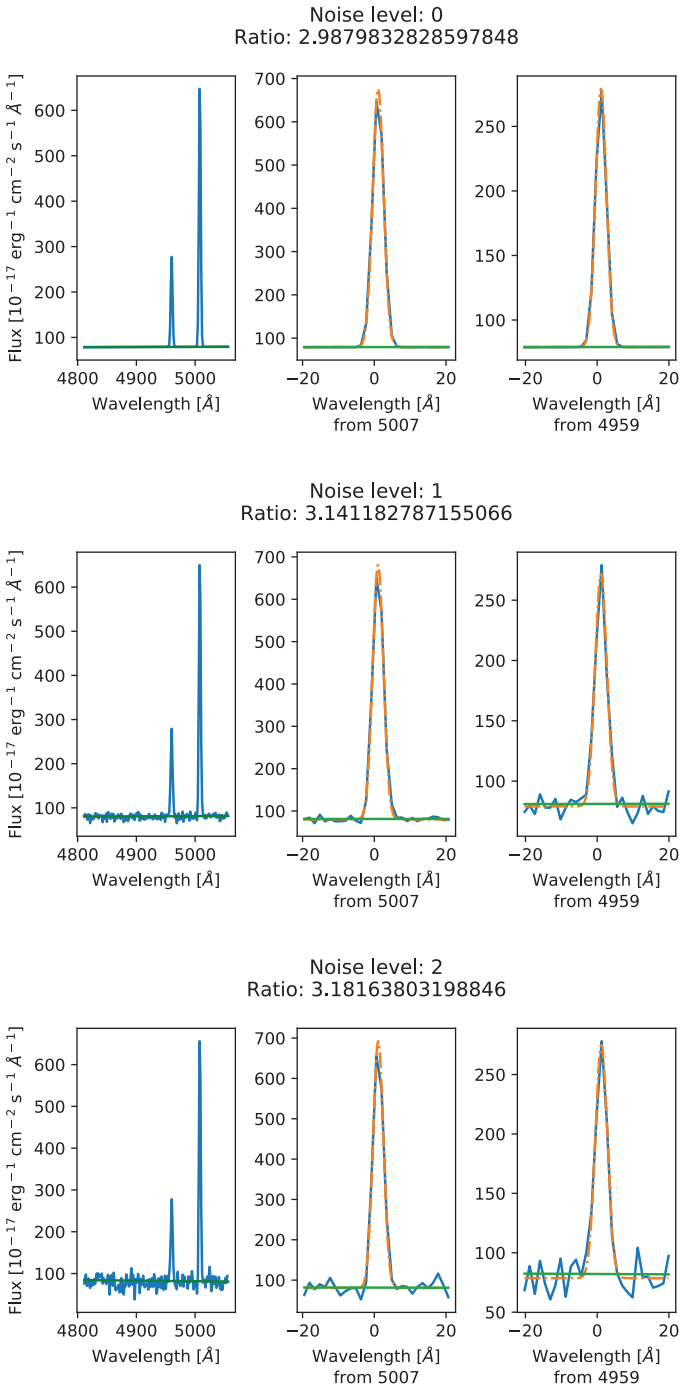

Figure 2.4 Plots of a generated spectrum and its emission line Gaussian fits for Level 0 - Level 2. The blue curve represents the data, the orange curves represent the model Gaussian fits, and the green lines represent the model linear continuum. The flux ratio is extremely sensitive to very small amounts of noise and slight variations in the location of the fitted continuum.

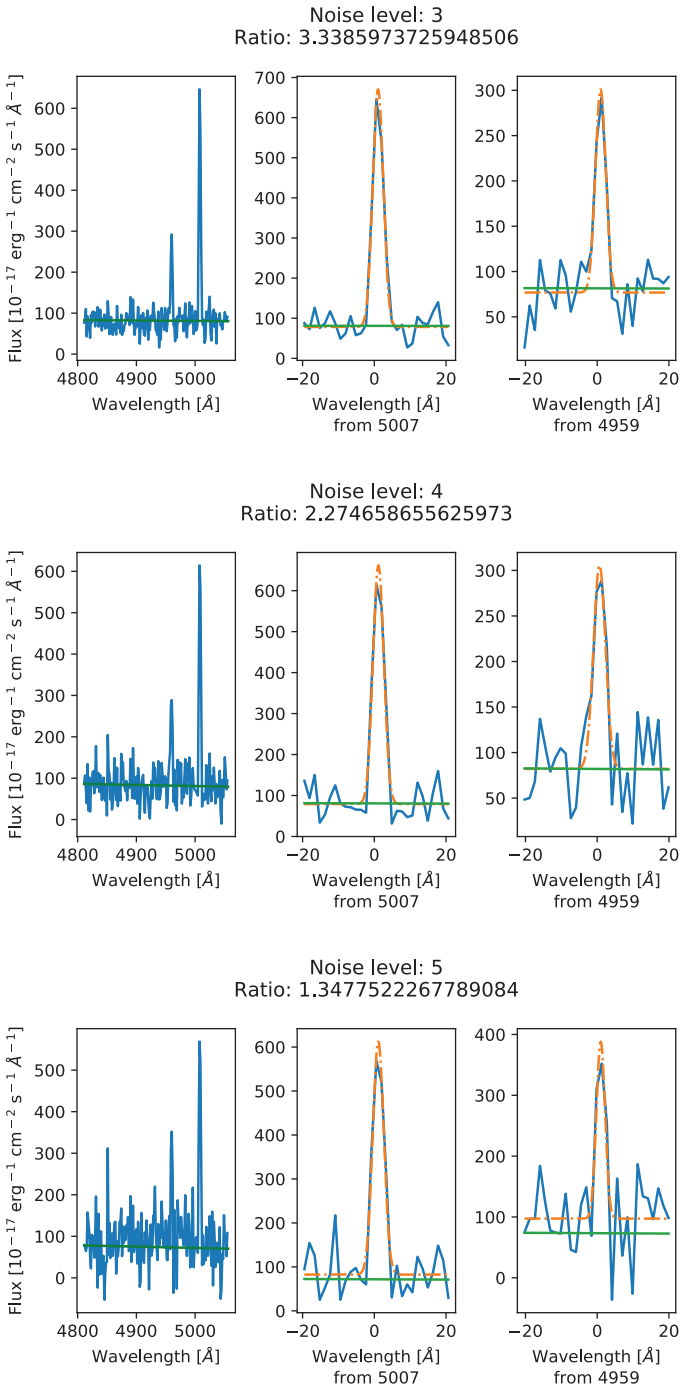

Figure 2.5 As Figure 2.4, for Level 3 - Level 5 noise. The flux ratio becomes increasingly unstable as noise increases and the the emission lines and continuum become harder to fit correctly.

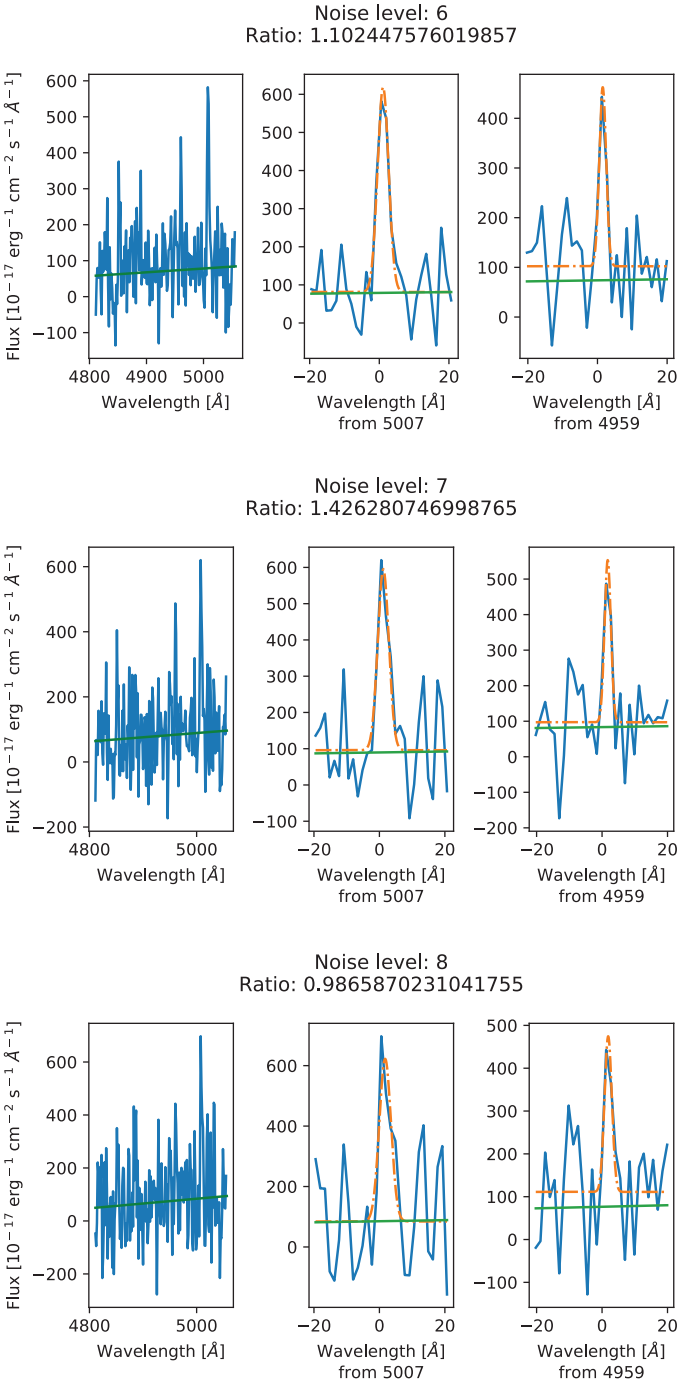

Figure 2.6 As Figures 2.4 and 2.5, for Level 6 - Level 8 noise. The flux ratio is nonphysical at this point as the fitting routines cannot distinguish between the emission lines and the noise surrounding them.

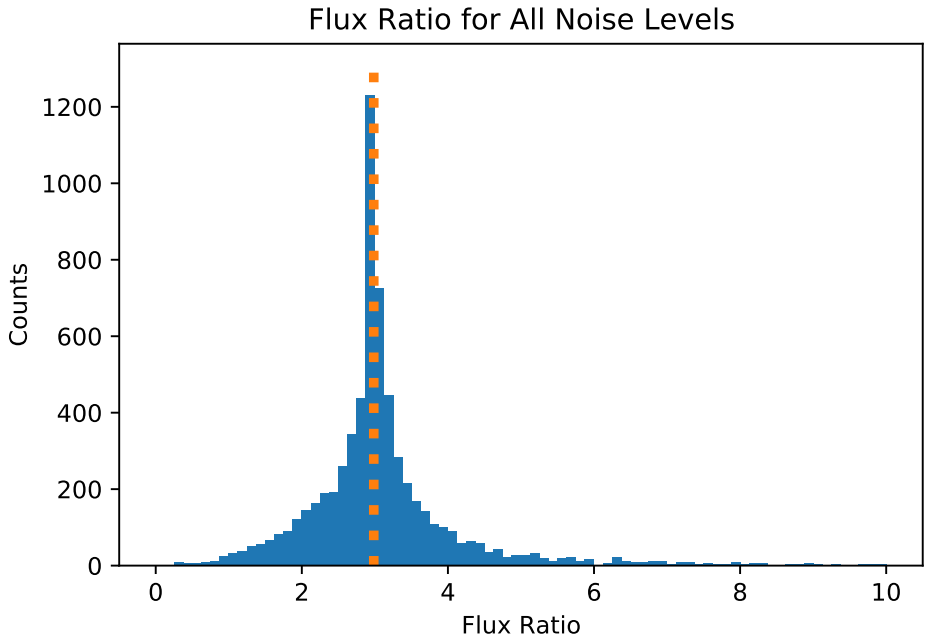

Figure 2.7 Histogram showing the calculated flux ratio for all simulated spectra which made it through the fitting process, regardless of noise level. The dotted orange line shows the position of the input ratio value of 2.985. The asymmetry and long rightward tail are worth noting.

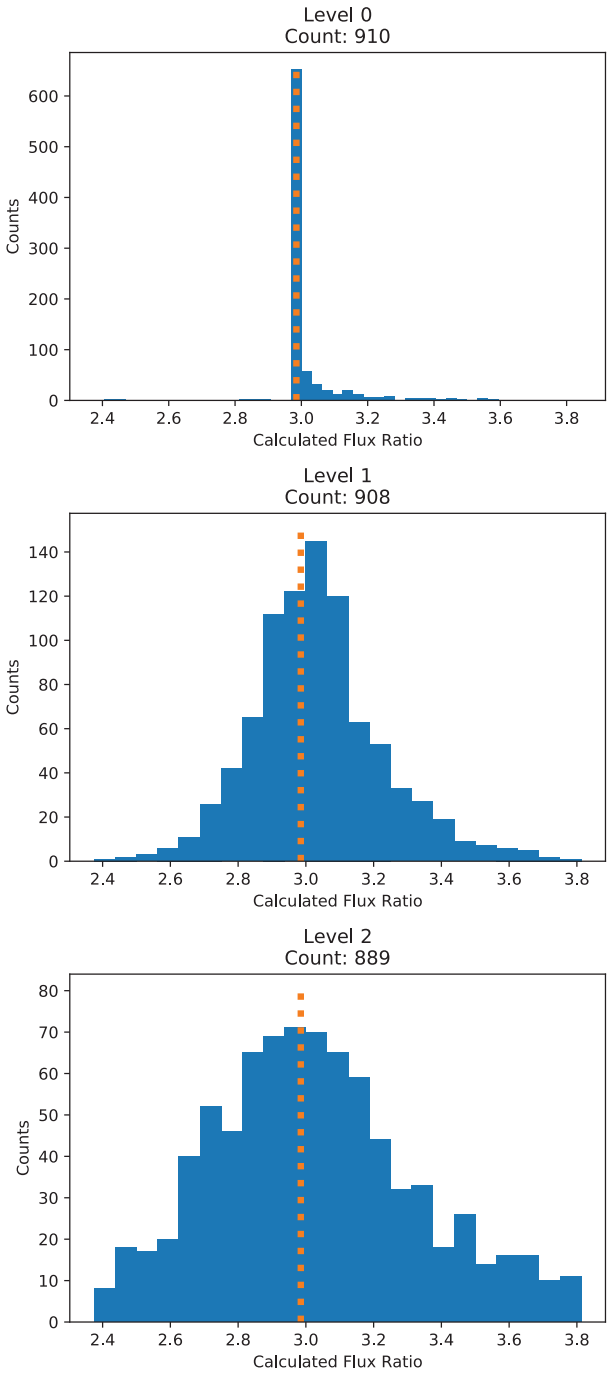

Figure 2.8 Histograms showing the calculated flux ratio for simulated spectra broken into noise levels, for Level 0 - Level 2. The Level 0 diagram has slight right tail, stemming from difficulty fitting the continuum. The Level 1 flux ratios are mostly normal distributed, while the Level 2 flux ratios are beginning to show rightward skew.

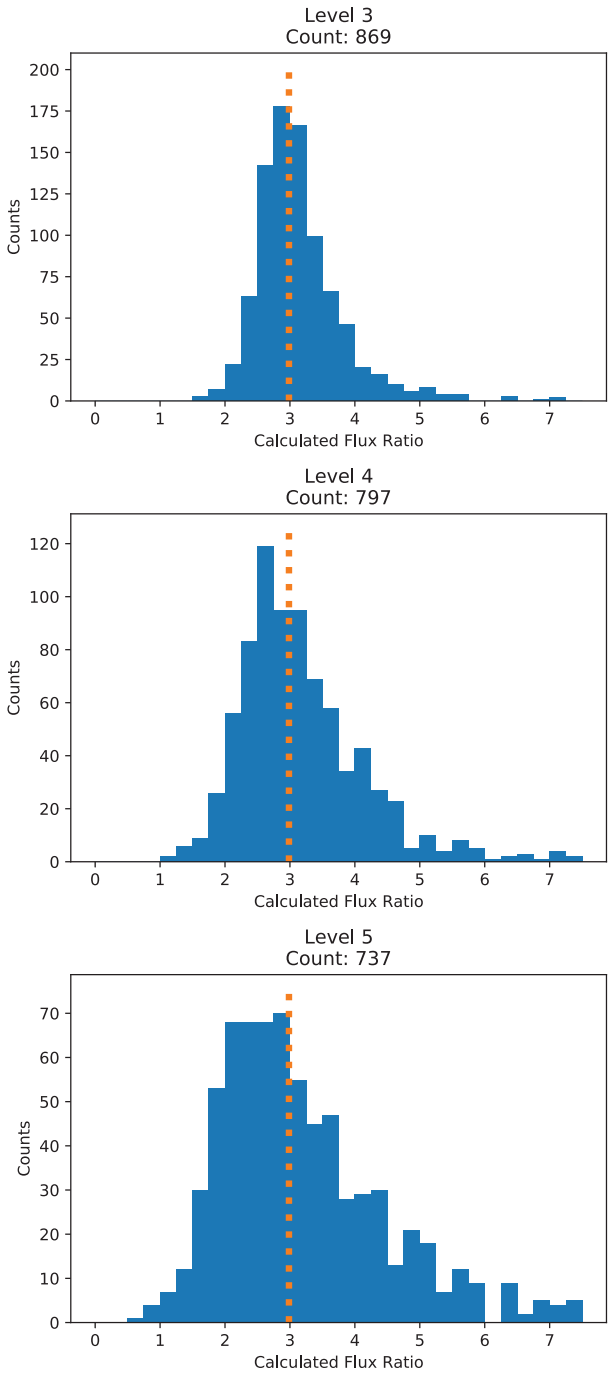

Figure 2.9 As Figure 2.8, for Level 3 - Level 5. With increasing noise, the flux ratio distribution becomes more skewed to the right and the peak ratio value decreases.

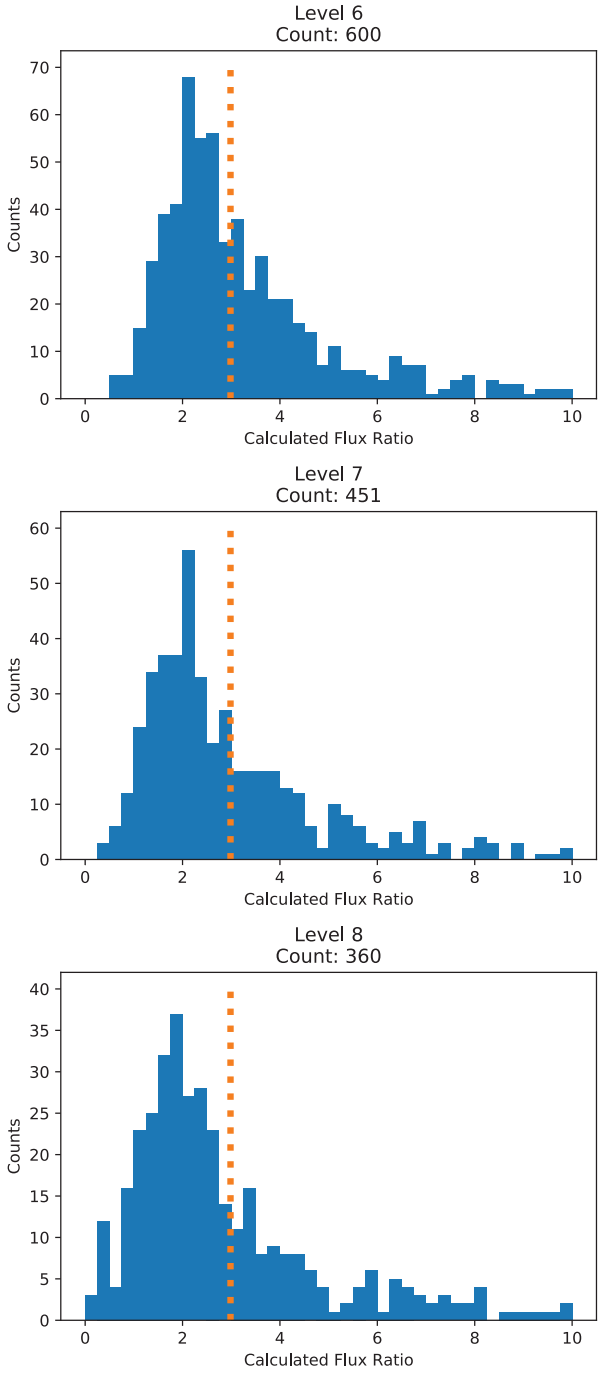

Figure 2.10 As Figures 2.8 and 2.9, for Level 6 - Level 8. The peak values for these noise levels have fallen well below the 2.985 set value and show significant right skew. Spectra with this much noise are definitely unsuitable for use in our final analysis.

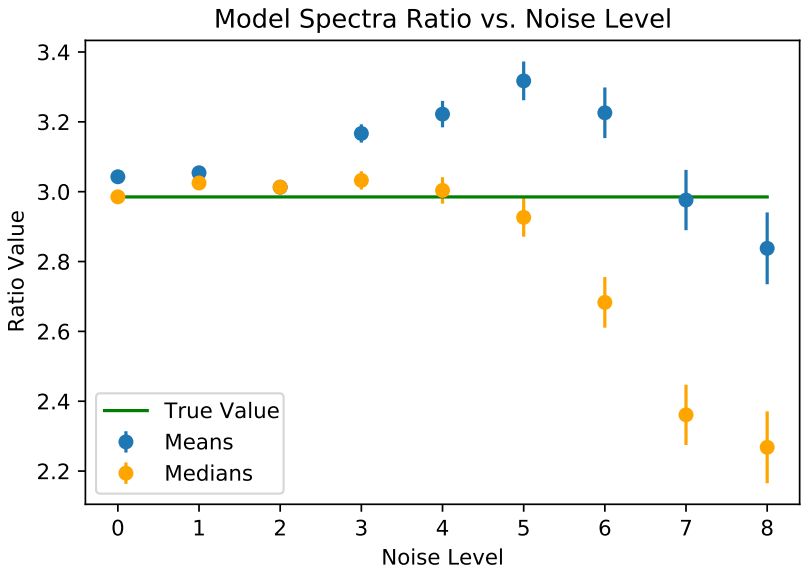

Figure 2.11 Plot of the calculated flux ratio of against the noise level for the generated spectra. The blue points represent the mean flux ratios in the noise level while the orange points represent the median flux ratios. The error bars in both cases are given by the standard error in the mean. The green line corresponds to the input ratio value, set at 2.985. Neither fit statistic matches the input ratio well for noise levels 6, 7, or 8. Noise levels 0, 1, 2, 3, 4, and 5 are all better approximated using the median than the mean.

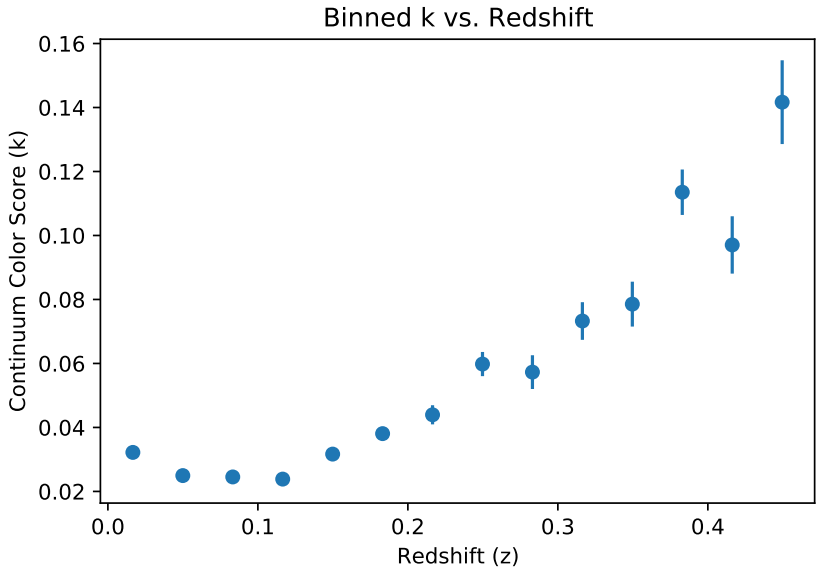

Figure 2.12 The continuum color score, binned into redshift bins of size $\Delta z = 0.033$, plotted against redshift. The color score is obviously correlated with redshift; a higher, bluer k value is linked to an increasing redshift.

Chapter 3

Results and Conclusions

Following the methods described in the previous chapter, the [O III] $\lambda\lambda$5007, 4959 flux ratios of 12,241 spectra were determined. Section 3.1 will detail the results and Section 3.2 will provide a discussion on the meaning of these results.

3.1 The [O III] $\lambda\lambda$5007, 4959 Å Flux Ratio

After proper statistics were applied to the 12,241 calculated [O III] $\lambda\lambda$5007, 4959 flux ratios, the results showed no significant trend of the [O III] ratio with increasing redshift, as expected for a cosmology in which the physical constants have not changed.

The distribution of calculated ratios was very near normal, indicating an adequate treatment and elimination of biases. Table 3.1 and Figure 3.2 present the median-measured [O III] flux ratio value in redshift bins of width 0.033 in z. The overall variance for each bin was calculated from the standard deviation of each bin divided by the square root of the number of objects in the bin. We truncated the data at $z = 0.467$ as there were only 10 objects beyond this. The distribution of calculated ratios is very nearly Gaussian, with only a slight right skew (Figure 3.1). For this reason we took the median ratio value as the better representation of the true average.

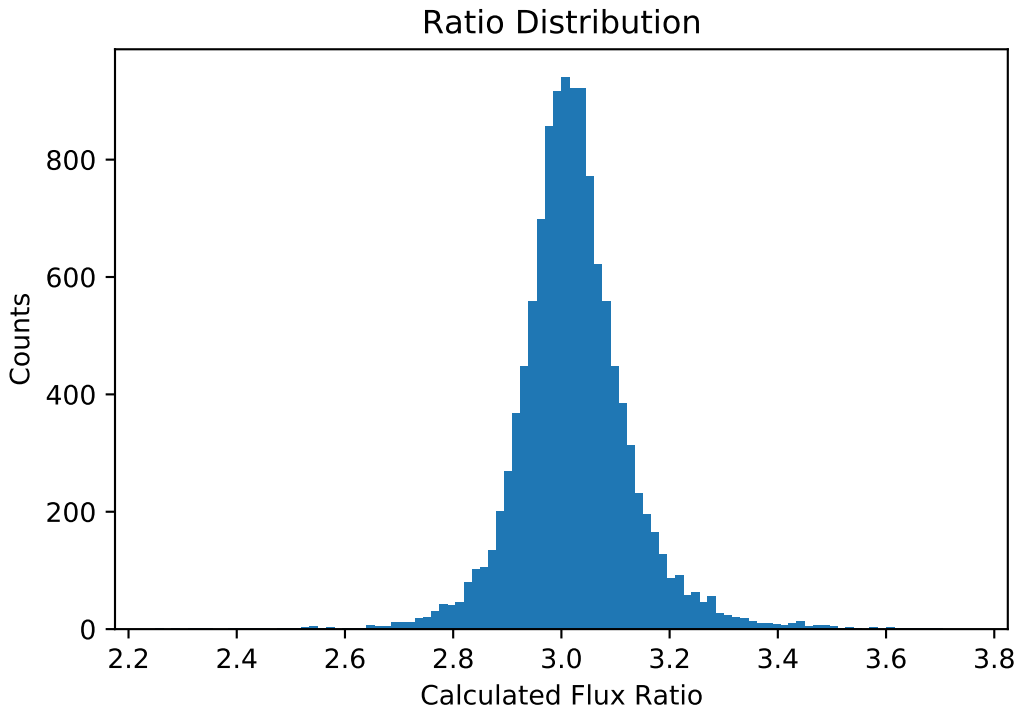

Figure 3.1 Histogram showing the calculated ratio distribution of all spectra, after all cuts.

As can be seen in Table 3.1, the properties of the [O III] emission doublet show little change between redshift bins. Column 1 gives the redshift range. Rows are broken into redshift bins, with the final row representing all objects in the survey, excluding the ten data points truncated due to their redshift being higher than $z = 0.467$. Column 2 lists the number of objects in each bin. Columns 3 and 4 give the mean and median ratio within the bin respectively. Column 5 lists the standard deviation of the mean. Column 6 gives the average SNR of the $\lambda 5007$ line within each bin.

As shown by the plot in Figure 3.2, the value of the flux ratio stays approximately constant with increasing redshift. The larger uncertainty with increasing redshift reflects smaller numbers at greater distances. Applying a weighted least squares regression to the binned data and using the inverse of the errors in the bin means as weights, we find no trend in [O III] flux ratio with redshift. A 95% confidence interval gives the slope between (-0.049, 0.049) and the y-intercept between

Table 3.1. Emission line properties with distance

Redshift range	Count	Mean Ratio	Median Ratio	Std Dev	Mean SNR
$0.000 < z < 0.033$	3329	3.007	3.003	0.152	31.4
$0.033 < z < 0.067$	2775	3.026	3.022	0.099	31.3
$0.067 < z < 0.100$	2104	3.025	3.018	0.117	30.6
$0.100 < z < 0.133$	1383	3.031	3.025	0.158	30.5
$0.133 < z < 0.167$	998	3.045	3.033	0.125	29.5
$0.167 < z < 0.200$	657	3.031	3.030	0.154	30.4
$0.200 < z < 0.233$	395	3.019	3.025	0.124	29.2
$0.233 < z < 0.267$	211	3.049	3.044	0.120	30.4
$0.267 < z < 0.300$	146	3.010	3.005	0.154	28.9
$0.300 < z < 0.333$	98	3.015	3.024	0.120	27.6
$0.333 < z < 0.367$	56	3.023	3.021	0.084	28.1
$0.367 < z < 0.400$	27	3.030	3.014	0.084	31.8
$0.400 < z < 0.433$	39	3.020	3.012	0.134	29.8
$0.433 < z < 0.467$	23	3.028	3.000	0.088	31.7
$0.000 < z < 0.467$	12241	3.023	3.017	0.133	30.8

(3.012, 3.040). This interval encompasses both the mean and median ratio values.

Results were very different when the continuum was handled less carefully. Those analyses showed a strong upward trend in the ratio as redshift increased, with a ratio of only 2.91 at $z = 0$. This trend was found to be conflated with the continuum color score, and was thus rejected in favor of the methods described above. Table 3.1 shows a comparison of selected statistics between the

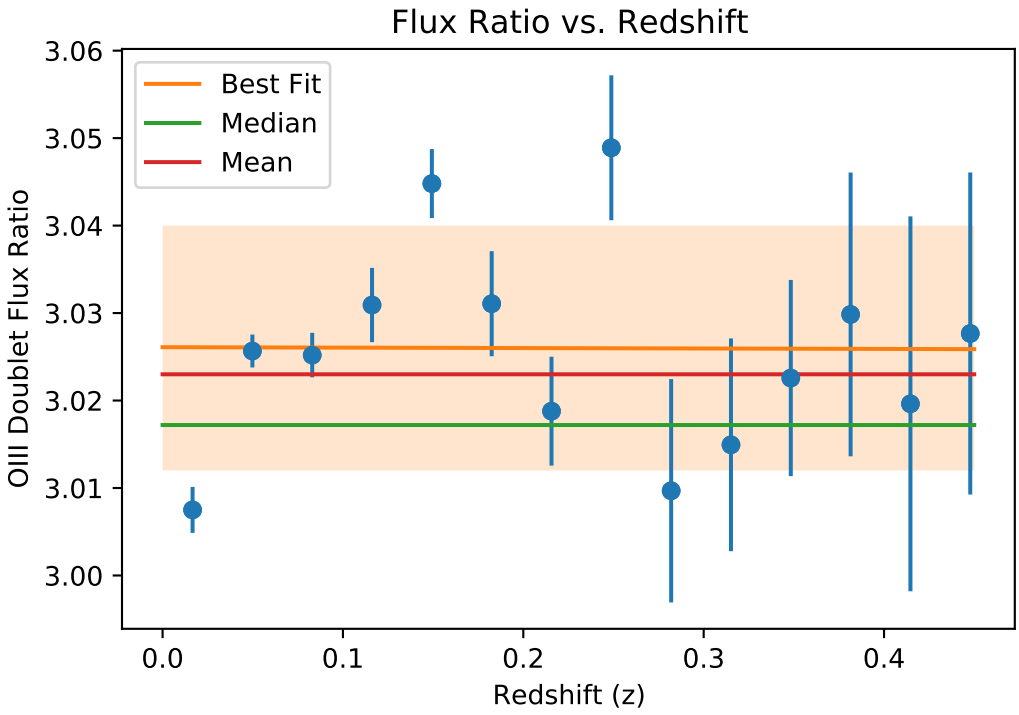

Figure 3.2 Data showing the $\lambda\lambda$5007, 4959 ratio increase with redshift. The errors bars are given by the errors in the binned means, shown in Table 1. The orange line shows the trend line fitted to the binned data, weighted by the bin's standard deviation. The orange box shows a 95% confidence interval for the for the y-intercept.

linear and Legendre polynomial continuum methods.

3.2 Discussion

Our results show that the proper treatment of the continuum is critical to the final results. When a simple linear fit was used for the continuum, the results were biased by inflections to almost nonphysical results. When we fit with the third-order Legendre polynomial, not only did the fit of each individual spectrum improve, the standard error of the data set as a whole decreased and the mean and median values of the set were closer to the expected values. The median is still slightly higher than expected, and we believe this to be a zero-point bias that likely comes from the

Table 3.2. Linear vs. Legendre polynomial continuum results

Statistic	Linear Continuum	Legendre Continuum
Count	12644	12241
Mean	2.928	3.023
Median	2.938	3.017
Std. Dev.	0.203	0.133
Std. Error	1.801×10^{-3}	1.198×10^{-3}
Intercept	2.924	3.026
Slope	0.186	-0.001
p-Value	4.11×10^{-40}	0.984
Significant	yes	no

continuum fit. Our difficulties precisely fitting the continuum show a need for higher scrutiny for large-scale analyses of star-forming galaxies' spectra to prevent biases.

Because there was no non-zero trend between galaxy's [O III] $\lambda\lambda$5007, 4959 flux ratio and its redshift, we found no evidence for a change in the emission ratio out to a redshift of $z = 0.467$. We thus conclude that there is no change in the Einstein spontaneous emission coefficients of the [O III] λ5007 and λ4959 Å emission lines to the precision of our study.

Appendix A

Rejected Spectra

Figures A.1 - A.24 show spectra that met the criteria for rejection during the visual fitting stage. In each figure, the left panel shows the continuum in the region on which it was fit with a linear model. The center and right panels show the $\lambda 5007$ Å and $\lambda 4959$ Å emission lines, respectively. In each panel, the blue line represents the SDSS-calibrated data, the green line represents the linear continuum fit, and the orange dashed curves represent the emission line Gaussian fit. The figure captions indicate which criterion or criteria indicated removal of that spectrum.

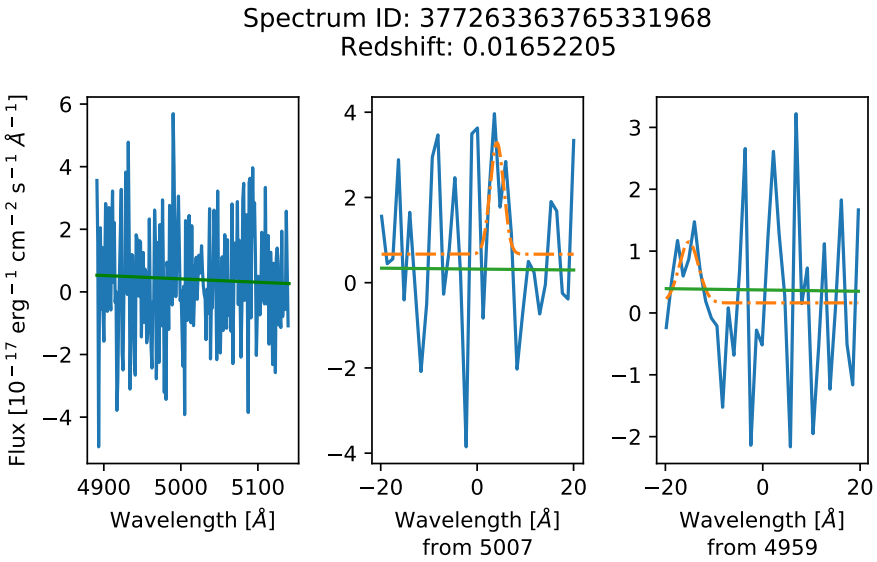

Figure A.1 Rejected by criterion 2.

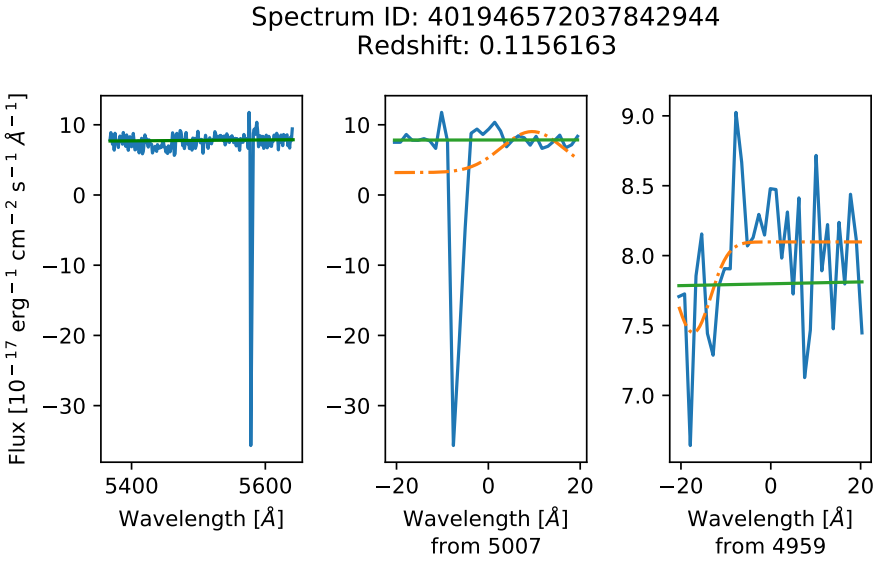

Figure A.2 Rejected by criteria 1, 2, and 8.

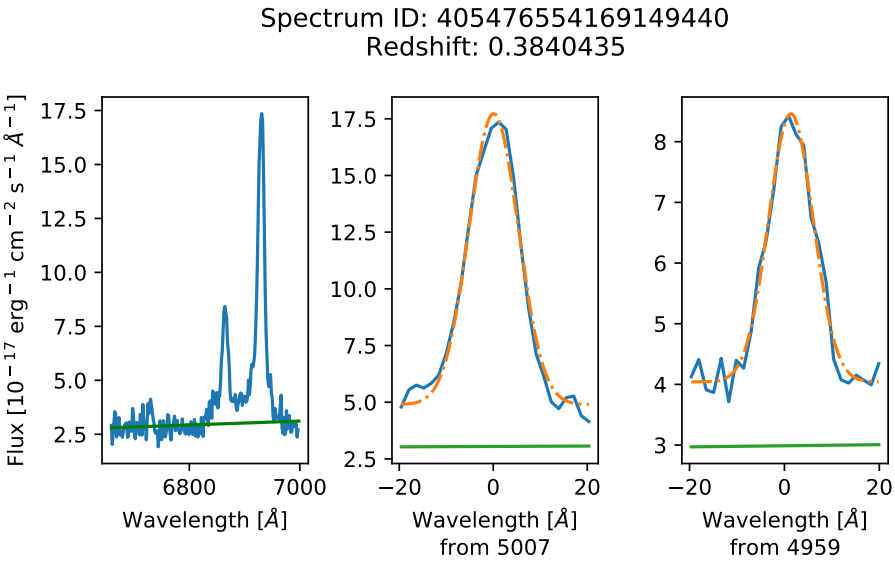

Figure A.3 Rejected by criteria 4 and 7.

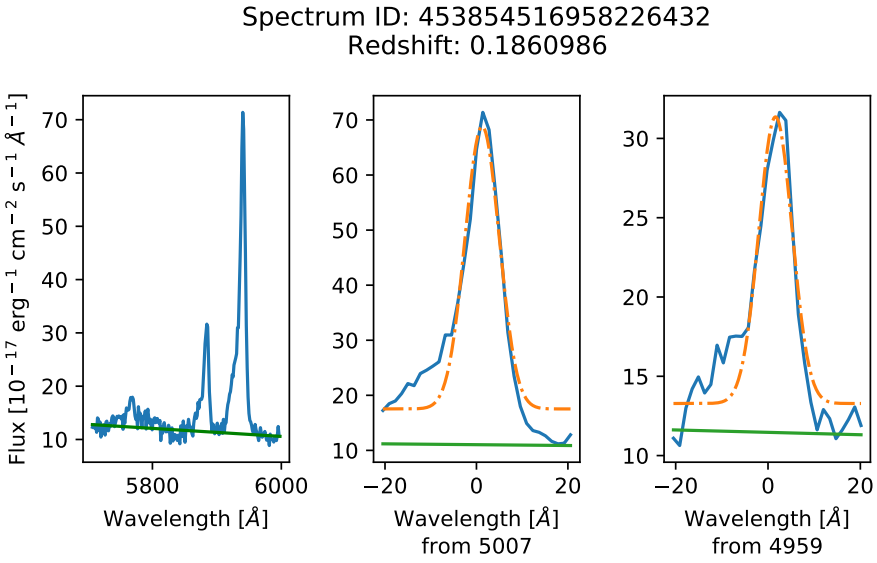

Figure A.4 Rejected by criteria 1, 4, and 8.

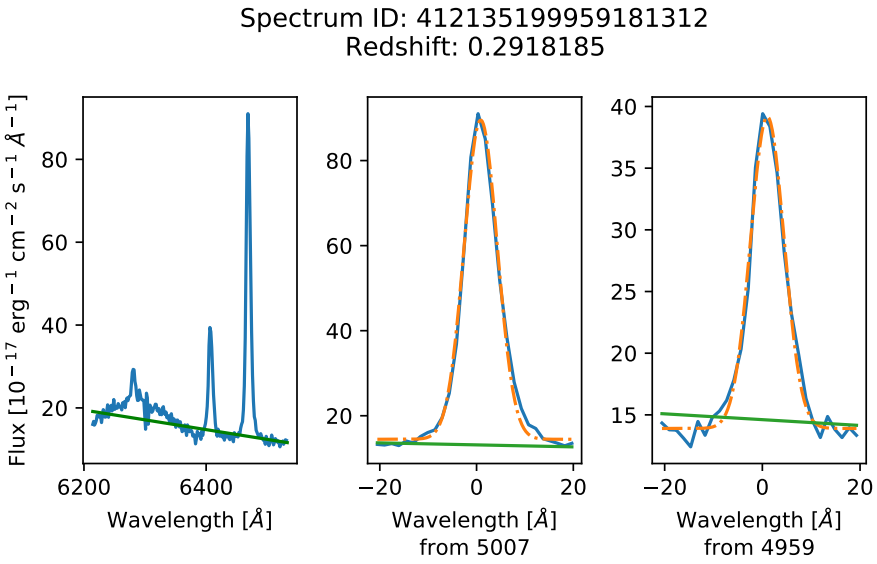

Figure A.5 Rejected by criteria 3 and 5.

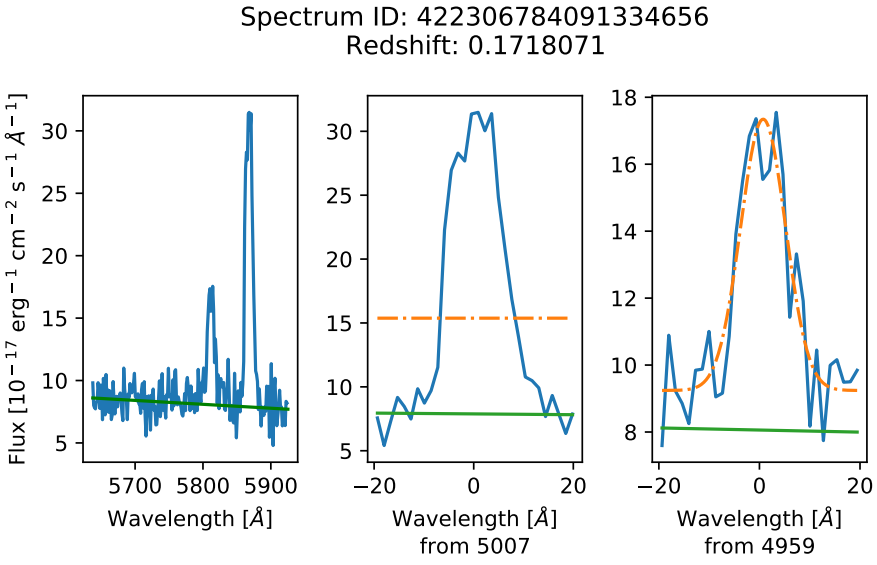

Figure A.6 Rejected by criterion 1.

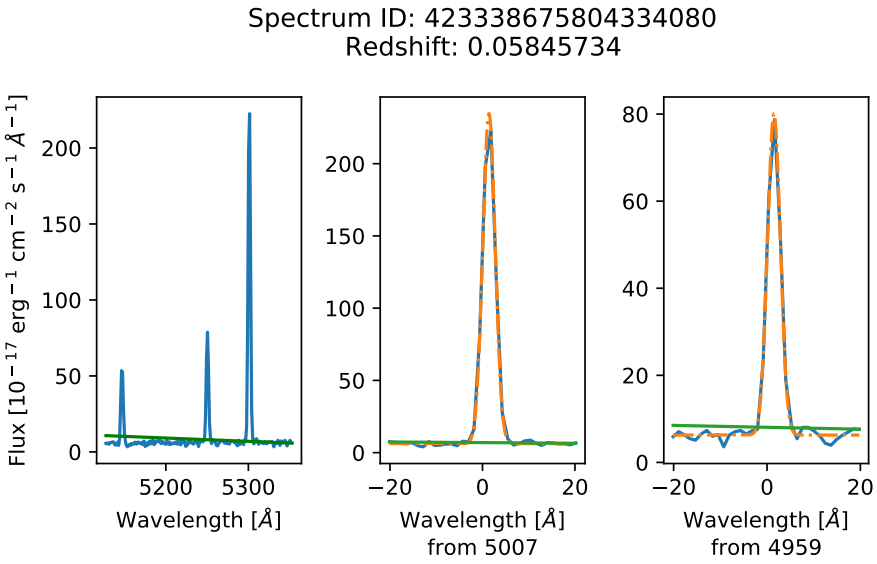

Figure A.7 Rejected by criterion 3.

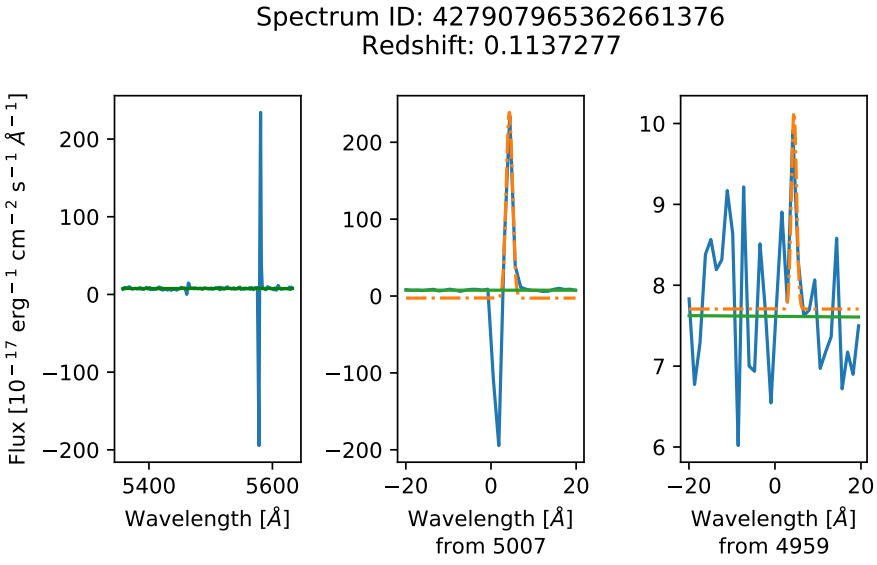

Figure A.8 Rejected by criterion 2.

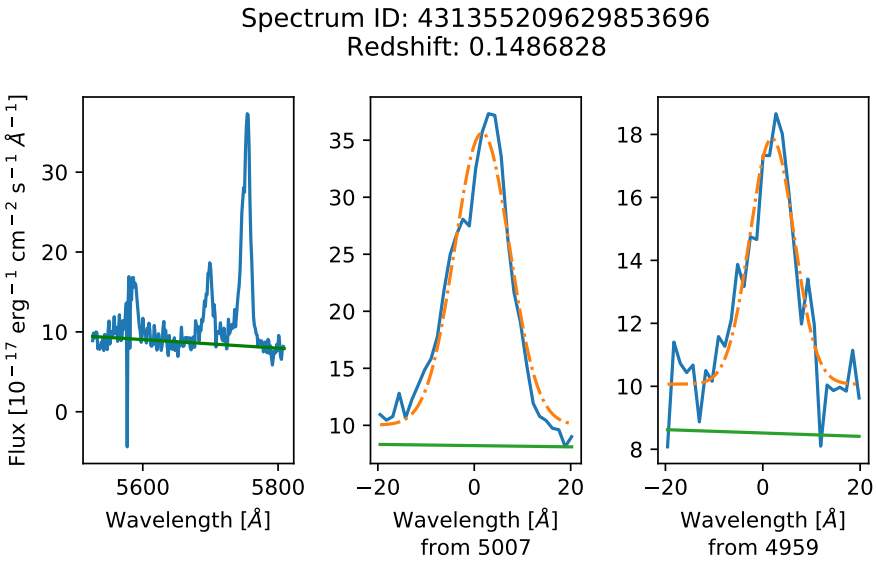

Figure A.9 Rejected by criteria 4 and 7.

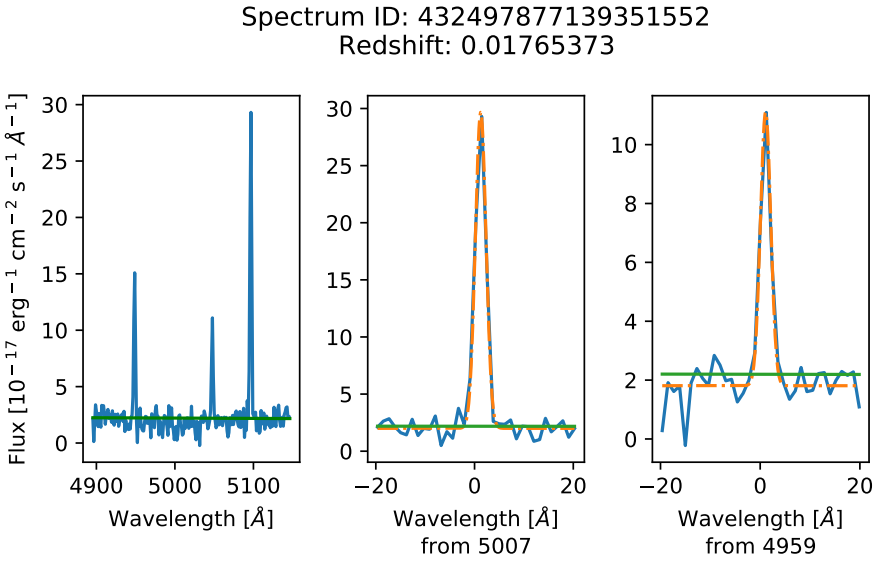

Figure A.10 Rejected by criterion 3.

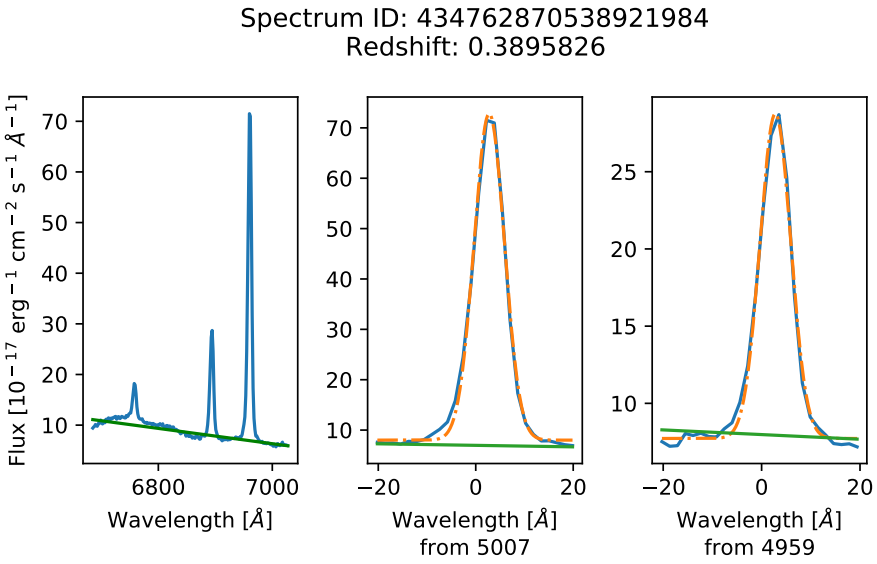

Figure A.11 Rejected by criteria 3 and 5.

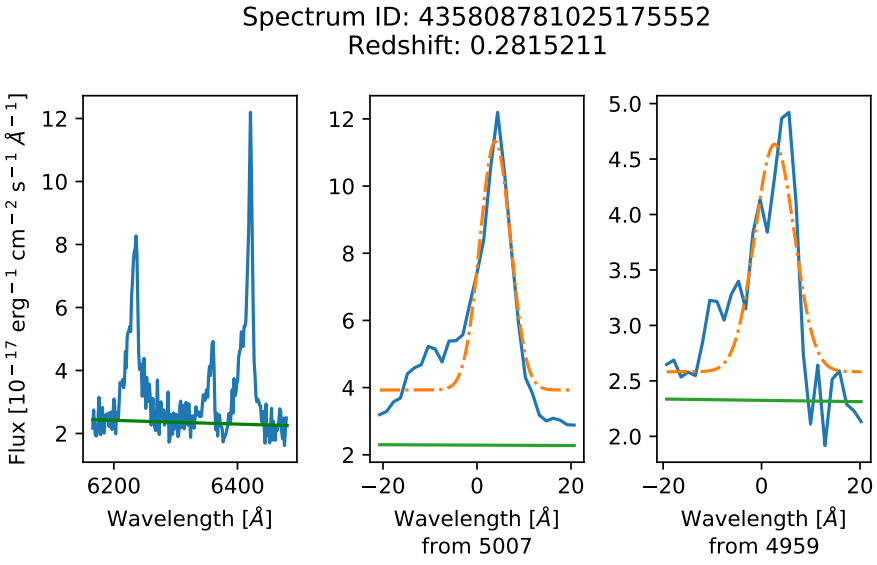

Figure A.12 Rejected by criteria 4 and 8.

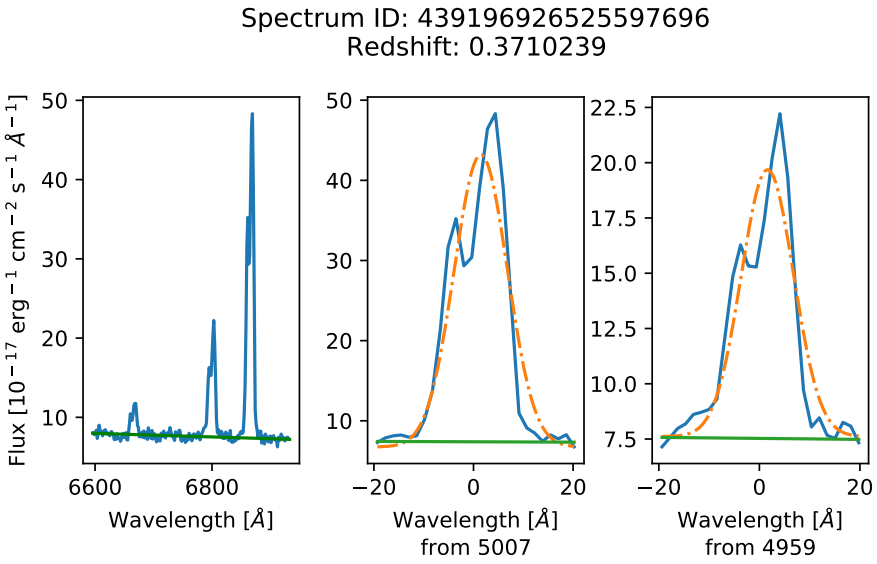

Figure A.13 Rejected by criterion 6.

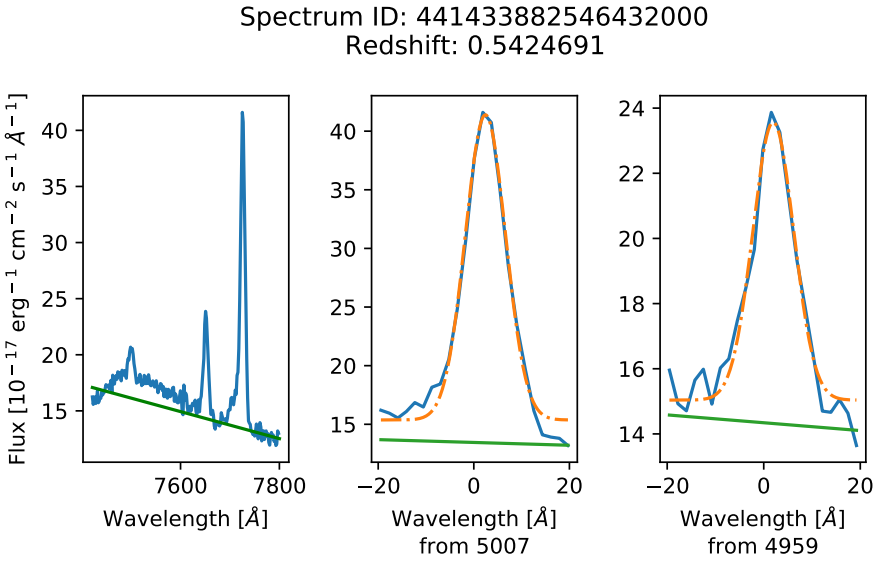

Figure A.14 Rejected by criteria 4 and 5.

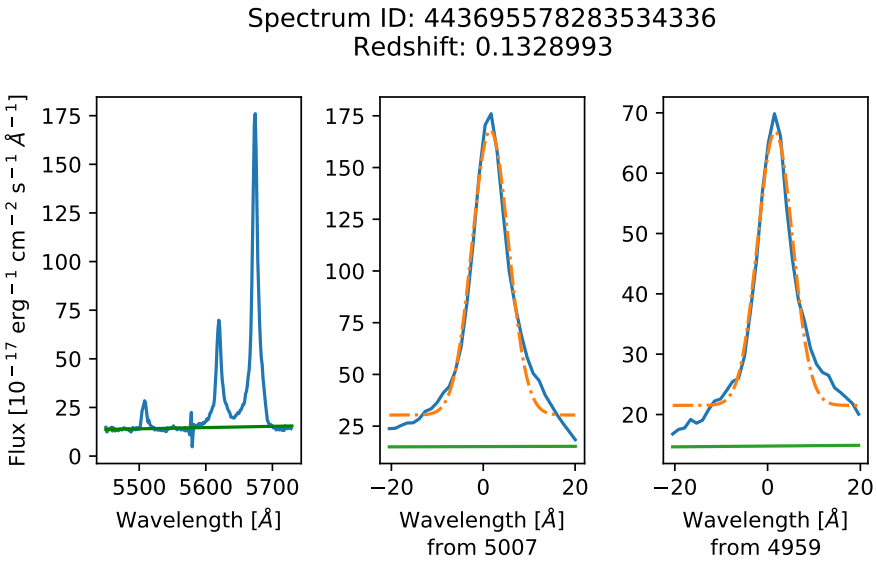

Figure A.15 Rejected by criteria 4 and 7.

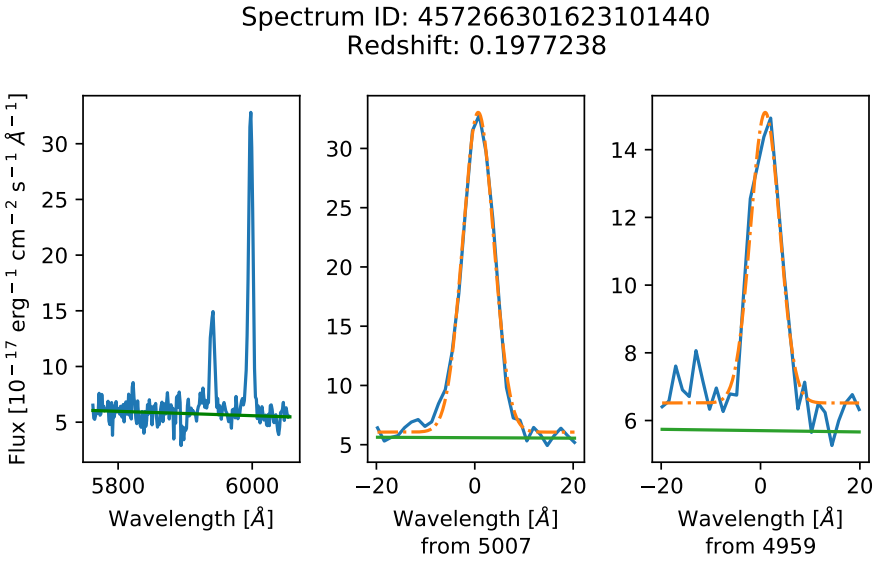

Figure A.16 Rejected by criterion 4.

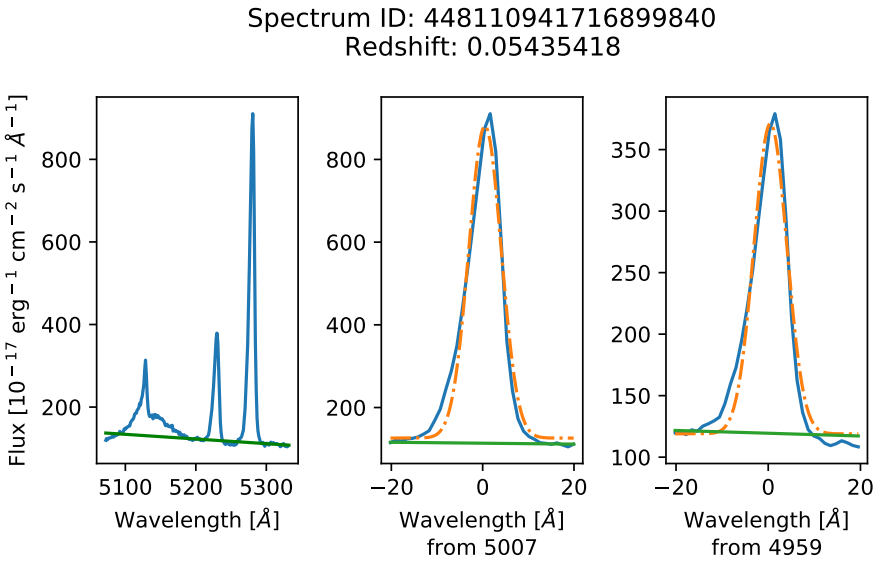

Figure A.17 Rejected by criteria 3 and 5.

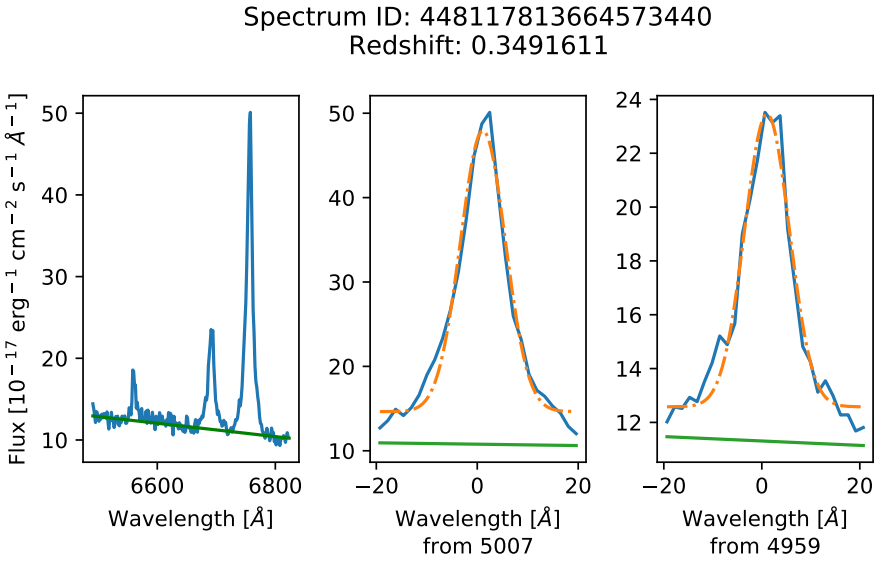

Figure A.18 Rejected by criteria 3 and 5.

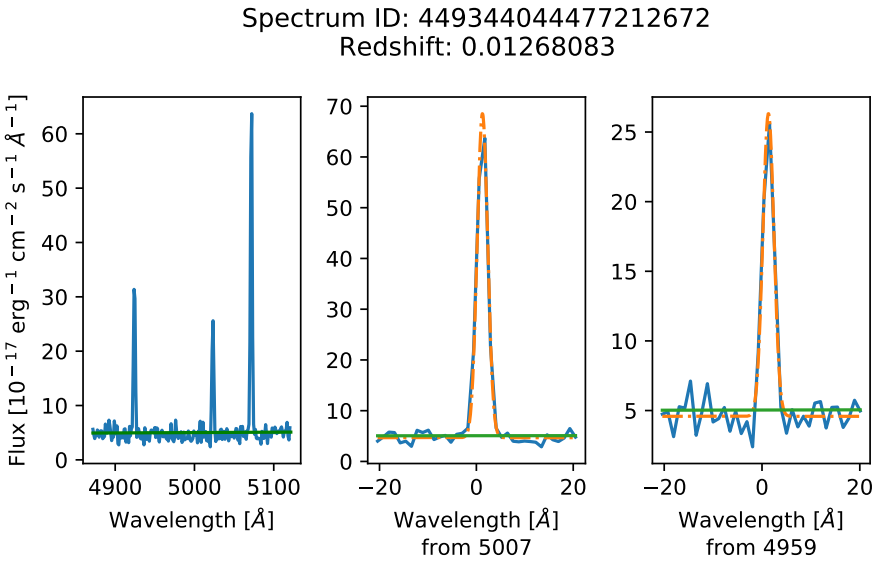

Figure A.19 Rejected by criterion 3.

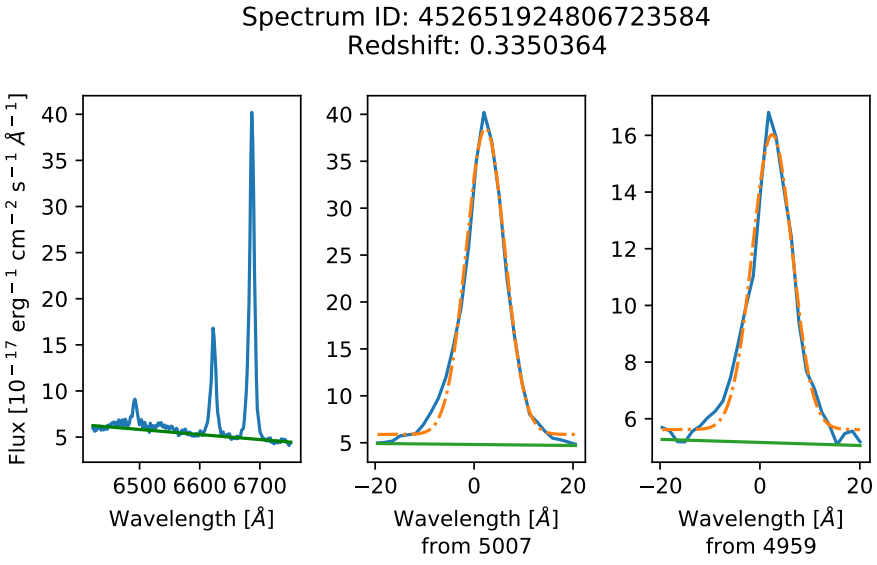

Figure A.20 Rejected by criterion 5.

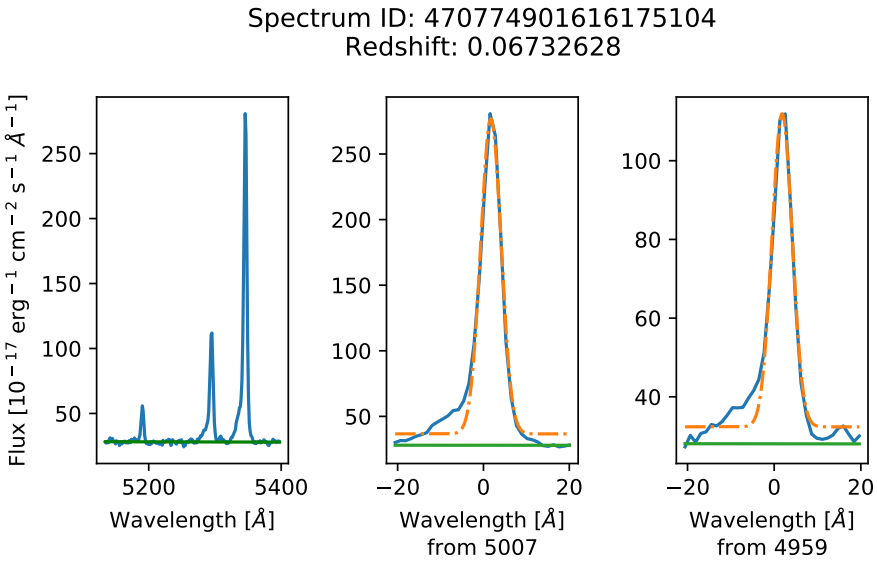

Figure A.21 Rejected by criterion 8.

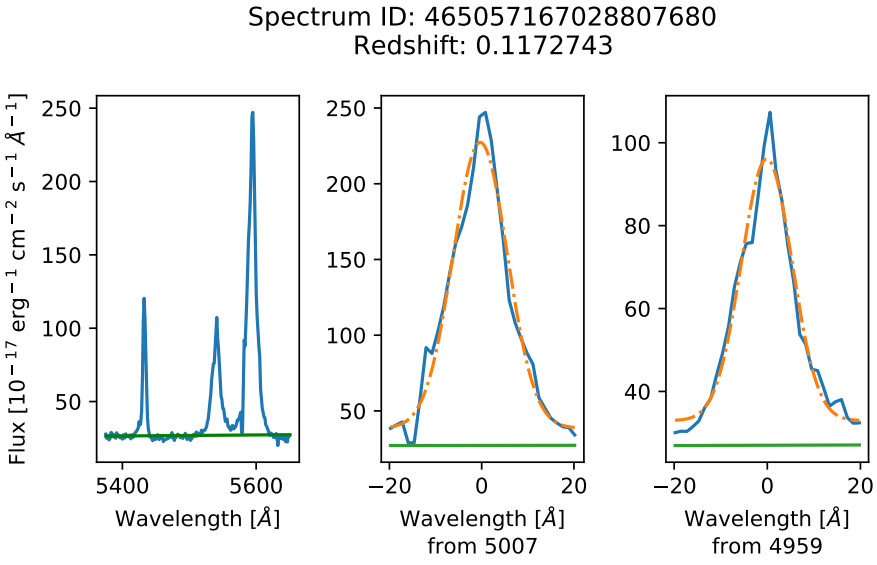

Figure A.22 Rejected by criterion 7.

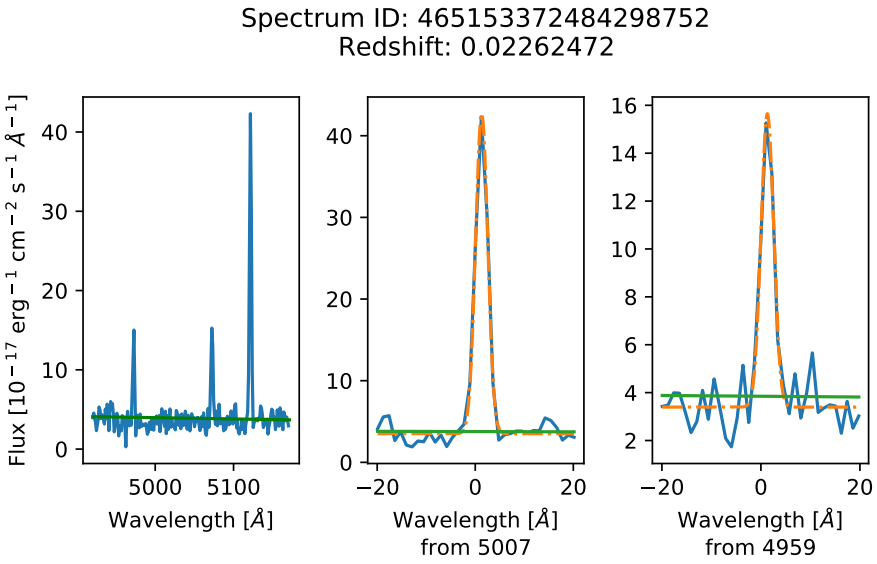

Figure A.23 Rejected by criterion 3.

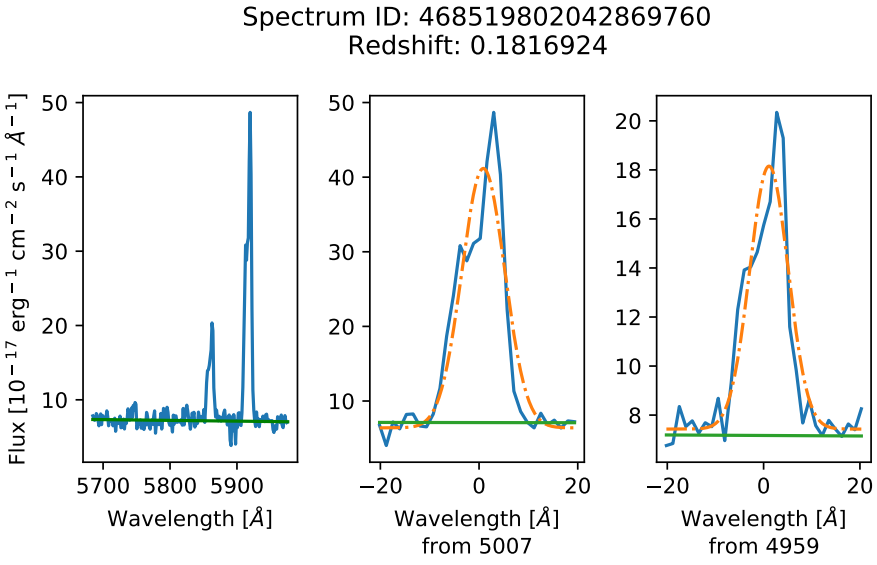

Figure A.24 Rejected by criteria 3, 6, and 8.

Appendix B

Noteworthy Spectral Features

During the visual review process (see Section 2.2.2), we found a small percentage of spectra that, while inappropriate for our study, were interesting for other reasons and deserve mention. In all plots in this appendix, the blue line represents the calibrated data and the orange markers represent the redshifted centroid wavelength of the Hβ and [O III] $\lambda\lambda$4959,5007 Å emission lines.

The first feature of note was what appeared to be double peaks. In some spectra, there were two clear flux maxima in each of the [O III] emission lines. We were intrigued, and initially wondered if this was an artefact of the data collection or the MPA-JHU pipeline. However, the not insignificant number of cases (on the order of tens in the dataset), as well as the matching profiles of the 5007 and 4959 lines in the overwhelming majority of this subset led us to believe that something more was going on. With further research into the field, we found that these are likely the result of galaxy merger events. They could also be due to rotational broadening or bi-conical outflows fro the central AGN.

Another noteworthy feature was a broadened set of [O III] emission lines. Typically, O III emission arises from the narrow line region of galaxies and have very small widths. An example of this can be seen in Figure B.3. In many broad-line cases, we also observed moderate to strong asymmetries, such as in Figures B.4 and B.5. These features probably arise from motion of the gas

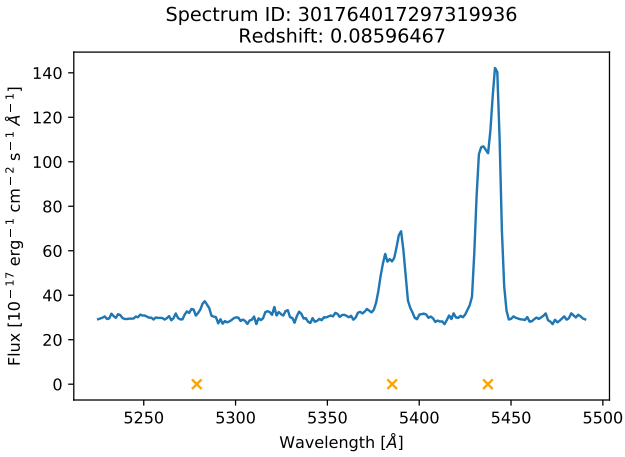

Figure B.1 An example of a spectrum showing double peaks on both of the [O III] emission lines.

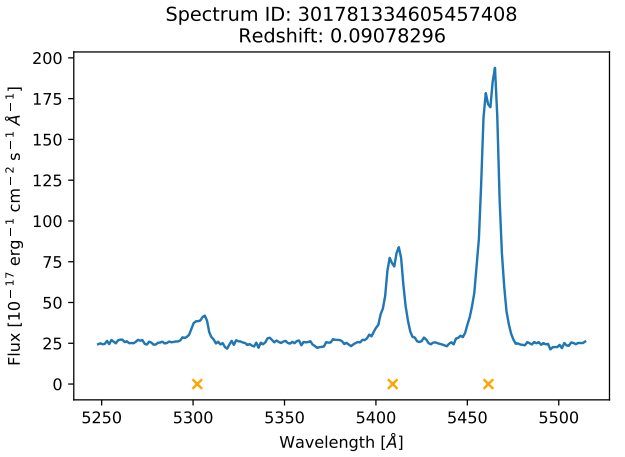

Figure B.2 Another example of a spectrum with double peaks in the [O III] emission lines.

regions. Future studies could be done to determine if this is the case and whether these lines can be used as tracers of mass motion.

Figure B.6 shows the final feature of note: a missing spectral line. In a few dozen spectra, either the λ4959 Å or the λ5007 Å line was simply missing, which should not be physically possible based on the emission processes which create these lines. Upon closer inspection, we found that the region which should contain the line was masked out and the flux values across several Å were

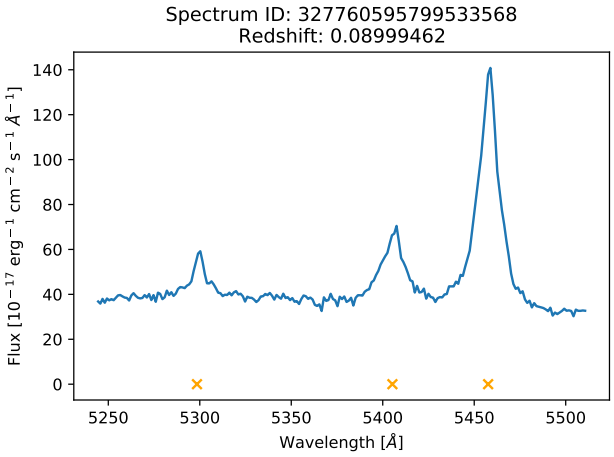

Figure B.3 Example spectrum with broadened [O III] emission lines, especially when compared the the Hβ line.

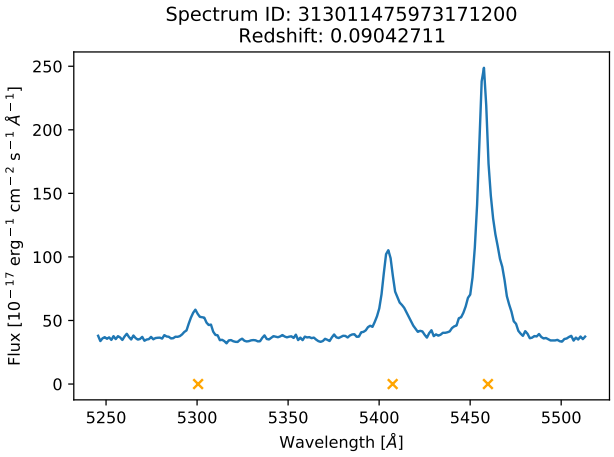

Figure B.4 Spectrum with somewhat broadened [O III] emission lines, as well as strong red-wing asymmetries. The Hβ line also shows a large red wing.

identical. We believe this to be an artifact of either the SDSS data collection or the MPA-JHU pipeline. Our leading theory is that the emission line peaked so steeply that it was mistaken for a cosmic ray and was thus masked out of the data.

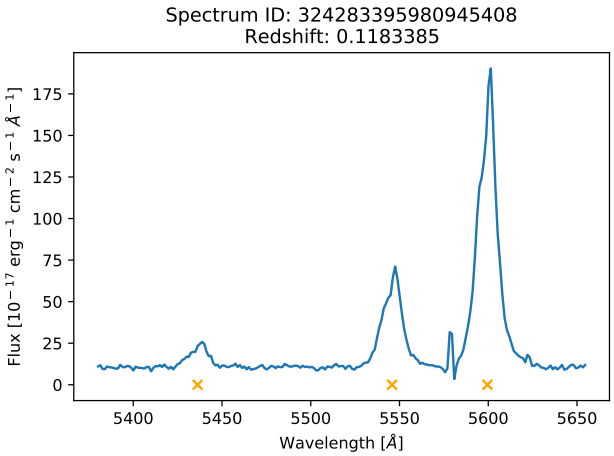

Figure B.5 Example of a spectrum with strong blue-wing asymmetries in the emission lines. The [O III] lines appear somewhat broadened as well.

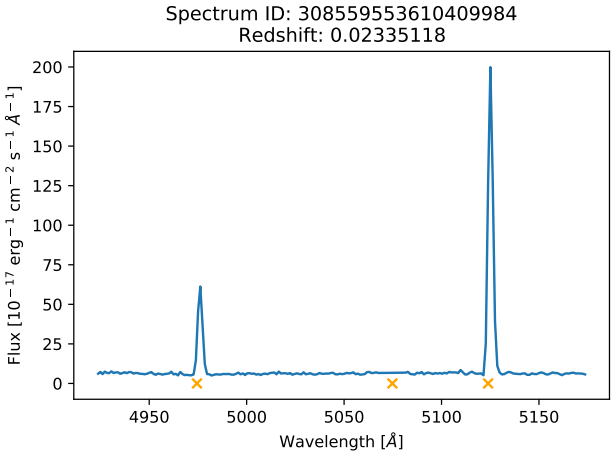

Figure B.6 Spectrum in which the [O III] λ4959 Å emission line is missing and seems to have been clipped or masked out.

Bibliography

[1] Franco D. Albareti et al. “Constraint on the time variation of the fine-structure constant with the SDSS-III/BOSS DR12 quasar sample”. In: *MNRAS* 452.4 (Oct. 2015), pp. 4153–4168. DOI: 10.1093/mnras/stv1406. arXiv: 1501.00560 [astro-ph.CO].

[2] John N. Bahcall, Charles L. Steinhardt, and David Schlegel. “Does the Fine-Structure Constant Vary with Cosmological Epoch?” In: *ApJ* 600.2 (Jan. 2004), pp. 520–543. DOI: 10.1086/379971. arXiv: astro-ph/0301507 [astro-ph].

[3] Burke Boyer et al. “Are More Distant Galaxies Hotter?” In: *Research Notes of the AAS* 3.8 (Aug. 2019), p. 119. DOI: 10.3847/2515-5172/ab3a41. URL: https://doi.org/10.3847/2515-5172/ab3a41.

[4] Yan-Mei Chen et al. “Evolution of the most massive galaxies to z= 0.6 – I. A new method for physical parameter estimation”. In: *Monthly Notices of the Royal Astronomical Society* 421.1 (Mar. 2012), pp. 314–332. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2011.20306.x. eprint: https://academic.oup.com/mnras/article-pdf/421/1/314/3132359/mnras0421-0314.pdf. URL: https://doi.org/10.1111/j.1365-2966.2011.20306.x.

[5] L. R. Colaço, R. F. L. Holanda, and R. Silva. “Probing variation of the fine-structure constant in runaway dilaton models using Strong Gravitational Lensing and Type Ia Supernovae”. In: *arXiv e-prints*, arXiv:2004.08484 (Apr. 2020), arXiv:2004.08484. arXiv: 2004.08484 [astro-ph.CO].

[6] M. Dietrich, D. M. Crenshaw, and S. B. Kraemer. "Probing the Ionizing Continuum of Narrow-Line Seyfert 1 Galaxies. I. Observational Results". In: *ApJ* 623.2 (Apr. 2005), pp. 700–720. DOI: 10.1086/429121. arXiv: astro-ph/0501483 [astro-ph].

[7] M. S. Dimitrijevic et al. "The flux ratio of the [OIII] 5007,4959 lines in AGN: Comparison with theoretical calculations". In: (Oct. 2006). DOI: 10.1111/j.1365-2966.2006.11238.x. arXiv: astro-ph/0610848 [astro-ph].

[8] M. E. Galavis, C. Mendoza, and C. J. Zeippen. "Atomic data from the IRON Project. XXII. Radiative rates for forbidden transitions within the ground configuration of ions in the carbon and oxygen isoelectronic sequences". In: *A&AS* 123 (May 1997), pp. 159–171. DOI: 10.1051/aas:1997344.

[9] C. M. Gutiérrez and M. López-Corredoira. "The Value of the Fine Structure Constant over Cosmological Times". In: *ApJ* 713.1 (Apr. 2010), pp. 46–51. DOI: 10.1088/0004-637X/713/1/46. arXiv: 1002.4777 [astro-ph.CO].

[10] P. Leisy and M. Dennefeld. "Planetary nebulae in the Magellanic Clouds. I. Carbon abundances and TypeI PNe." In: *A&AS* 116 (Mar. 1996), pp. 95–116.

[11] Donald E. Osterbrock and Gary J. Ferland. *Astrophysics of gaseous nebulae and active galactic nuclei*. 2006.

[12] Yu. V. Petrov et al. "Natural nuclear reactor at Oklo and variation of fundamental constants: Computation of neutronics of a fresh core". In: *Phys. Rev. C* 74.6, 064610 (Dec. 2006), p. 064610. DOI: 10.1103/PhysRevC.74.064610. arXiv: hep-ph/0506186 [hep-ph].

[13] H. Rahmani, N. Maheshwari, and R. Srianand. "Constraining the variation in the fine-structure constant using SDSS DR7 quasi-stellar object spectra." In: *MNRAS* 439 (Mar. 2014), pp. L70–L74. DOI: 10.1093/mnrasl/slt183. arXiv: 1312.5324 [astro-ph.CO].

[14] M. Rosa. "A Possible Nonlinearity in IDS / Image Dissector Scanner / Data". In: *The Messenger* 39 (Mar. 1985), p. 15.

[15] P. J. Storey and C. J. Zeippen. "Theoretical values for the [OIII] 5007/4959 line-intensity ratio and homologous cases". In: *MNRAS* 312.4 (Mar. 2000), pp. 813–816. DOI: 10.1046/j.1365-8711.2000.03184.x.

[16] Chris Stoughton et al. "Sloan Digital Sky Survey: Early Data Release". In: *The Astronomical Journal* 123.1 (Jan. 2002), pp. 485–548. DOI: 10.1086/324741. URL: https://doi.org/10.1086/324741.

[17] Jean-Philippe Uzan. "Varying Constants, Gravitation and Cosmology". In: *Living Reviews in Relativity* 14.1, 2 (Mar. 2011), p. 2. DOI: 10.12942/lrr-2011-2. arXiv: 1009.5514 [astro-ph.CO].

[18] J. K. Webb et al. "Further Evidence for Cosmological Evolution of the Fine Structure Constant". In: *Phys. Rev. Lett.* 87.9, 091301 (Aug. 2001), p. 091301. DOI: 10.1103/PhysRevLett.87.091301. arXiv: astro-ph/0012539 [astro-ph].

[19] John K. Webb et al. "Search for Time Variation of the Fine Structure Constant". In: *Phys. Rev. Lett.* 82.5 (Feb. 1999), pp. 884–887. DOI: 10.1103/PhysRevLett.82.884. arXiv: astro-ph/9803165 [astro-ph].

Index