Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Nathan Rauch April, 08, 2025

Searching for Post-Merger Recoiling Black Holes

by

Nathan Rauch

Erin Bonning, Ph.D.

Advisor

Department of Physics

Erin Bonning, Ph.D.

Advisor

Alissa Bans, Ph.D.

Committee Member

Marco Tezzele, Ph.D.

Committee Member

2025

Searching for Post-Merger Recoiling Black Holes

By

Nathan Rauch

Erin Bonning, Ph.D.

Advisor

An abstract of a thesis submitted to the Faculty of Emory College of Arts and Sciences of Emory University in partial fulfillment of the requirements of the degree of Bachelor of Science with Honors

Department of Physics

2025

Abstract

Searching for Post-Merger Recoiling Black Holes

By

Nathan Rauch

When galaxies merge, their central supermassive black holes will ultimately merge as well. During the merger, they can gain velocities with respect to their host galaxies. This occurs when the emission of gravitational wave radiation during the merger is non-isotropic. In this project, we search for a spectroscopic signature of recoiling black holes in quasars. Quasars have multiple Doppler broadened emission lines. These lines can be identified as originating in different parts of the quasar depending on if the line is broad or narrow. The broad lines originate from gas gravitationally bound to the black hole, and narrow lines from further out in the host galaxy. We use spectral data from the Sloan Digital Sky Survey, and fit the emission lines with the software pyQSOFit. We search for objects of interest by calculating the velocity of the broad line with respect to the narrow line. We fit approximately 50,000 quasar spectra, focusing on broad Hβ and the narrow forbidden line [OIII]. The resulting velocities are consistent with the predicted range of recoil velocities of post-merger black holes. A number of outliers, with distinct spectral features such as double-peaked broad lines, were identified within the sample.

Searching for Post-Merger Recoiling Black Holes

By

Nathan Rauch

Erin Bonning, Ph.D.

Advisor

A thesis submitted to the Faculty of Emory College of Arts and Sciences of Emory University in partial fulfillment of the requirements of the degree of Bachelor of Science with Honors

Department of Physics

2025

Acknowledgments

I would like to first acknowledge my advisor, Dr. Erin Bonning, for her support throughout my undergraduate education and this project. After graduating from Oxford and arriving at the main campus of Emory University, her classes on astrophysics and general relativity inspired me to take up research, and as a result this honors thesis project. Her insight and assistance in this project was crucial to its success both in the process of analysis and the writing process. I would also like to acknowledge Dr. Alissa Bans for her amazing lectures and feedback on this project. I particularly enjoyed her observable astrophysics course, which provided the unique opportunity to take astrophysical observations as a student. I would like to thank my committee member, Dr. Marco Tezzele, who has helped to further my understanding of mathematics and provided great feedback on this project. I will always remember the great lectures in his course on partial differential equations.

Contents

List of Figures

1 Introduction

When galaxies merge, their central black holes will also end up merging. This merging of black holes has the potential to cause a kick velocity, which results from an uneven distribution of gravitational wave radiation at the time of the black hole merger. The purpose of this project is to use the spectra of many quasars to identify quasars that potentially contain recoiling black holes. This will be done by using the spectral properties of a quasar as they relate to the physical regions that make up a quasar.

This section will identify what a quasar is, what its spectrum looks like, and what gravitational wave radiation is, as these concepts relate to this project. This will be done by describing the unified model of Active Galactic Nuclei (AGN), examining and explaining the spectra of quasars, describing gravitational wave radiation, and explaining how these concepts can be used identify quasars that are recoiling with respect to their host galaxy.

1.1 What is an AGN?

An AGN is a small extremely luminous object in the center of a galaxy. The structure of AGN is described in the unified model: a black hole surrounded by several regions of material each with their own defining features. Furthermore, in the unified model, the classification of an AGN is determined by the angle at which it is viewed. AGN are also known for having some spectral features including, but not limited to, broad-band continuum emission, strong emission lines, and the potential for strong radio emission. The most luminous AGN are classified as Quasars.

The unified model of AGN presents a model for the structure of an AGN. This structure is made up of a central black hole surrounded by an accretion disk, a dusty torus surrounding the accretion disk, the broad line region, the narrow line region, and a radio jet [12], seen in Figure 1. The furthest inward region is the black hole region. This region

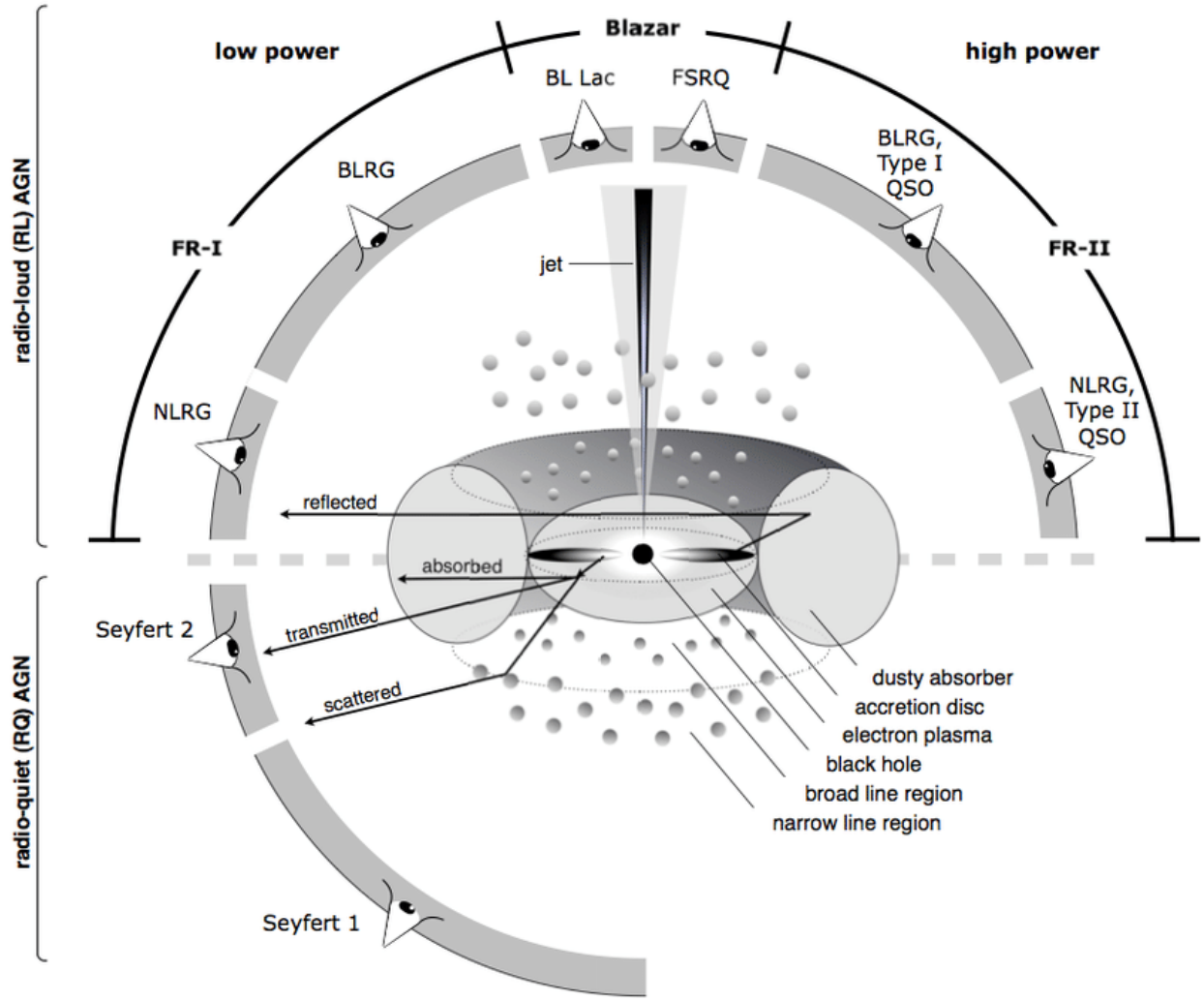

Figure 1: A visual representation of the unified model of AGN. This figure also shows the angle at which the AGN is viewed that would lead to specific classifications of the AGN. This figure comes from the Book *Active Galactic Nuclei* [2].

mainly outputs very high energy and mostly continuum, X-ray emission, which we will not directly measure for this project. The broad line region is just outside of the black hole region and has lower energy emission, strongest in UV and optical wavelengths. However, the broad line region is close enough to the central black hole that the gas has high enough velocities that the spectral lines that are produced will be significantly broadened due to Doppler broadening. Much further out from the black hole is the narrow line region. The narrow line region is much less dense and lower velocity than the preceding regions,

allowing for the emission of much narrower lines. One of the narrow lines emitted in the narrow line region is the forbidden line [OIII] [2]. It is relevant to note that depending on the orientation of the AGN, the observed properties and, as a result, the classification, will change, this can be seen in Figure 1. Many AGN are bright radio sources because of the two radio lobes directed in opposing directions with the AGN between them. These lobes are non-stellar in nature and are connected to the AGN that lies between the lobes. This indicates that the source of the radio lobes is the AGN, but not all AGN have strong radio emission [10].

AGN also have a very small angular size, extremely high luminosity, broad-band continuum emission, and strong emission lines. The small angular size of an AGN comes into play when the AGN is observed with the host galaxy surrounding it. The AGN is much smaller than its host galaxy and will often appear simply as a bright point within the galaxy. Typical AGN luminosity can range from about 1% of the power output of a typical galaxy to being 10^4 times the power output [10]. Some AGN will be much brighter and stand out, while others will be much harder to distinguish within the galaxy. It is important to also note that the observed brightness of an AGN can be affected by things such as extinction due to large amounts of dust and the effects of gravitational lensing. An example spectrum of a quasar can be seen in Figure 2. Emission lines are common features in the spectra of many objects. Emission lines are very abrupt increases in the emission of a spectrum focused around a specific wavelength, as opposed to absorption lines which are decreases in the emission around a specific wavelength. This wavelength corresponds to the energy of a photon from the de-excitation of an electron in a specific atom, ion, or molecule. It is very common in the spectra of stars and galaxies to see weak absorption lines as the primary lines. This is because stars and galaxies are sources of light, and materials that absorb and re-emit light at specific wavelengths will cause these absorption lines. In contrast, AGN stand out because of their strong emission lines. AGN have strong emission lines because the light coming from the accretion disk of an AGN

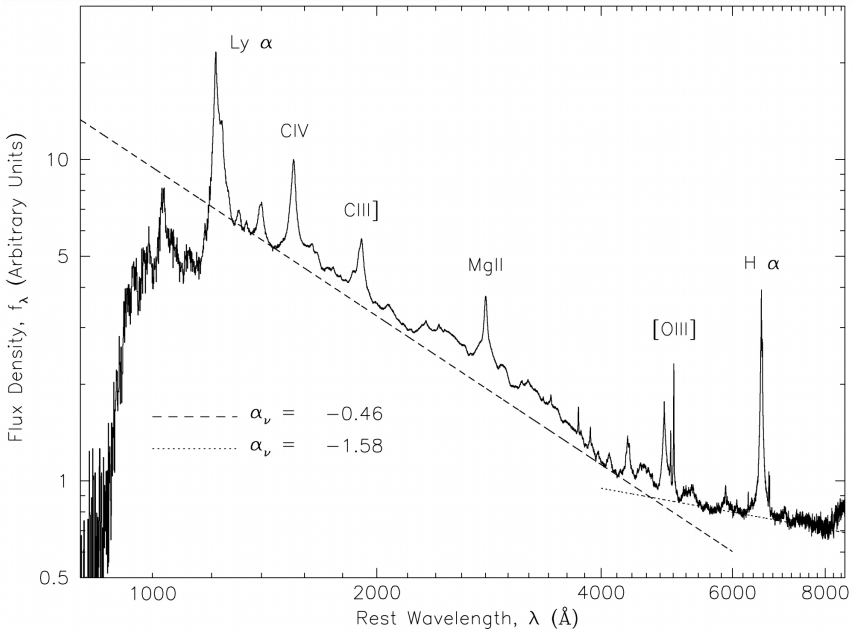

Figure 2: An example of a spectrum of a quasar. This figure comes from Vanden Berk et al. [17].

excites the material surrounding the black hole. There is also a line width distribution where some emission lines within AGN spectra are much more broad than others. There are even some emission lines that can be seen in both the narrow and broad classes of lines [10]. The line width can be linked to the origin of the emission within the topography of the AGN, which will be discussed later. An example of a spectrum with broad and narrow lines can be seen in Figure 3, where Hβ is visibly broad near the base of the line with a narrow point at the top because of the contributions of the broad Hβ and narrow Hβ lines respectively.

Quasar is a general term for an AGN that is high luminosity and emits light over the entire spectrum. Quasars have all of the AGN properties listed above to varying degrees [10]. The class of AGN that are dealt with in this paper are called quasars. Specifically,

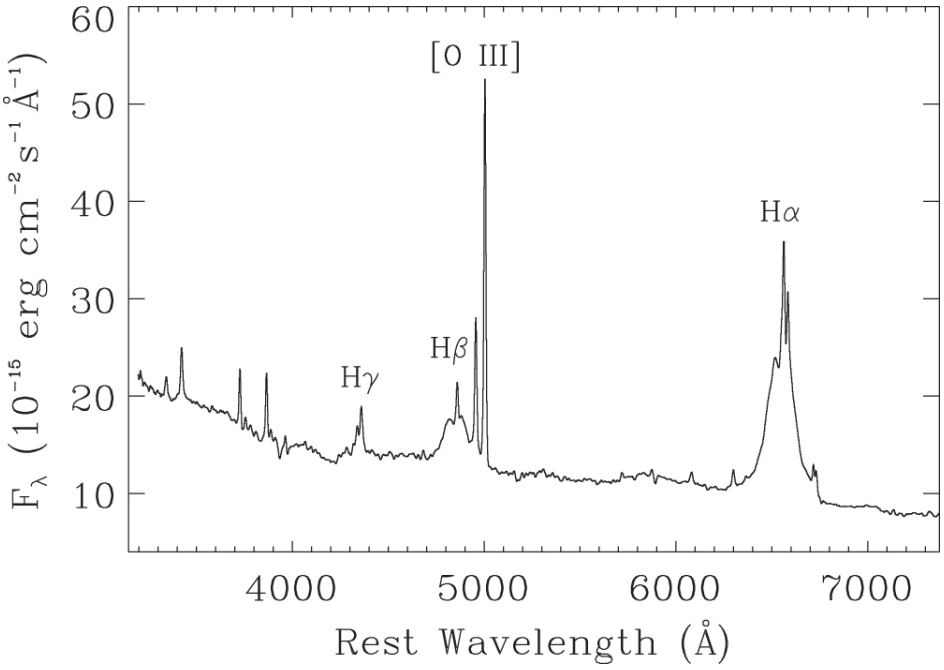

Figure 3: An example of a spectrum with visible broad and narrow lines. This figure comes from Pei et al. [13].

type 1 quasars because type 1 quasars have visible broad lines while type 2 quasars do not [12].

1.2 What is Gravitational Wave Radiation?

Gravitational wave radiation is the result of masses accelerating. This radiation is the change in the effect on spacetime by the acceleration of mass that can be observed. The sources of radiation that are large enough to have been detected are limited to very dense and massive objects interacting with one another, such as a binary black hole system. The luminosity from such a system can be calculated within the theory of general relativity.

Gravitational waves are ripples in spacetime that are caused by the acceleration of

masses. The release of gravitational wave radiation from a merger of black holes is the reason that the mass of the black holes is not conserved, it is radiated away as energy in the gravitational waves. In order to model the energy lost due to gravitational wave radiation, it is necessary to introduce the reduced quadrupole moment tensor.

$$\text{Ŧ}^{jk} \approx \sum_{objects\ i} m_i(x_i^j x_i^k - \frac{1}{3}\eta^{jk} r_i^2) \tag{1}$$

Where indices are indicative of Einstein summation notation. This tensor is defined, like moment of inertia, in relation to the rotational motion of the black holes in the binary system. Define the trace-reversed metric perturbation,

$$H_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} \tag{2}$$

where $h_{\mu\nu}$ is the metric perturbation and $\eta_{\mu\nu}$ is the metric of flat spacetime. The metric perturbation is a measure of how the metric, $g_{\mu\nu}$ is changed from that of flat space because of gravitational waves.

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{3}$$

Now, through using the weak-field Einstein equation satisfying the Lorentz gauge condition with the trace reversed metric perturbation with raised indices, it is possible to approximate for slow moving objects, v«c, the trace-reversed metric perturbation with raised indices through analogy with electrodynamic electric potential, seen Moore p.384 [11].

$$H^{\mu\nu} = 4G \int_{src} \frac{T^{\mu\nu}}{R} dv|_{t-R} \tag{4}$$

$T^{\mu\nu}$ is the stress energy tensor containing source information, and R is the distance between the point of interest and the black hole merger. Using the identity

$$\int_{src} T^{jk} dv = \frac{1}{2}\frac{d^2}{dt^2}\int_{src} T^{tt} x^j x^k dv = \frac{1}{2}\frac{d^2}{dt^2}\int_{src} \rho x^j x^k dv = \frac{1}{2}\ddot{I}^{jk} \tag{5}$$

it is found that

$$H^{jk} = \frac{2G}{R}\ddot{I}^{jk} \tag{6}$$

and $H^{tt} = \frac{4GM}{r}$ and $H^{it} = 0$. Since, $H^{xx}_{TT} = \frac{1}{2}(H^{xx} - H^{yy})$ it can be shown that,

$$H^{jk}_{TT} = \frac{2G}{R}\ddot{\mathbb{I}}^{jk} \tag{7}$$

Now, since this is since this is transverse traceless $h^{jk}_{TT} = H^{jk}_{TT}$. It is possible to find the flux by finding the tt component of the stress energy tensor.

$$Flux = T^{tt}_{GW} = \frac{1}{32\pi G} < \dot{h}^{jk}_{TT}\dot{h}^{TT}_{jk} >= \frac{G}{8\pi R^2} < \dddot{\mathbb{I}}^{jk}_{TT}\dddot{\mathbb{I}}^{TT}_{jk} > \tag{8}$$

Now the luminosity is found by integrating the flux over the surface of the sphere at R.

$$L = \frac{-dE}{dt} = \int_{sphere} Flux \cdot dA = \frac{G}{5} < \dddot{\mathbb{I}}^{jk}_{TT}\dddot{\mathbb{I}}^{TT}_{jk} > \tag{9}$$

The quadrupole moment tensor is related to the rotational motion of the binary mass system. The luminosity is the rate at which energy is radiated away as gravitational waves [11]. From the equation above, and the general understanding that energy is leaving the system in the form of gravitational wave radiation, it can be inferred that the changing quadrupole moment tensor is related to the coalescing orbits of the masses resulting from a loss of energy in the system. An uneven distribution of this emission is caused by differences in mass and spin of the masses.

1.3 Quasar Mergers

It is not uncommon for galaxies to gravitationally interact and merge in the universe. When they do, the AGN in said galaxies will gravitate towards the center of the new merged galaxy and gravitationally interact with one another. This causes the central black

holes of the AGN to coalesce and eventually merge.

When quasars merge, the central black holes within the quasars will begin to gravitationally interact with each other. This gravitational interaction leads to a binary black hole system, where the black holes will orbit each other as they coalesce. When black holes are in close orbit, they emit gravitational wave radiation. Furthermore, when the black holes reach the point of merging, they release a violent burst of gravitational wave radiation. [1]. This phenomenon can be seen in LIGO data for merging black holes, an example of this is in Figure 4. It is expected that an abrupt release of energy would occur at the moment of black hole merger, leading to a kick velocity in the resulting black hole, depending on the direction of gravitational wave emission at the moment of merger.

1.4 What are Kicks and How can we Identify Them?

In the case of non-isometric radiation of gravitational wave radiation, there can be a reactionary, "kicked", momentum in the merged black hole system caused by Newton's third law. As a result of the extreme amount of energy release from a black hole merger, the kick velocity of the merged black hole can be significant enough to eject the black hole from its host galaxy [4]. Quasars are a great candidate to demonstrate this because they contain black holes with the possibility of merging, and they have host galaxies around them that contain the emission line regions with differing distance scales with respect to the central black hole. Figures 5 and 6 give a visualization of the difference in distance scales of the broad line region and narrow line region.

Kick velocities can be measured using the properties of the AGN that have merged, most notably, the use of the redshift of spectral lines. The spectrum of an AGN has spectral lines that come from the broad line region and narrow line region respectively. Using the fact that the broad line region is relatively close to the kicked black hole, and gravitationally bound, and that the narrow line region is significantly farther away, and not gravitationally bound, the redshift of the broad line with respect to the redshift of the

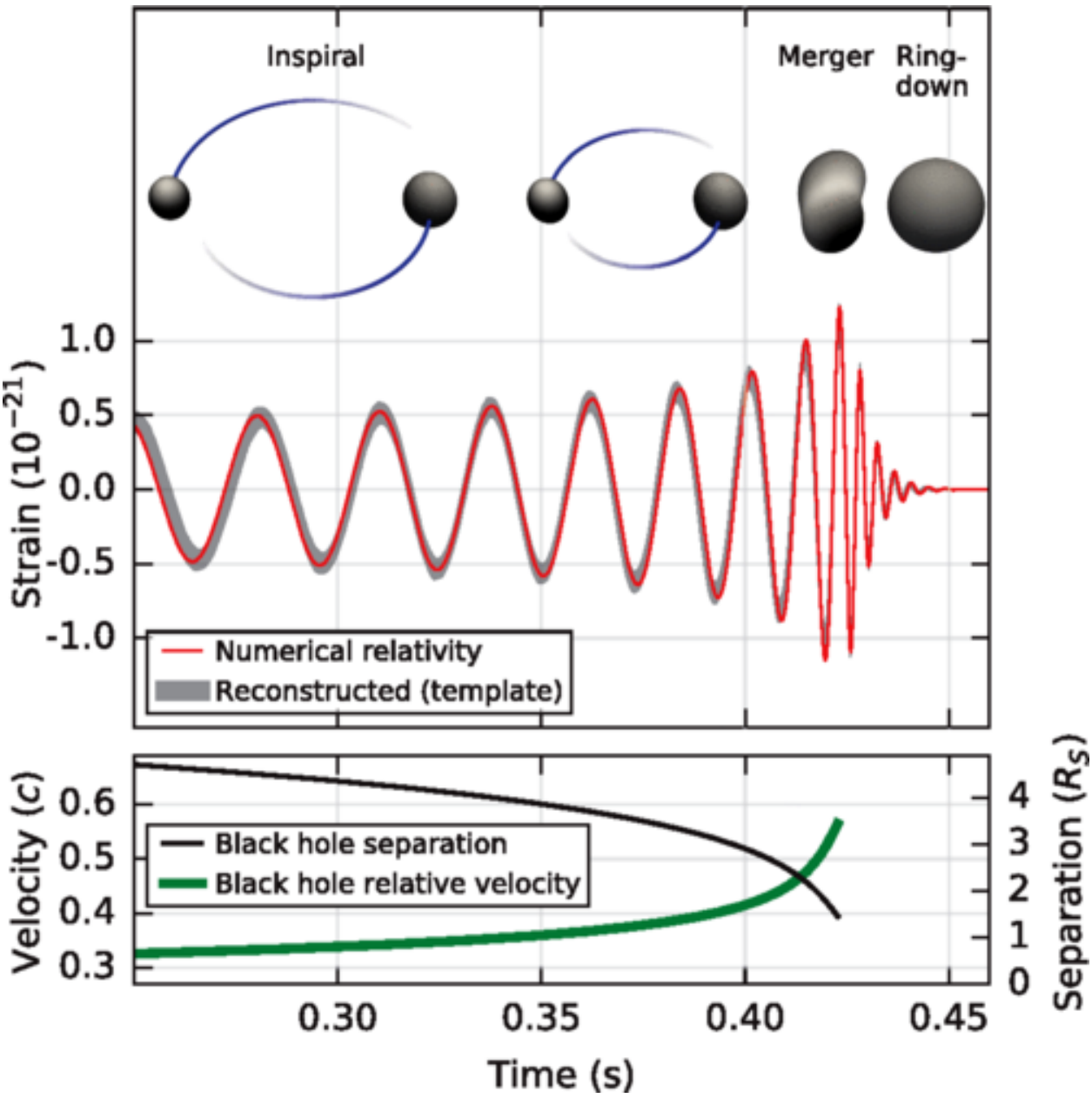

Figure 4: A visual representation of what the LIGO data for merging black holes looks like. Depicts the progress of coalescence at each stage of the waveform. Figure from Abbott et al. [1].

narrow line can be used as a proxy for the redshift of the black hole with respect to the host galaxy. With this knowledge, it is easy to find the velocity of the black hole with respect to the host galaxy.

Redshift is the Doppler shift of light. This means that the redshift of an object comes

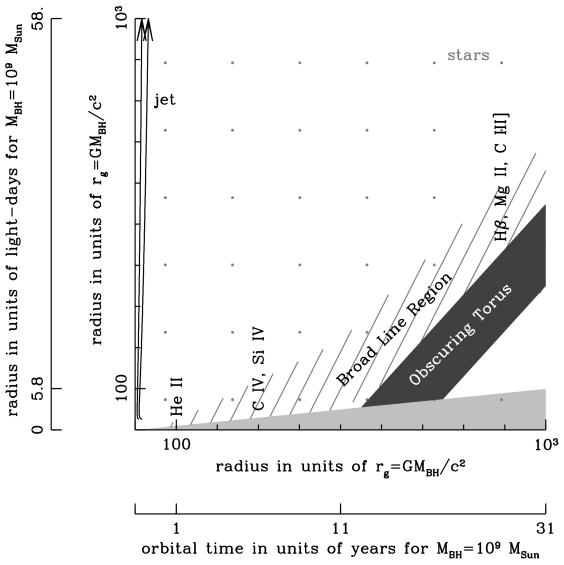

Figure 5: Figure displaying scale of the broad line region with the black hole being at zero on the x-axis. Figure from Hall [9].

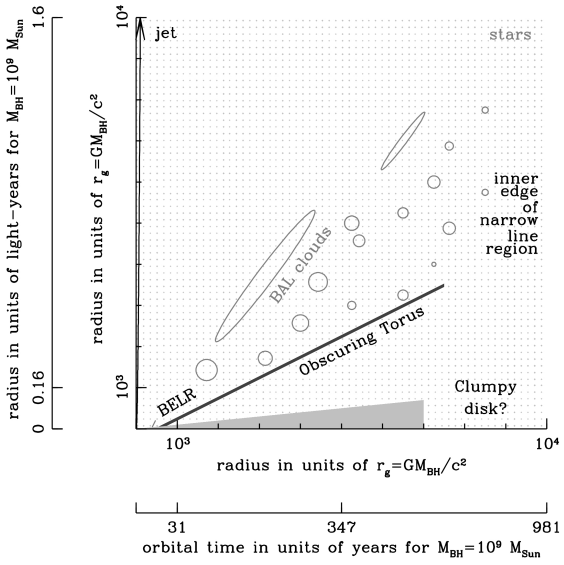

Figure 6: Figure displaying the scale of the narrow line region with the black hole being at zero on the x-axis. Figure from Hall [9].

from the change in frequency of the wave of light as a result of the motion of the source of the light with respect to the observer. The general formula for this is given as follows

$$z = \frac{\Delta\lambda}{\lambda} \tag{10}$$

where z is the redshift, $\Delta\lambda$ is the difference in the observed wavelength of light and the rest wavelength, and λ is the rest wavelength of light. The velocity is the redshift times the speed of light. As a result of this project dealing with extragalactic objects such as quasars, it is important to note that there is a redshift associated with the expansion of the universe. Hubble's law states that as objects get increasingly farther away they recede quicker due to the expansion of the space between observer and object, seen in the equation

$$v = H_0 d \tag{11}$$

where v is the velocity of recession, H_0 is the Hubble constant, and d is the distance of the object [6]. As a result of this property and the relation between velocity and redshift, it has become common practice to use redshift as a proxy for the distance of very distant objects because the velocity that the object has through space is inconsequential when compared with the velocity from Hubble's law. Using this property, if there is a systemic redshift of the quasar as a whole and a different redshift of the central black hole, the Doppler shift formula could be used to calculate the velocity of the black hole with respect to its host galaxy.

Since the black hole is not directly observable, a proxy for the black hole is required to measure the redshift of the black hole. This proxy will be the broad line region. Because the broad line region is very close to the black hole itself, the redshift of the broad line Hβ, and as a result, the motion of the broad line region will be used as a stand-in for the motion of the black hole itself. The spectral line Hβ was chosen to be the representative for the broad line region and black hole because it is near a strong narrow line and the most prominent broad line in most spectra, Hα, has a peak that is difficult to distinguish from nearby narrow lines. Measuring the redshift of the host galaxy itself can prove to be challenging because of the immensely luminous nature of the quasar. Therefore, a representative is required for the redshift of the host galaxy as well. This representative can be a spectral line from the narrow line region since the narrow line region is far enough out from the black hole that it should not be affected by the motion of the black hole. This is because the broad line region is 100 times further from the black hole when compared to the broad line region. Since gravity can be thought of as an inverse square law with respect to distance, this means a particle in the narrow line region will experience 10000 times less gravity from the black hole than an identical particle in the broad line region. The redshift of the [OIII] narrow line will be used as a stand-in for the host galaxy of the quasar. This project will also use the SDSS redshift to compare with the narrow line region's redshift. This will show that there is a slight difference between the redshift of

the narrow line and the redshift of the object that SDSS calculates. Using the comparative redshift of the broad and narrow line region will create an estimate for the motion of the black hole with respect to its host galaxy.

This project will look for evidence of recoiling black holes in a large sample of quasars by measuring the relative velocity of the quasars' broad line regions with respect to their narrow line regions.

2 Methods

The project searches for evidence of post-merger recoiling black holes by measuring the velocity of the broad lines of quasars with respect to their narrow line regions. This requires spectral data of quasars, which will be obtained from the Sloan Digital Sky Survey. The emission lines will be fit to each spectra by using the program pyQSOFit and the velocity of the broad line region with respect to the narrow line region will be calculated using the peak wavelengths of the Hβ broad line and [OIII] narrow line. Once these velocities are calculated, it will be necessary to identify potential errors in the fitted spectra of specific objects and errors in the methodology used to fit the spectra and calculate velocities. This section will focus on how this project used SDSS and pyQSOFit, and how the initial data analysis was performed.

2.1 Sloan Digital Sky Survey (SDSS) & Spectra

The Sloan Digital Sky Survey (SDSS) is one of the largest and most comprehensive astronomical surveys. It has created the most detailed three dimensional maps of over one third of the night sky. Furthermore, the survey has spectra for more than three million objects [15]. The Sloan database is made up of the culmination of many individual surveys, including but not limited to, the Imaging and Legacy Survey, Supernova Survey, SEGUE 1/2, and (e)BOSS. SDSS releases its data in batches, which are named by sequential number. This project only deals with quasar spectra, so it will only use the data from data release 16 (DR16), as it was the last update to quasar spectra. SDSS releases updates for their quasar spectra in the form of the quasar catalog, the most recent being DR16Q. Within DR16Q there are data for 750,414 quasars with 225,082 being new [15]. The spectra for quasars that are given in SDSS cover the range of wavelengths from 3650 to 10400 Å, but all surveys cover the redshift of interest for this project [18]. An example spectrum can be seen in Figure 7.

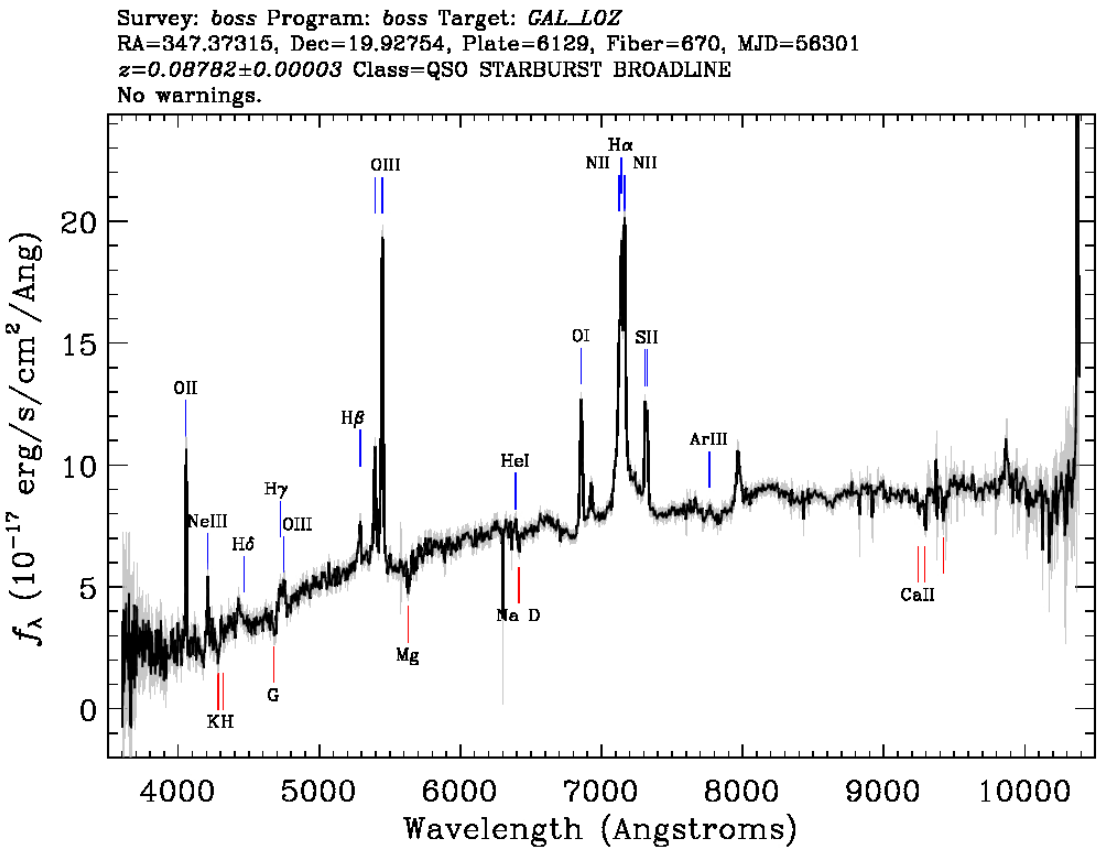

Figure 7: Sample SDSS quasar spectrum. Shows typical wavelength range of data and emission lines detected.

In order to download the data from SDSS, the first step is to submit a spectroscopic query [1] to SDSS for a CSV file with the redshift, plate number, MJD, and fiber of all quasars with the given input parameters. Plate number, MJD, and fiber are commonly used to identify the specific observation of the spectrum of an object. The plate and fiber are related to the technology that was used for observing, where there would be plates with fibers in them all oriented towards an object of interest, so each unique plate and fiber combination would be a unique observation. The MJD is the modified Julian date, Julian date - 2,400,000.5, of the observation, which is represented as an integer.

For the purposes of maintaining as many quasars in the sample as possible, this project only limited the query to quasars with broad lines, type 1 quasars with a redshift of at or

[1] `https://skyserver.sdss.org/dr16/en/tools/search/SQS.aspx`

below z=0.55. SDSS has a total redshift range for quasars of z=0.0333 to 7.00 [15]. This project will limit the redshift to be less than 0.55 because of the spectral lines used in this survey. The spectral line of Hβ has a rest wavelength of 4861 Å and the spectral line [OIII] has a rest wavelength of 5007 Å. Since the [OIII] line has the higher rest wavelength, the maximum redshift allowed for the survey needs to keep the [OIII] line within the range of wavelengths of the SDSS data. The shortest wavelength for the maximum wavelength of an SDSS spectrum is 9000 Å, but in the interest of being safe consider the maximum wavelength allowed to be 8000 Å. Therefore, to ensure that the [OIII] line is always below this wavelength in the data from SDSS

$$z < \frac{8000 - 5007}{5007} = 0.598. \tag{12}$$

Once again, in the interest of keeping the [OIII] line well within the wavelength range of the data, the maximum redshift was rounded down to z=0.55.

After recieving the csv file, it was opened in excel to create the url for each spectrum, in order to download them using the wget function. Each object could have a spectrum from any one of the surveys of SDSS, and the survey affects the proper url of the spectra. In order to account for this, there was a column in the excel file made for each of the four surveys mentioned in the above paragraph, and the concatenate function along with the plate, MJD, and fiber numbers of each object was used to create the urls for each survey for every object. Since there is useful data in all of the surveys, it is necessary to identify a way to match objects with their proper survey. In order to match the objects to their proper surveys, and therefore, their proper urls, the plate numbers were used. Plate number was used because of the association found between plate number and the SDSS survey. A plate number of greater than 3519 was associated with the eBoss survey, a plate number of exactly 2957, 2640, 2962, 3000, 3002, 3003, 3005, or above 3006 and below 3520 was the Segue survey, and all else was DR12. The code in Appendix A was used to take the excel

file as the input and to output a txt file with the list of urls that can be fed to the wget function.

With the urls, it is possible to now use the wget function to download the spectrum of every object. The function used to download the files is,

```
wget -i QSO_URLs.txt -P C:\Users\robos.000\Downloads\Spectra\ -o
wget.txt --rejected -log=wget-rejected.txt
```

where QSO_URLs.txt is the file with the urls, the directory after the -P is the output directory, -o wget.txt creates a file that the output text is stored in, and the final piece makes a file for the error outputs specifically. This helped when troubleshooting because sometimes there will be a server side error while downloading massive amounts of spectra. The spectral data for all quasars within the targeted redshift range of 0.55 or under were downloaded, with a resulting 51,000 quasars. After downloading all of the spectra, the next step was to fit the spectral lines to them.

2.2 PyQSOFit & Fitting Data

In order to fit the lines of the spectra that were downloaded, it was necessary to use a fitting program. This project used a program called PyQSOFit. This program was made by Dr. Hengxiao Guo in collaboration with Dr. Yue Shen and Dr. Shu Wang [8] [16]. The program pyQSOFit fits the continuum, Fe II, spectral lines, and the host galaxy as seen in Figure 8. First, the host galaxy and quasar components of the spectrum are found by using principle-component analysis (PCA). This means that there are spectra of both host galaxies and quasars that are used as eigenspectra, in order to find the least squares combination of these eigenspectra that recreates the data input into the program. The continuum, meaning the spectrum without emission and absorption lines, is fit with several components including power-law fit, polynomial fit, and Balmer continuum [8]. Fe II is also removed from the spectrum by pyQSOFit. Fe II is an ion that has spectral

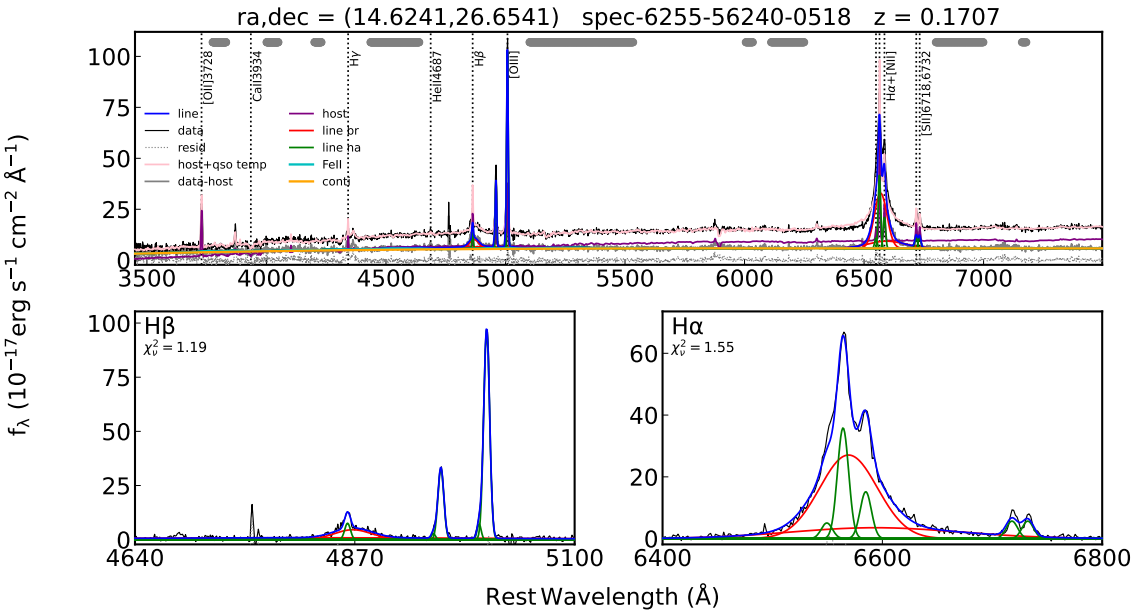

Figure 8: Quasar spectrum fitted using pyQSOFit. This displays the fitted narrow lines, broad lines, SDSS data, and continuum in the main image. The line complexes are the images below, and give a zoomed in view of the wavelength ranges shown on the bottom of each plot.

lines that show up throughout the spectrum of quasars, and as a result it is necessary to subtract these contributions from the spectrum to see other features. Since this project is most interested in the broad and narrow lines, it is important to note that this program produces a least-squares fit for the spectral lines of a quasar by using Gaussian references for broad and narrow lines.

When the user starts the program, pyQSOFit, they set wavelength boundaries for where they want the program to search for each spectral line. The program takes the input file for the quasar's spectrum and uses the input redshift to find a rest frame version of the spectrum. The program then searches within the user-defined boundaries for the best fit in comparison with the model for each line within the program. Once the program has deciphered what within the data best fits each line, it calculates the peak wavelength, the full width half-maximum of the line, and the equivalent width of the line. (The equivalent width of a spectral line is the width of the continuum, at the position of the line, that

the area of the spectral line would take up under the curve.) In this project, the peak wavelength is the quantity of interest.

SDSS calculates the redshift of an object by using PCA to make a least squares fit of the object's spectrum at different redshifts of the eigenspectra. The lowest chi-square value of these fits will ultimately be the SDSS redshift of the object [15]. Details can be found on the SDSS webpage [2]. We used this program to find the SDSS redshift and the [OIII] line redshift for every quasar in the survey. Then we made a histogram of the difference of these values, seen in Figure 9.

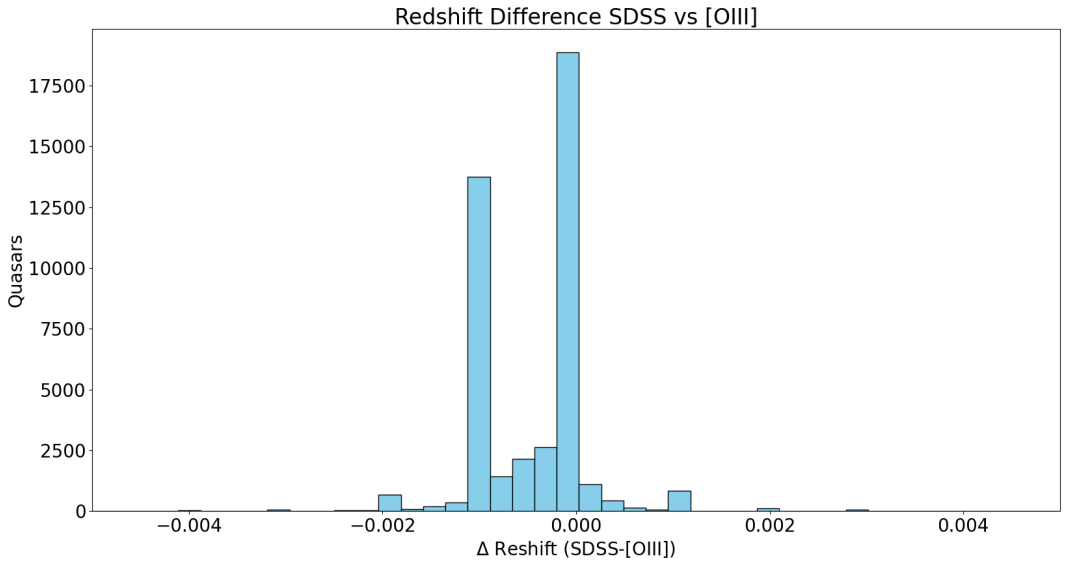

Figure 9: Histogram of the difference in redshift values from the [OIII] fit spectral line and the SDSS redshift for all objects in the sample.

The histogram shows that there does not appear to be a very significant difference in the redshifts of SDSS and the [OIII] lines. However, the [OIII] narrow line redshift seems to tend to be very slightly greater than that of the SDSS redshift. This means that the SDSS given redshift is a slight under approximation of the redshift of the narrow line region.

This project used the redshift of [OIII] as a stand-in for the redshift of the host galaxy. This is because the narrow line region is significantly separated from the quasar itself

[2] https://www.sdss4.org/dr17/algorithms/redshifts/

(roughly 1000 light years away), which allows for the assumption that the narrow line region will see a negligible velocity resulting from the quasar's kicked velocity. The program will always fit something to be the spectral line from within the data inside the established wavelength range. That can also mean that if a line is not distinguishable from noise that the fitted line is not able to be trusted. However, it is also possible for the program to be unable to fit lines to the spectrum. In this case, it is because the spectrum is missing data. There is also a unique case where there is missing data, but not missing enough data to prevent the program from fitting the spectrum, seen in Figure 10.

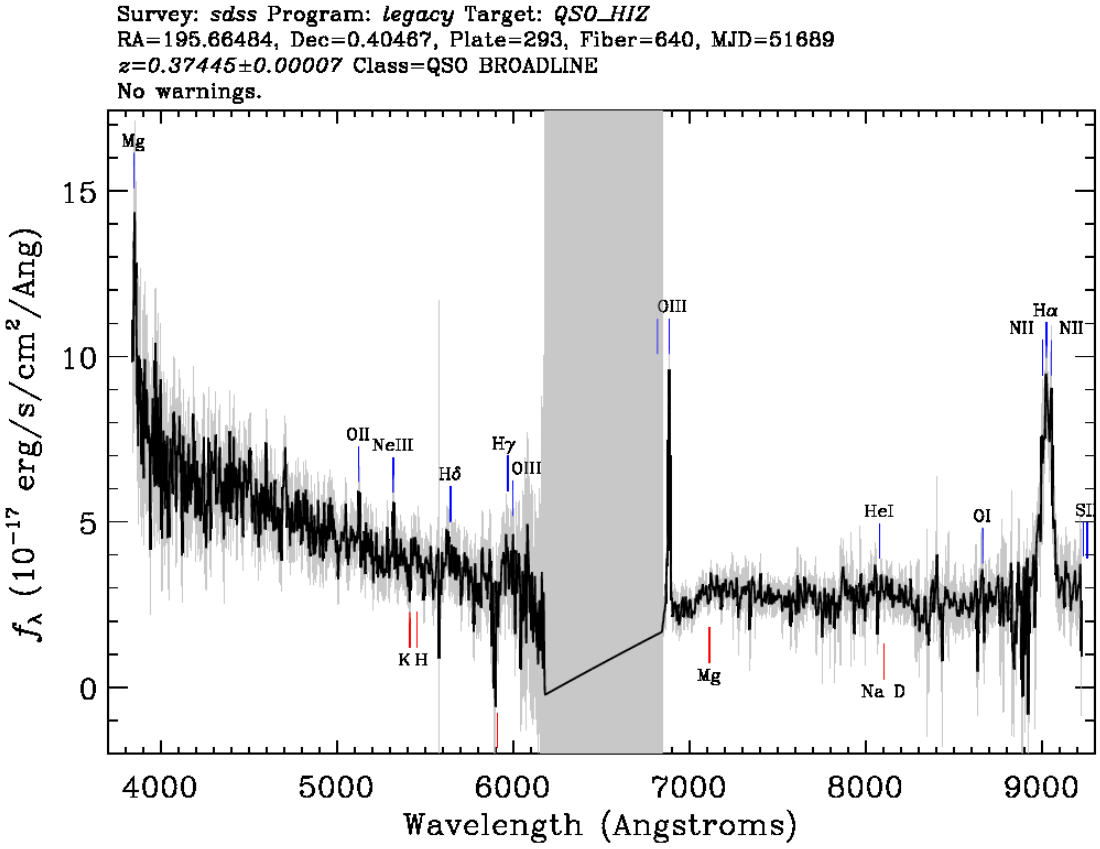

Figure 10: SDSS spectrum with missing data. This is a unique spectrum in that it can be fit with pyQSOFit, but it does not contain the spectral lines necessary for this project.

After removing the spectra that could not be fit or that were missing necessary data, there were 43,385 quasars left in the sample.

When pyQSOFit fits a spectrum, it outputs a .fits file, which contains the numerical

values for the fitted lines, and a pdf file that shows the visual representation of the data and fitted lines. The basic information of the object is contained at the top of the image,

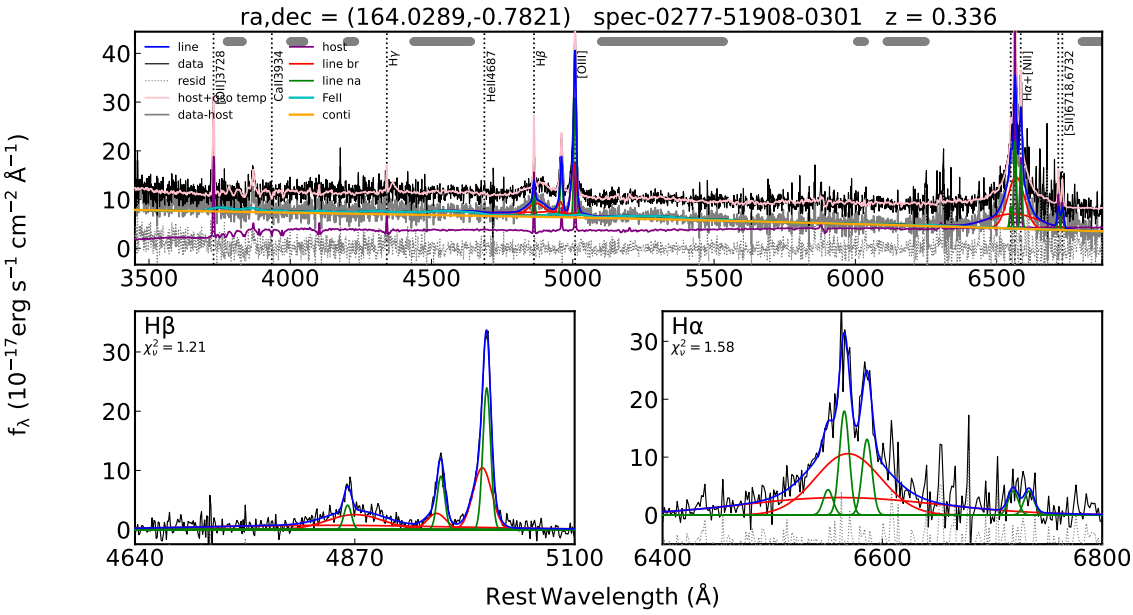

Figure 11: An example of the pdf image output from pyQSOFit for a spectra. Displays emission lines, continuum, SDSS data, and host galaxy fit.

including the position and redshift. The image is of the spectrum of the object after it has been shifted to rest-frame wavelengths, according to the redshift of the object as seen at the top of the image. Many spectral lines are labeled in the upper image, and the lower images are zoomed in views of the Hβ, Hα, and when available MgII line complexes. The x-axis is the wavelength in angstroms and the y-axis is the flux. Within the images there are various lines, with four being important for this project. The black line represents the data from SDSS for the spectra, the green line is the fit for narrow lines, the red line is the fit for broad lines, and the blue line is the cumulative fit line for the spectra.

The program pyQSOFit has some notable features for fitting the spectral lines to the data that are input. The fitting of each line is listed as three values in the output fits file. The first value is the scale, which is the amplitude of the line. The next value is called the centerwave. The centerwave is the peak wavelength of the line. The third value is

the sigma, which is the standard deviation of the line, as each line is approximated as a Gaussian. The lines that are fit are separated into line complexes, which are grouped by a range of wavelengths centered about the line that the complex is named after. These complexes include MgII, Hβ, and Hα. These complexes each have a calculated signal to noise ratio and chi squared fitting parameter, as well as an area, equivalent width, and full width half maximum for the line with the highest amplitude in the wavelength region.

In order to select quasars that have a significant velocity with respect to their host galaxy, I wrote a script that uses PyQSOFit to fit the lines of the spectrum and then calculate the redshift of the Hβ broad line with respect to the [OIII] narrow line. It was also important to flag the spectra of objects that result in fitting errors in order to ensure no oddities in the survey. The fitting function, from pyQSOFit's github example, was a Jupyter notebook file made for fitting the data of one spectrum given the input data for the spectrum from SDSS. Because this project required fitting many thousands of spectra, it was necessary to create a loop for the function so that it could fit all of the spectra within a directory without needing to be manually restarted. The loop functioned properly, but couldn't handle the sheer amount of data being fit. This was because Jupyter notebook saved the resulting output of the function as it was running, which took up too much memory and caused the device to crash. This issue was fixed by restructuring the code into two sections, seen in appendix B, that could be run in the command line as python files without using working memory.

2.3 Data Reduction

In the interest of flagging significant data it was necessary to make a cut for parameters that are more common in noisy or incoherent results. First, trying a chi-squared value cutoff of 10, this cutoff means that any fit with a chi-square value for Hβ of greater than 10 will be ignored because it is not a close enough fit. Secondly, adding an arbitrary cutoff for equivalent width and full width half-maximum for Hβ, this cutoff was made because

some quasars were fit with Hβ broad lines that were indistinguishable from the continuum itself because of how wide and small they were. Lastly, the cutoff was changed to a signal to noise ratio cutoff for the Hβ line complex of 5. This means that the fitted Hβ line needs to have a signal to noise ratio of at least five in order to be considered for flagging. When this cutoff was made, it became apparent that the other cutoffs were unnecessary, and therefore the only cut for parameters that is used in the final iteration of the survey is the signal to noise ratio of Hβ.

The above method worked quite well for the preliminary tests of the fitting program that was being used, but the primary goal of the project is to make a distribution of the velocities of all the objects, compare this distribution with the expected results, and analyze the spectra of the objects with the most extreme velocities. For this purpose, it is more useful to simply fit all of the spectra without conditions and separate the unusable data after the fact. Since a script for the fitting program to fit all of the data and store them in a directory was already designed, this was quite straightforward. After finding the values for the velocity of every object, they were plotted as a histogram of the velocities, seen in Figure 12.

There were multiple issues with this first histogram. Most notably, there are many velocities that do not agree with the expected values of the velocity of a recoiling black hole. This is because the velocities are much too high in magnitude. This is because of an issue with how the fitting program defines the line complexes. The Hβ line complex includes the [OIII] line, and the Hβ broad peak wavelength is the broad line with the greatest amplitude in the Hβ complex. There is also an issue where the wing component of the [OIII] line can be flagged as a broad line, despite [OIII] not being capable of being a broad line. When this [OIII] line is the highest amplitude "broad" line, the velocity is massively impacted because of the parameter that was used for the Hβ line peak, seen in Figure 13. In order to circumvent this, the peak wavelength of the highest amplitude of the two specific Hβ broad line fits was used as the peak wavelength of the Hβ broad line.

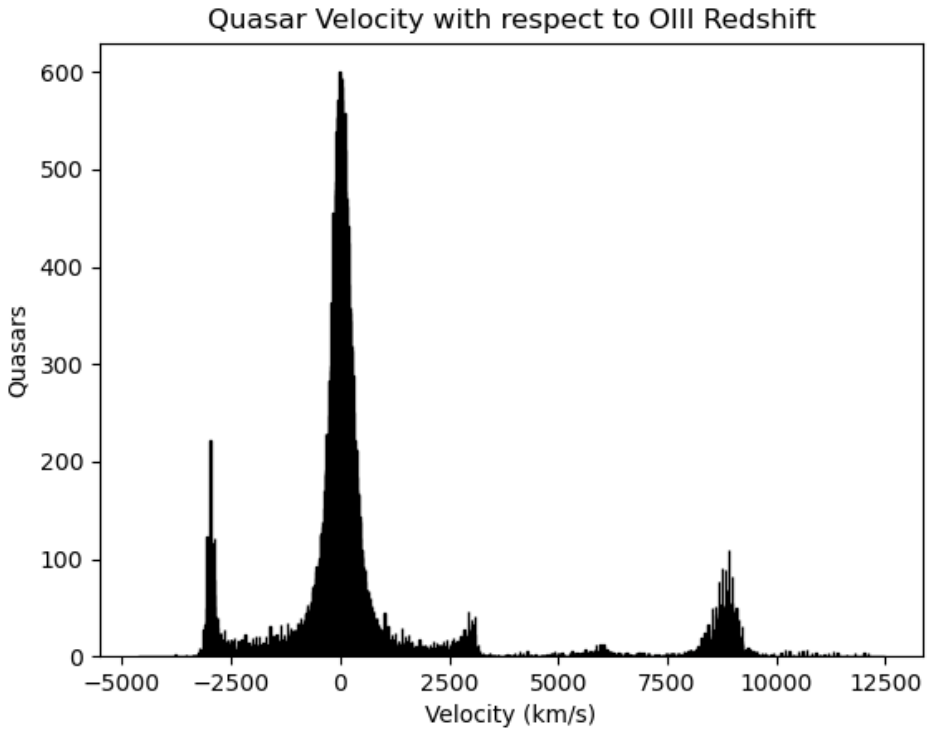

Figure 12: Initial histogram of all quasar velocities with respect to [OIII], with no conditions. Contains velocity values which are much greater than what is expected as the maximum recoil velocity of 4000 km/s. Also displays regions of unexpected density around ± 3000 km/s.

This solved the issue of velocities being far too high.

There was also an issue of there being two peaks at roughly 3000km/s of recoiling velocity, which is not an expected part of the distribution, under the assumption of isometry. This is because there is a lot of noisy data, and the fit for the peak wavelength cannot be trusted in these cases. Using the same method as in the preliminary testing of the fitting program, an SNR cut of the data, this problem was solved. An SNR cutoff of SNR>4, seen in Figure 14, seems to solve the issue of peaks, but it did not preserve as many spectra as would be optimal for the plot. This method reduced the data set from 43,385 quasars to roughly 11,000 quasars.

Afterwards, trying an SNR cutoff of 2, along with the method of using the most significant H beta broad line fit as the stand in for the quasar's redshift was effective at

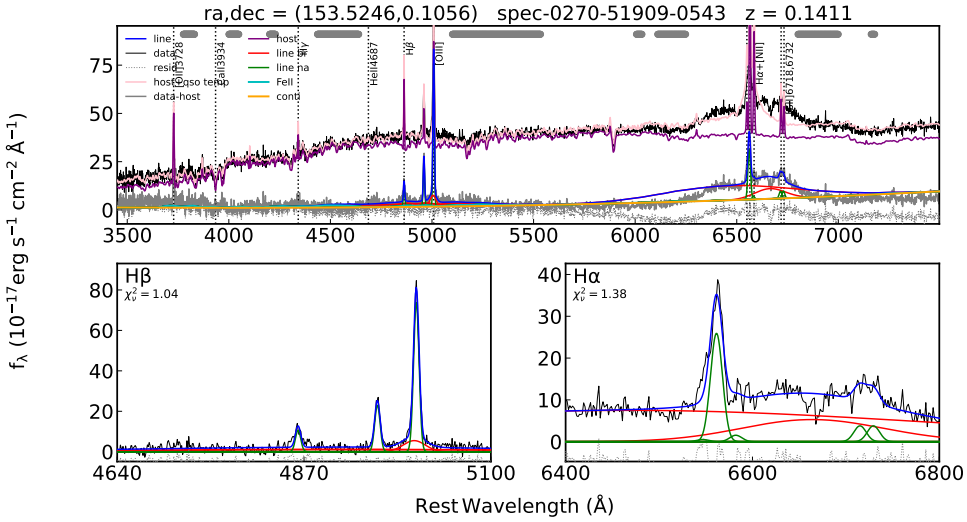

Figure 13: Spectrum with false "broad" wing component of [OIII]. This line is the visible red bump under the [OIII] narrow line at 5007 angstroms.

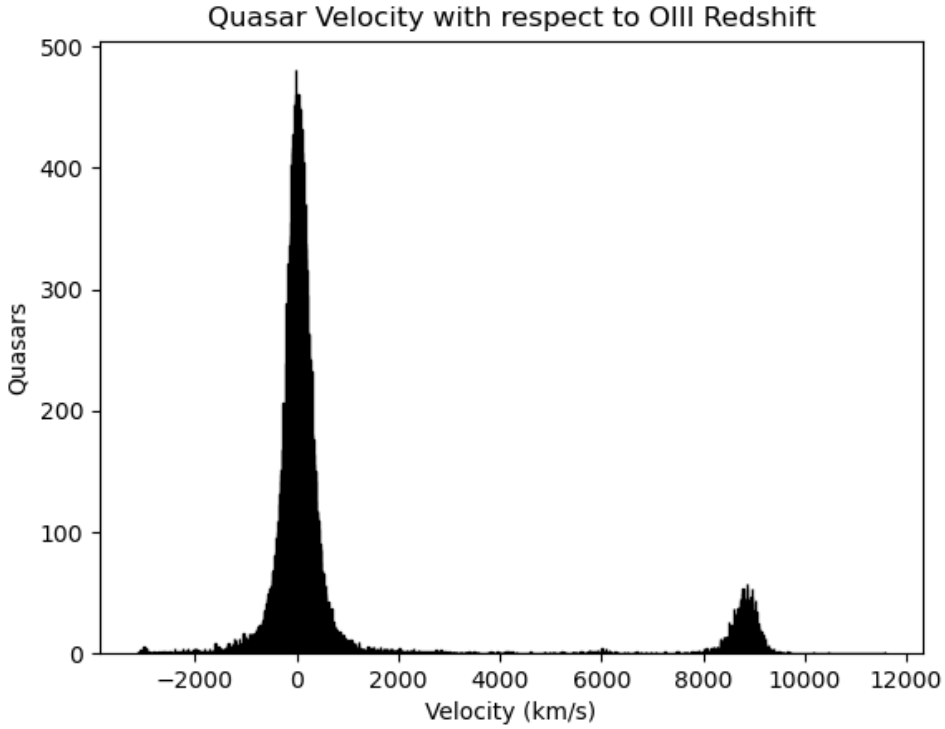

Figure 14: New histogram of velocities of the quasars with an SNR cutoff of 4. Still contains unexplained high velocity objects. Reduced number of objects in sample from roughly 43000 to 11000.

both reducing the peaks and eliminating nonsensical velocity values from the plot, while only limiting the sample to 26,000 quasars, seen in Figure 15. However, there was still an

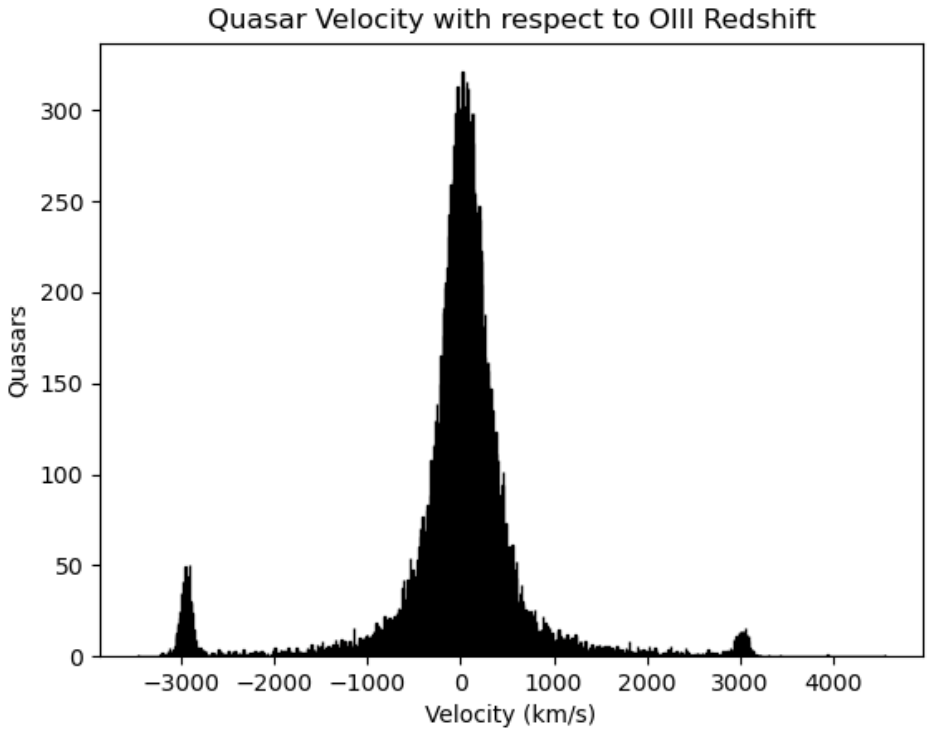

Figure 15: Histogram of velocities of the quasars with an SNR cutoff of 2 and using the specific Hβ broad lines as described above. Removes unexplained high velocity values, but does not eliminate ± 3000 km/s peaks. Sample size is roughly 26000 objects.

issue of peaks being present at roughly ±3000 km/s.

2.3.1 Manual Review of Spectra

Because the peaks were not completely eliminated, a manual review of the spectra around ± 3000 km / s and elimination of the spectra where the data were not meaningful or suitable for study in this capacity was required. This eliminated the peaks of the histogram while preserving the count of objects. There were four groups of objects that were not suitable for study. The first group were objects with broad $H\beta$ lines of such low amplitude that they were indistinguishable from the continuum or unable to assess an accurate peak

wavelength of the line, seen in Figure 16. This was the most common error in the spectra

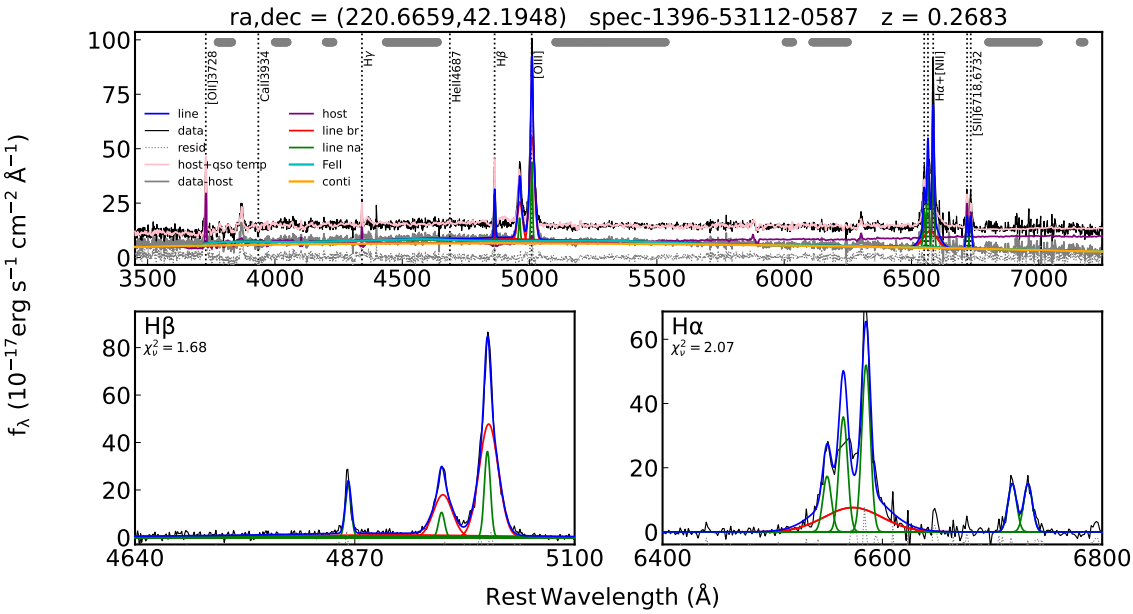

Figure 16: An example of a spectrum with fitted Hβ lines that are too weak to accurately fit the peak of the line.

that were flagged with velocities above 2800 km/s. The velocity that is associated with these spectra cannot be trusted because the broad line is so insignificant that there can be no confidence in the estimated peak wavelength.

The second most common group of error-full spectra is objects with noise that overpowers the fit of the hβ broad line, seen in Figure 17. Spectra such as this were unusable due to the inability to ascertain whether the fit matched the data well or if program were merely fitting noise as if it were data.

The third group of inaccurate spectra was less common, but was made up of spectra with poor fits due to a very strong narrow line and a very weak and low redshift broad line, seen in Figure 18. Visual inspection can show that there would be no significantly redshifted broad line.

The least common error in the high velocity spectra is a bad fit due to two disagreeing Hβ broad line fits, seen in Figure 19. This is an issue because of how the pyQSOFit

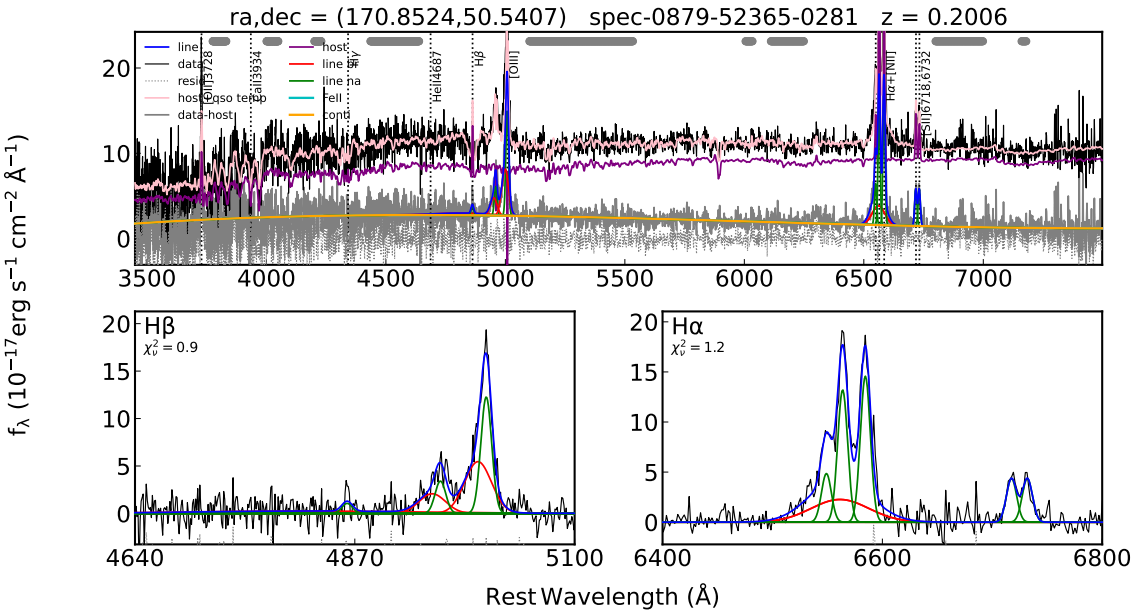

Figure 17: An example of a spectrum with fitted Hβ lines that are visibly contained within the noise of the data, meaning the fit is inaccurate.

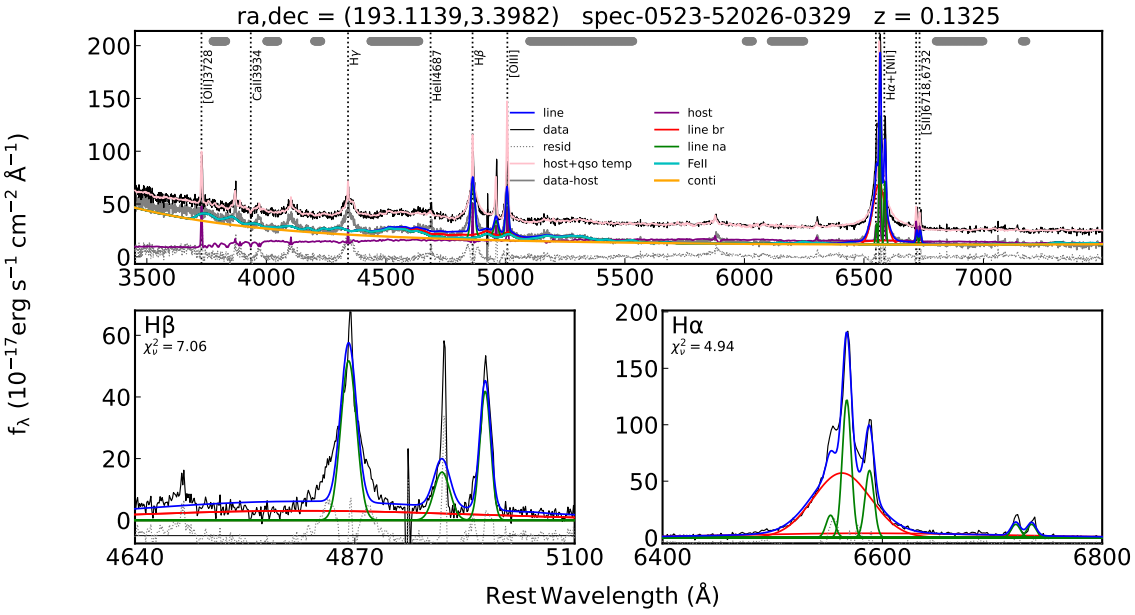

Figure 18: An example of a spectrum with fitted Hβ broad lines that are too weak in comparison to the fitted narrow Hβ to be fit well by pyQSOFit.

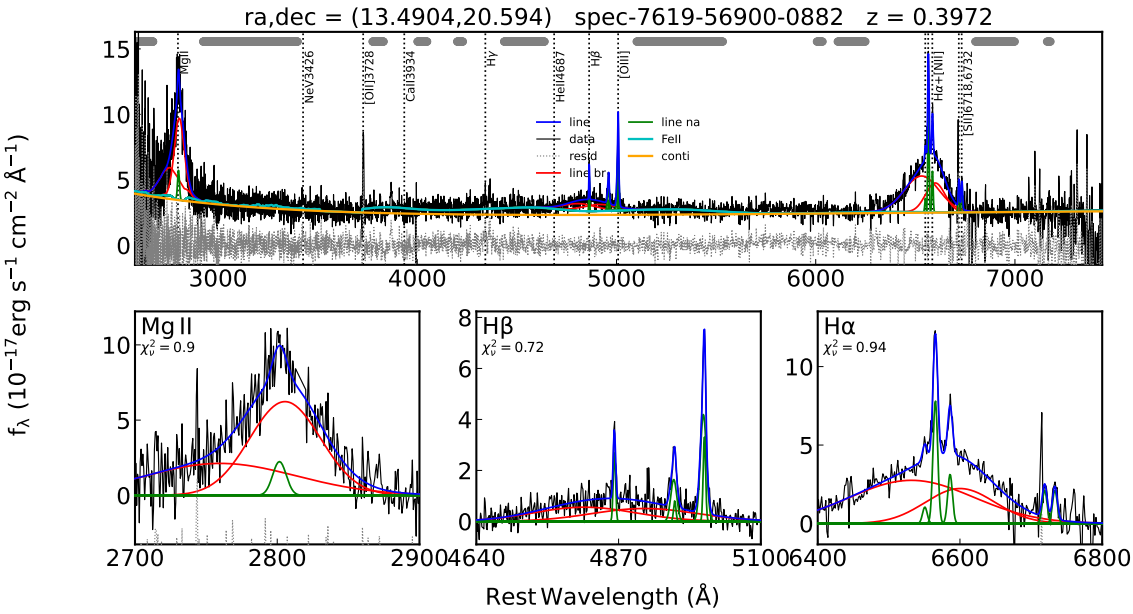

Figure 19: An example of a spectrum with two conflicting fits for broad Hβ. The data shows an Hβ broad line that does not appear redshifted, which does not agree with either of the broad line fits.

program adds all of the broad and narrow lines together for the cumulative fitted spectral line. With both specific Hβ broad lines in disagreement and with similar amplitudes, it can make a broad component with little redshift, appear as though it has a significant velocity. After the removal of these spectra with error from the dataset, the resulting histogram of all objects no longer has unexplained regions with high densities of objects, seen in Figure 20.

3 Results

This section breaks down the velocity histogram after data reduction. The purpose of discussing this histogram is to compare it with what the expected velocities of recoiling black holes. This will be done by comparing velocities with the maximum expected recoil velocity and the predicted velocity distribution from [14].

3.1 Comparison with Expected Results

The resulting histogram after manual review is seen in Figure 20. The expected results of

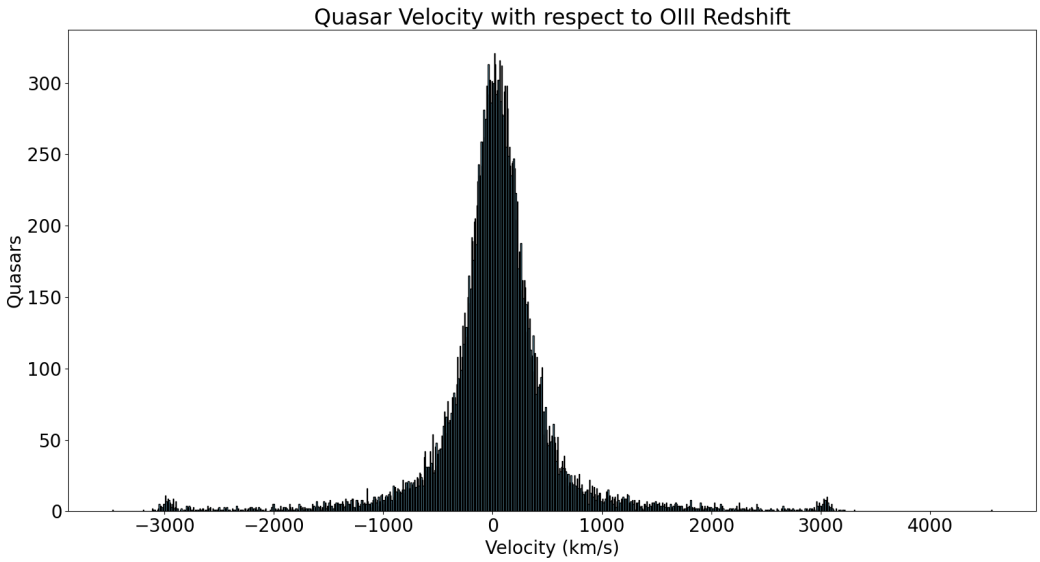

Figure 20: Histogram of velocities of the quasars after removing objects with fitting errors. This no longer has large peaks at ± 3000 km/s, but there remains some objects within this region to discuss.

the kick velocities are two-fold. First, there is the expectation that the distribution will be much more dense at lower velocities. This is a result of the probability of a random binary black hole merger having a specific kick velocity or greater goes down exponentially as that velocity increases [14], seen in Figure 21. Furthermore, these velocities come from observational data of redshift, meaning that the calculated velocity will be affected by the

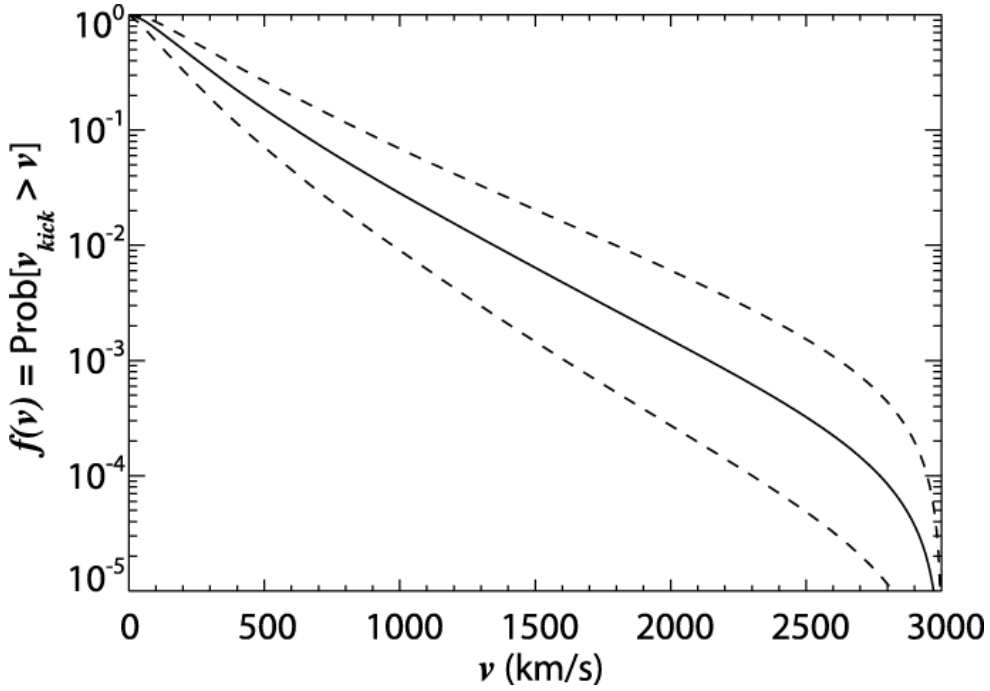

Figure 21: This is a plot of the probability of a kick velocity being greater than or equal to specific velocities from [14]. Displays an exponential drop in the probability of velocities as they increase.

orientation of the system. This is because the peculiar velocities of quasars, or any general object, are expected to be isotropic, as a result of no bias in the direction of phenomena that could result in a peculiar velocity on a universal scale. The second is that the maximum velocity of a recoiling quasar is expected to be roughly 4000 km/s resulting from the case where there are two black holes with maximal spins with counter-aligned spins [5]. Figure 20 displays both of these qualities in the data. However, there is one spectrum, the spectrum of the object 11353-58467-0222, that has a recoiling velocity of $\tilde{4}$570 km/s, which is slightly beyond what would be expected. Beyond just this one spectrum, there are other potentially incorrect spectra that exist in the high velocity, or greater than ± 2800 km/s range.

4 Analysis

This section discusses the remaining spectra in the high velocity (> 2800 km/s) region of the histogram. This is to note both the fitted spectra that are representing accurate data, but to also draw attention to potential errors in the spectra that remain. These errors take several forms, including potentially double-peaked broad lines, significant noise, and disagreeing broad lines. The reason these spectra were not eliminated in the initial data reduction is because they all have a possibility of being well fitted spectra, and as such, they are deserving of separate discussion from the poorly fitted spectra with clear errors. Furthermore, this section will discuss the velocity histogram after removing the objects, discussed in this section, that we cannot have full confidence in.

4.1 Interesting Spectra

Although the velocity distribution aligns with expectations for a large quasar sample, there are still spectra in the dataset that deserve attention. Some spectra have unique features that may raise doubts about the accuracy of their velocity, though their fit to the data is still good. This means that while the fit may be accurate, the velocity could be inaccurate for the black hole's actual motion relative to its host galaxy. These objects were kept in the dataset because they are potentially accurate, provide valuable information about the methodology used in identifying kick velocities, and their numbers are too small to significantly affect the overall velocity distribution. Objects with high velocities are notable because such velocities are rare in physical systems, requiring more scrutiny. There are also a few objects no longer in the dataset that are worth discussing.

The first object to discuss is the object in Figure 22. This object is significant for two reasons. First, this is the highest velocity object in the entire dataset according to the calculated value of velocities. Second, the velocity is actually the only velocity that exists outside of the predicted velocity range from the paper predicting velocities

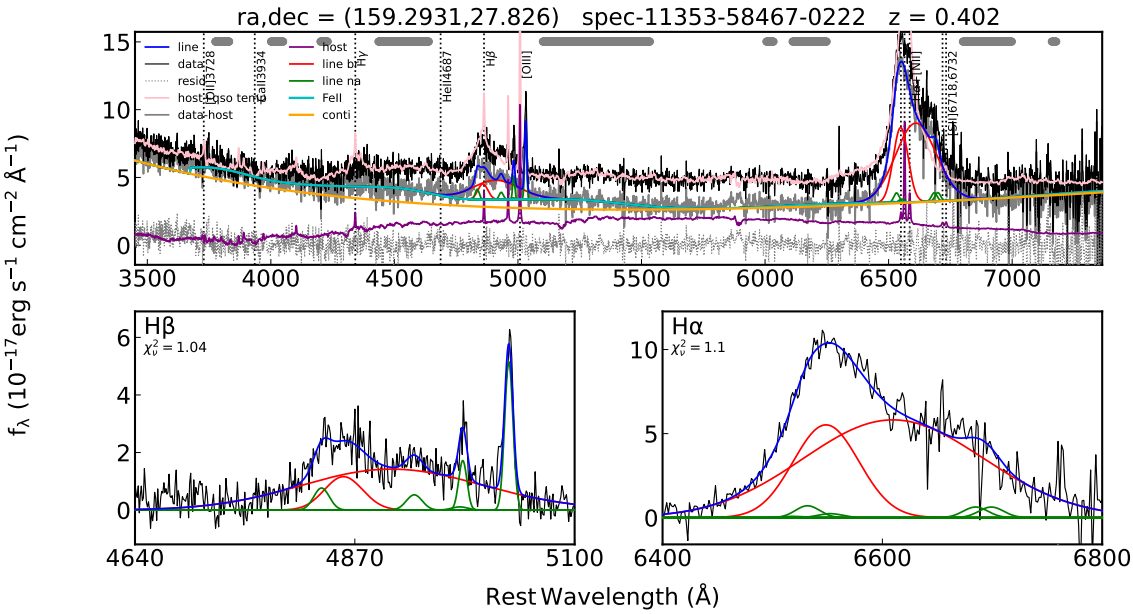

Figure 22: Fit of object 11353-58467-0222. This is the highest velocity spectrum in the sample. It appears to have a potential double peak in the Hβ broad line.

for this phenomenon [5]. There are a couple of potential reasons for this discrepancy, including a high amount of error in the fit value of the spectra, or a double peaked spectra. Double peaked spectra are spectra of objects where there appears to be two separate lines originating from where a spectral line should be.

There are multiple theories about double-peaked emission lines. One theory is the theory that the broad line double peaks have to do with a binary black hole system's rotation. In this model, the black hole with the broad line region rotating towards us would be blueshifted and the other would be redshifted. The prevailing theory seems to be that there is a rotation of the broad line region at the outer parts of the accretion disk [7]. Regardless, these motions have nothing to do with the recoiling black hole, and thus make it difficult to obtain a great deal of information about the kick velocity when there is a present double peaked Hβ line. Furthermore, the fitting program can experience some difficulty with properly fitting broad lines to these objects. The possibility of a double peaked spectrum seems most apparent in the Hβ region of wavelengths, as evidenced

by the two peaks before the [OIII] narrow lines. Figures 23 and 24 also have some characteristics that could be evidence of a double peaked spectra.

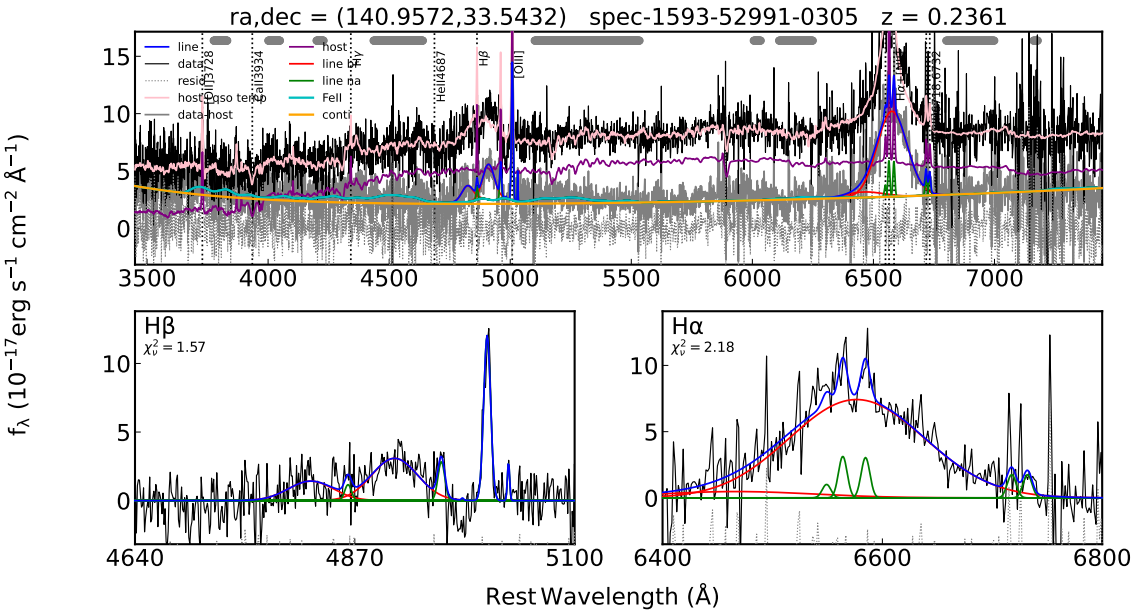

Figure 23: Fit of object 1593-52991-0305. This is another object that appears to be double-peaked in broad Hβ.

Figure 23 is another object of high velocity (3000 km/s) with two peaks in the Hβ region than that of Figure 22. The object in Figure 23 has no apparent double peak in the data of the Hα region, which suggests that there is a possibility that the double peak in the Hβ region could be more to do with the noise in the spectrum rather than an actual double peaked spectrum.

As opposed to the previous two candidates for double peaked spectra, the object in Figure 24 has a much more pronounced double peak. This is because the double peak is visible in both the data in the Hβ and Hα regions. Furthermore, the double peak seems to have the same shape in both regions, which further increases the possibility of this being the case of this spectra. Out of the three spectra, it is likely that the only accurate fits are that of Figure 24, and potentially 23 if the double peak is not noise. This prediction is gained from visually inspecting that the broad line fits are accurate to where the two

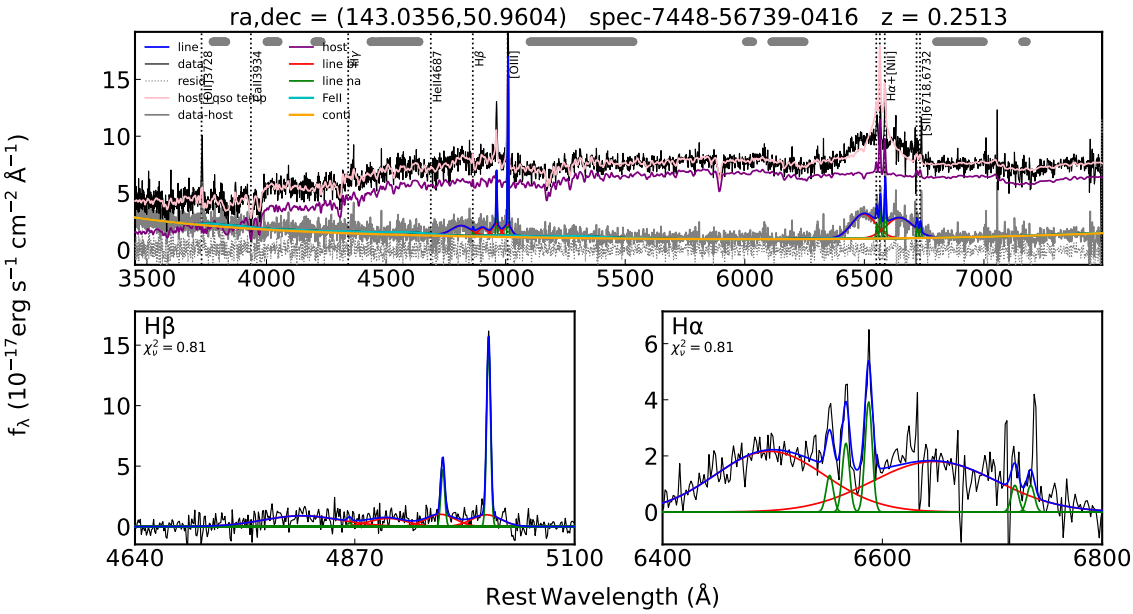

Figure 24: Fit of object 7448-56739-0416. This is the most rponounced double-peaked Hβ broad line in the sample. It also seems to have double-peaked Hα.

lines would be in Hβ in either of these two figures. If we were to interpret the velocity of these spectral lines as orbits around the black hole, as it seems visually plausible for Figure 24, this would be a velocity of roughly 3000 km/s both towards and away from us. These velocities are well under what could possibly be orbiting a supermassive black hole, meaning that this is a possibility.

It is important to note potential errors in spectra fitting, especially for the high velocity (3000+ km/s) spectra, but it is equally important to identify high velocity spectra that do not seem to be in error. One such spectrum is that in Figure 25. This object has a low amplitude of its broad line, but the amplitude is still large enough to overcome the noise that is existent in the spectra. The redshift of the Hβ broad line is also apparent in the Hα broad line. The combination of these two pieces of information result in a high confidence that this object is genuinely a high velocity black hole with respect to its host galaxy.

Another interesting spectrum from the high velocity group is the spectrum represented in Figure 26. This spectrum shows broad lines with different characteristics in each of the

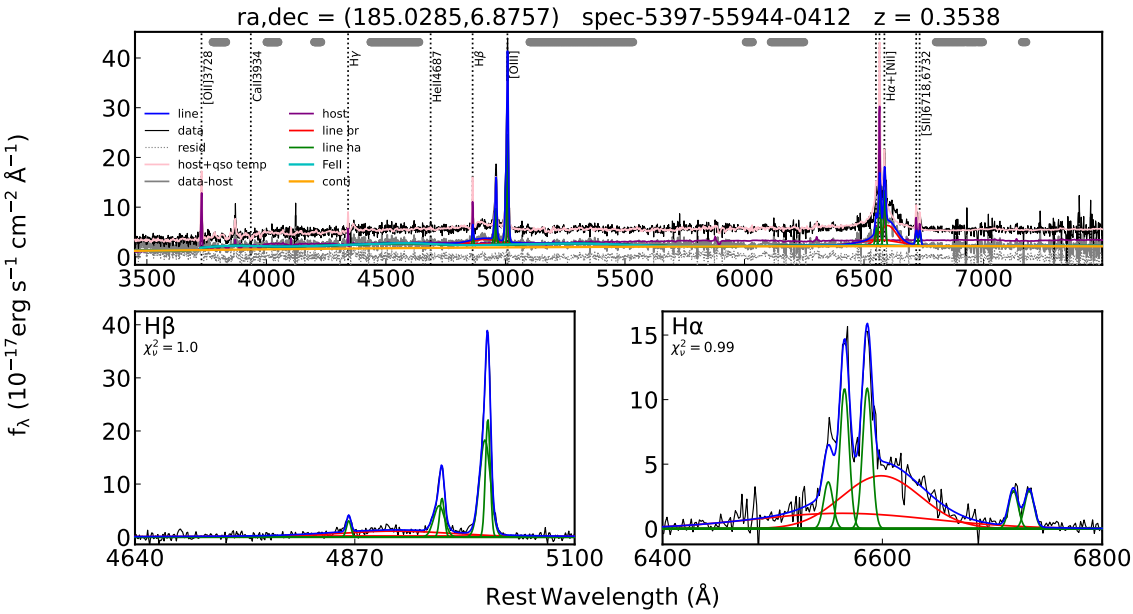

Figure 25: Fit of object 5397-55944-0412. An example of a well fit spectrum with a high velocity broad line region with respect to its narrow line region.

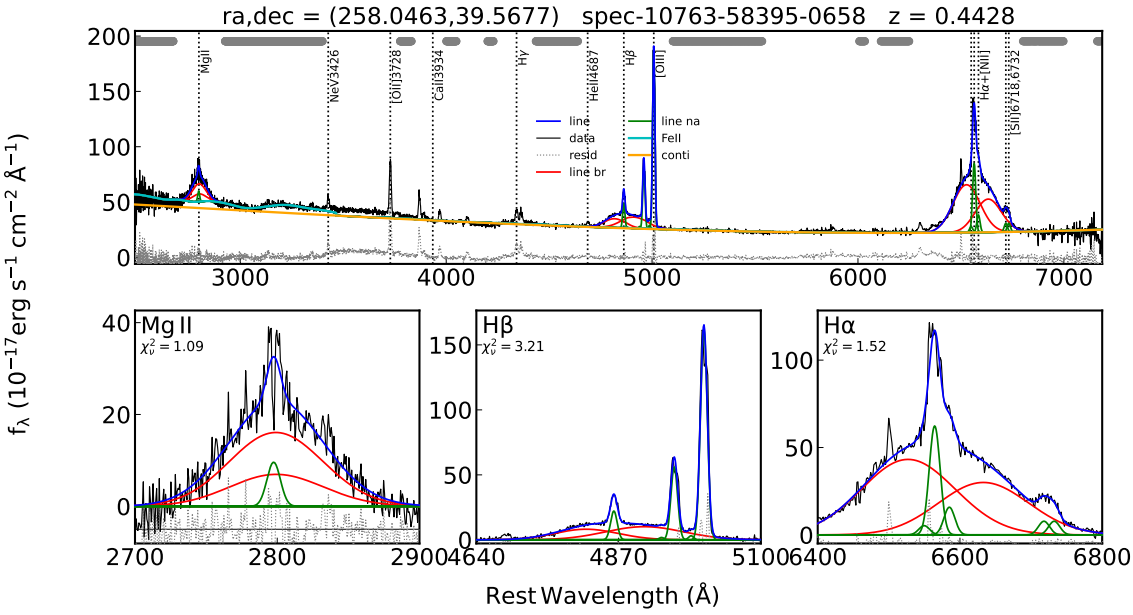

Figure 26: Fit of object 10763-58395-0658. Displays Hβ broad line fits that disagree, but the dominant broad line could give an accurate peak wavelength.

three regions, MgII, Hβ, and Hα. The fit broad line in Hβ seems to be slightly redshifted, while the broad line in Hα seems to be slightly blueshifted, and the broad line in MgII seems to be centered about the narrow line. This suggests that there is an issue with fitting in this spectrum despite the low noise and relatively significant amplitude of the broad lines. As a result, this is another high velocity spectrum that does not display characteristics that instill large amounts of confidence in the calculated velocity.

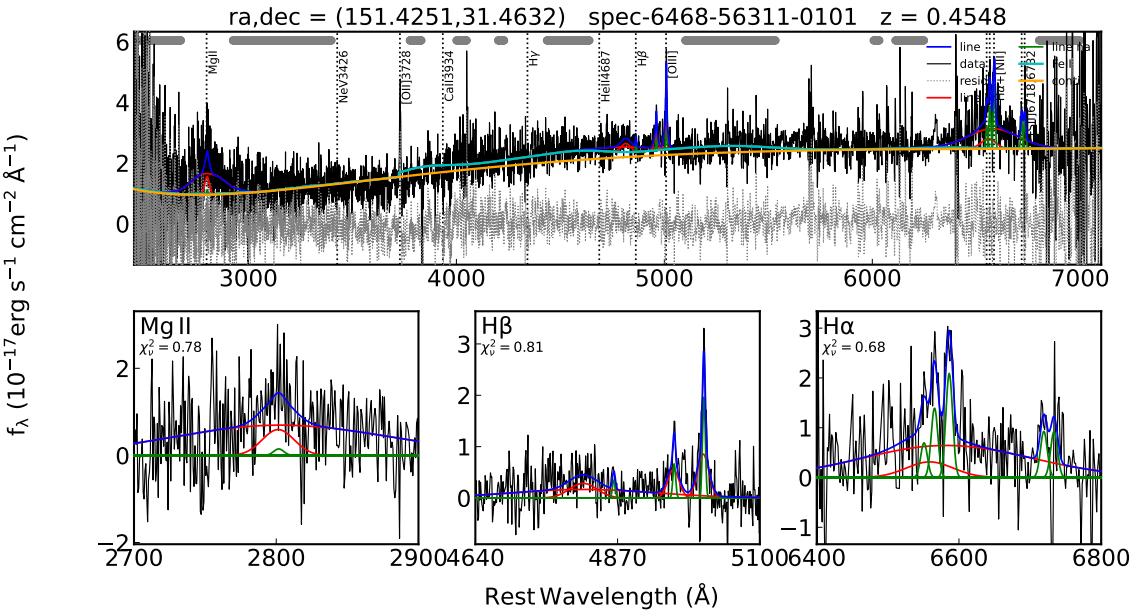

Figure 27: Fit of object 6468-56311-0101. Displays a noisy fit that has a reasonable fit for Hβ based on the shape of the data.

The spectrum shown in Figure 27 is an example of a quite noisy spectrum with data significant enough to consider the possibility of the velocity being accurate for estimating the order of magnitude. This is because, there is a distinct shape in the data that can be seen even through the noise of the spectrum and because the velocity resulting from the peak of this feature is significant enough that it is distinguishable from that of a low velocity object.

The object in Figure 28 is not included in the final velocity distribution because it does not seem to have a distinguishable broad line peak due to noise, it has a very low amplitude

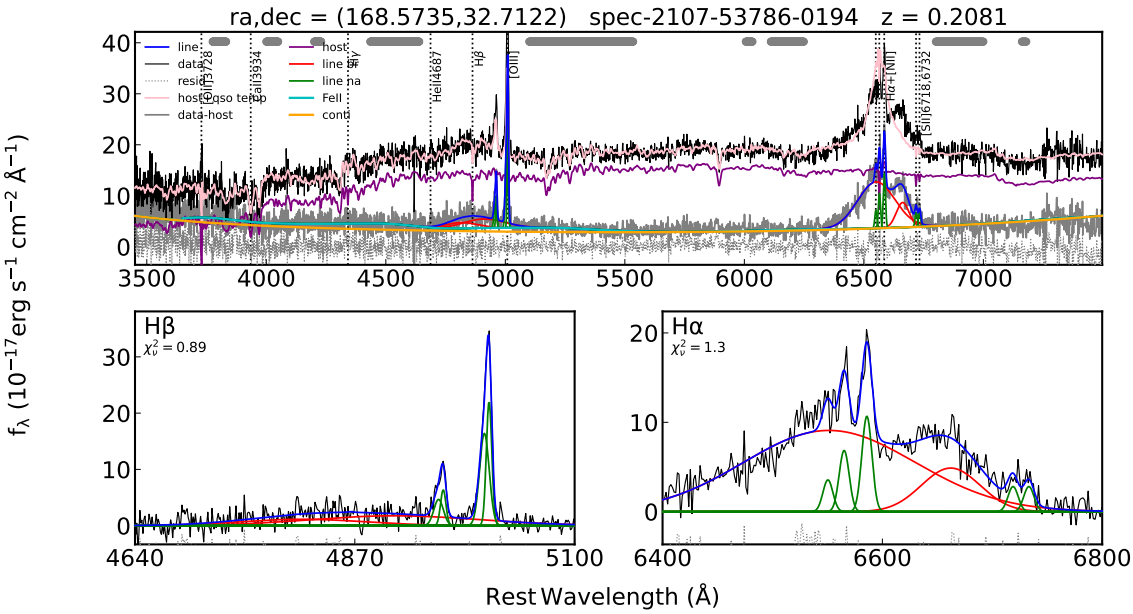

Figure 28: Fit of object 2107-53786-0194. This object is not in the histogram because it has a combination of the potential errors discussed. These errors include potential double-peaked lines, significant noise, and disagreeing broad lines.

of broad line, and it seems to potentially be double peaked. However, this spectrum is a great example for discussing the issues of fitting narrow Hβ line. This is important because the Hβ narrow line could be a candidate for being the stand-in of the narrow line region's redshift. This figure displays just one example of an Hα narrow line that is far too insignificant to be properly fit. If the survey were to use the Hα line's redshift as the narrow line region's redshift, there would be a significant decrease in the available objects for study due to the inability to properly fit narrow Hα, but this is not a problem for the [OIII] spectral line.

4.2 Improved Velocity Plot

As shown through several examples of spectra with the potential to be improperly fit and have an incorrect velocity, it is possible to remove some of these spectra from the plot and compare. The reason that this was not done initially, is that these objects have the

potential for significant data and the distribution is negligibly affected, seen in Figure 29.

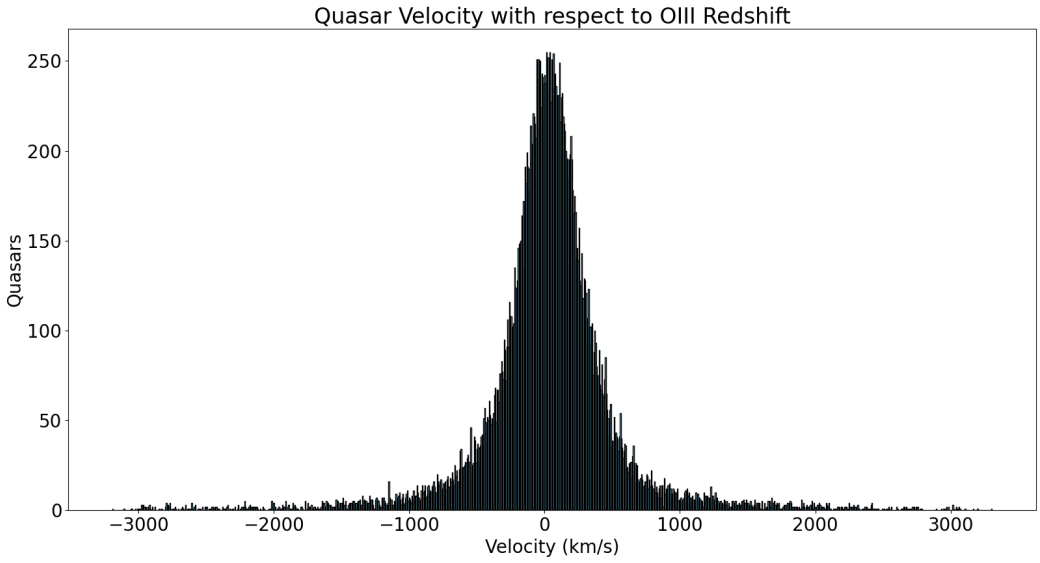

Figure 29: Histogram of kick velocities after removing low confidence spectra. This histogram no longer contains any notable unexplained features.

5 Next Steps

This section is meant to discuss what can be done in the future to improve and further evidence the results of this project. This includes the comparison of the velocities calculated in this project using [OIII] as the broad line with velocities calculated using MgII as the broad line. This comparison reveals that we cannot be confident in the broad line region's velocity representing the velocity of the recoiling black hole, but also that there is a trend in the data that suggests there may be a relationship between the broad line region's velocity and the black hole's. This comparison could be further explored and improved which is why it is in this section. Lastly, this section discusses methods that were not completed in this project, but would be useful for further developing this project.

5.1 Cross Examination with Broad Lines

The next step for this project would be to compare the results of using the Hβ broad line to calculate velocities with the results of using a different broad line as the stand-in for the redshift of the broad line region. This would allow for a cross examination of results between the broad lines redshifts with respect to the narrow line region. The expected result of this comparison would be that the redshift is roughly the same for all objects examined if the broad line region is moving with the black hole as was assumed in this survey.

Comparing the resulting velocities of the broad line region with respect to the narrow line region using separate broad lines works a sanity check of the results obtained in this project. This method was used in the paper Bonning et al. [3]. In this paper, the respective redshifts were plotted against each other in a graph for the objects, but the trend was not linear, suggesting that the velocity may not be entirely accurate, or due to the recoiling of black holes after a merger. Using the data from this survey, some of the objects contain fits for the MgII broad line. This means that it is possible to make a comparison of the

velocities obtained from the two broad lines for some of the objects in this paper.

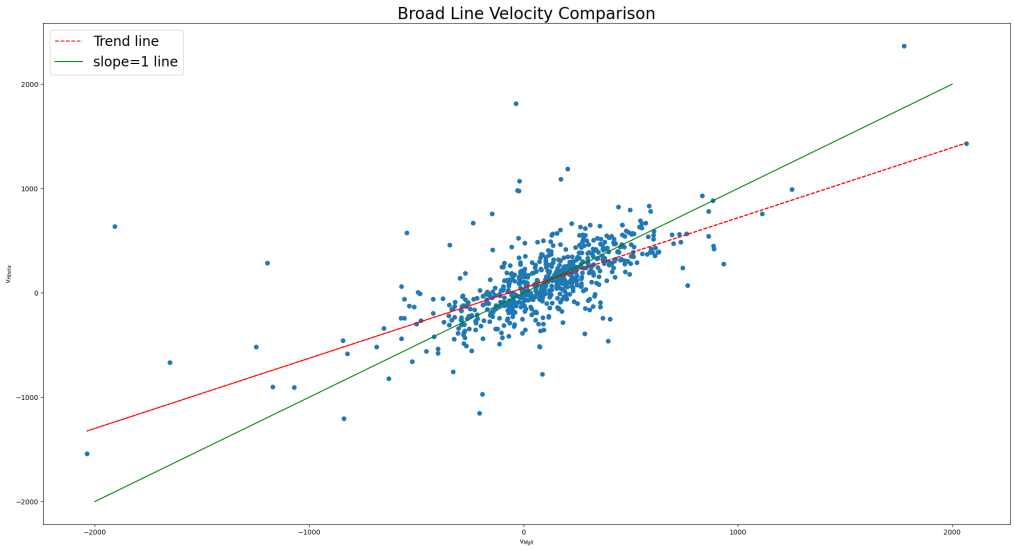

Figure 30: A comparison of the velocities of objects' broad line regions with respect to their narrow line regions calculated with the MgII broad line and Hβ broad line. The trend line, represented in red, has a slope of 0.67273. There is a line of slope 1, represented in green.

The plot in Figure 30 gives a demonstration of what this could be for this dataset. However, there were criteria required for the objects used in the making of this plot that may affect the data contained. First, and most importantly, the MgII broad line is only contained in the higher redshift spectra of this dataset and the minimum redshift of the objects in this plot is z=0.4. Secondly, the MgII broad line is a doublet, meaning that it has two lines at roughly 2800 Å, so the rest wavelength was assumed to be the 2798 Å line. Lastly, the amplitude of the broad line that is used in calculating the velocity was required to be twice the amplitude of the secondary fit broad line, in order to not allow for double peaked lines in this comparison. The slope of the trend line in the plot in Figure 30 is 0.67273 which does not imply the expected linear relationship, but the central portion of the data in the figure appears more linear, suggesting that this trend line might be skewed by outliers. This trend is most easily seen when comparing with the green line that has

a slope of 1. However, the values of velocity should be the same if the broad line region is truly recoiling with the black hole, which suggests that there is more to uncover. In a future implementation of this method, it would be important to do more analysis in what type of objects may be outliers in a plot such as this, as well as to find or create a new fitting program that can fit broad lines more accurately and fit other broad lines for comparison such as NeV or NeIII.

5.2 Model Velocity Histogram

Another future continuation of this project would be to model what the histogram of velocity should theoretically look like. This would be done by taking the probability distribution for kick velocities, and randomly generating a large number of objects with random spatial orientation and velocities based on this distribution. Then converting these velocities to what an observer would witness from one angle of observation, similar to the survey done in this project.

6 Conclusion

The purpose of this project is to search for post-merger recoiling black holes. This was done by using spectra to find the velocity of the broad line region with respect to the narrow line region by using the emission lines Hβ and [OIII] to represent the broad and narrow line regions respectively. This velocity is assumed to be representative of the central black hole's velocity with respect to its host galaxy because the broad line region is orders of magnitude closer to the central black hole and gravitationally affected by it, while the narrow line region is not. This section summarizes how the velocities were obtained and how the results were analyzed.

This project has taken the spectra for all quasars within a redshift of z = 0.55 from Earth, using SDSS, and fitted them with emission lines using the program pyQSOFit. After fitting all of the spectra available, the peak wavelength values of the broad Hβ line and the [OIII] narrow line were used to estimate the velocity of the broad line region with respect to the narrow line region. This was done by calculating the redshift of the broad line region with respect to the narrow line region and then using the Doppler shift formula to convert into a velocity.

With velocities of all of the objects within the sample, a histogram of the velocities was made. Initially, this plot contained results that would conflict with our understanding of what would be possible due to kick velocities from black hole mergers. However, through identification of errors in the techniques used to fit the data and find velocities and careful manual review of the spectra represented in the plot, the histogram was reduced to a plot that is entirely capable of being explained by our current understanding of kick velocities in post-merger black holes.

Using the MgII broad line to find kick velocities of the objects and compare with the velocities calculated with the Hβ broad line proved to be a worthwhile endeavor. This revealed that, while a trend suggesting that the velocities are due to post-merger kicks is possible, there is likely another variable involved in explaining these velocities. This

would require further testing with other broad lines as well as a more accurate method for fitting the spectral lines within the data for these objects. Lastly, in order to confirm that the histogram fully conforms to the current understanding of post-merger kick velocities, it would be important to test what a theoretical histogram would appear as given random orientations of velocities, the currently accepted distribution of kick velocities, and a single observation angle.

7 Appendix

Appendix A: URL Sorting Code

```
import numpy as np

data=np.loadtxt('QSOSpectra.txt',skiprows=2,dtype=str)
plate=data[:,0]
platenum=plate.astype(dtype=float)
dr12=data[:,4]
eBoss=data[:,5]
Segue=data[:,7]
urls=np.array(() ,dtype=str)
i=0
while (i<np.size(dr12)):
    if platenum[i]>3519:
        x=eBoss[i]
        urls=np.append(urls,x)
        i=i+1
    elif platenum[i]>3006 and platenum[i]<3520:
        z=Segue[i]
        urls=np.append(urls,z)
        i=i+1
    elif platenum[i]==2957 or platenum[i]==2640 or platenum[i]==2962 or platenum[i]==3000 or platenum[i]==3002 or platenum[i
        ]==3003 or platenum[i]==3005:
        z=Segue[i]
        urls=np.append(urls,z)
        i=i+1
    else:
       y=dr12[i]
       urls=np.append(urls,y)
       i=i+1


np.savetxt('QSO_URLs.txt', urls, delimiter='', fmt='%s')
```

Appendix B: Fitting Program

Code adapted from Guo et al, 2018. https://github.com/legolason/PyQSOFit/blob/master/example/example.ipynb

Initialization:

```
import os
os.chdir('C:\\Users\\robos\\Downloads\\Anaconda Stuff\\Honors thesis programs')

# matplotlib inline
import glob, os, sys, timeit
import matplotlib
import numpy as np

sys.path.append('../')
from pyqsofit.PyQSOFit import QSOFit
from astropy.io import fits
from astropy.table import Table
import matplotlib.pyplot as plt
import warnings

warnings.filterwarnings("ignore")

QSOFit.set_mpl_style()

# Show the versions so we know what works
import astropy
import lmfit
import pyqsofit

print(astropy.__version__)
print(lmfit.__version__)
print(pyqsofit.__version__)

import emcee  # optional, for MCMC

path_ex = '.' #os.path.join(pyqsofit.__path__[0], '..', 'example')

# create a header
hdr0 = fits.Header()
hdr0['Author'] = 'Nathan Rauch'
primary_hdu = fits.PrimaryHDU(header=hdr0)

"""
In this table, we specify the priors / initial conditions and boundaries for the line fitting parameters.
"""

line_priors = np.rec.array([
    (6564.61, 'Ha', 6400, 6800, 'Ha_br', 2, 0.0, 0.0, 1e10, 5e-3, 0.004, 0.05, 0.015, 0, 0, 0, 0.05, 1),
    (6564.61, 'Ha', 6400, 6800, 'Ha_na', 1, 0.0, 0.0, 1e10, 1e-3, 5e-4, 0.00169, 0.01, 1, 1, 0, 0.002, 1),
    (6549.85, 'Ha', 6400, 6800, 'NII6549', 1, 0.0, 0.0, 1e10, 1e-3, 2.3e-4, 0.00169, 5e-3, 1, 1, 1, 0.001, 1),
    (6585.28, 'Ha', 6400, 6800, 'NII6585', 1, 0.0, 0.0, 1e10, 1e-3, 2.3e-4, 0.00169, 5e-3, 1, 1, 1, 0.003, 1),
    (6718.29, 'Ha', 6400, 6800, 'SII6718', 1, 0.0, 0.0, 1e10, 1e-3, 2.3e-4, 0.00169, 5e-3, 1, 1, 2, 0.001, 1),
```

```
    (6732.67, 'Ha', 6400, 6800, 'SII6732', 1, 0.0, 0.0, 1e10, 1e-3, 2.3e-4, 0.00169, 5e-3, 1, 1, 2, 0.001, 1),

    (4862.68, 'Hb', 4640, 5100, 'Hb_br', 2, 0.0, 0.0, 1e10, 5e-3, 0.004, 0.05, 0.01, 0, 0, 0, 0.01, 1),
    (4862.68, 'Hb', 4640, 5100, 'Hb_na', 1, 0.0, 0.0, 1e10, 1e-3, 2.3e-4, 0.00169, 0.01, 1, 1, 0, 0.002, 1),
    (4960.30, 'Hb', 4640, 5100, 'OIII4959c', 1, 0.0, 0.0, 1e10, 1e-3, 2.3e-4, 0.00169, 0.01, 1, 1, 0, 0.002, 1),
    (5008.24, 'Hb', 4640, 5100, 'OIII5007c', 1, 0.0, 0.0, 1e10, 1e-3, 2.3e-4, 0.00169, 0.01, 1, 1, 0, 0.004, 1),
    (4960.30, 'Hb', 4640, 5100, 'OIII4959w',   1, 0.0, 0.0, 1e10, 3e-3, 2.3e-4, 0.004,  0.01,  2, 2, 0, 0.001, 1),
    (5008.24, 'Hb', 4640, 5100, 'OIII5007w',   1, 0.0, 0.0, 1e10, 3e-3, 2.3e-4, 0.004,  0.01,  2, 2, 0, 0.002, 1),
    #(4687.02, 'Hb', 4640, 5100, 'HeII4687_br', 1, 0.0, 0.0, 1e10, 5e-3, 0.004,  0.05,   0.005, 0, 0, 0, 0.001, 1),
    #(4687.02, 'Hb', 4640, 5100, 'HeII4687_na', 1, 0.0, 0.0, 1e10, 1e-3, 2.3e-4, 0.00169, 0.005, 1, 1, 0, 0.001, 1),

    #(3934.78, 'CaII', 3900, 3960, 'CaII3934' , 2, 0.0, 0.0, 1e10, 1e-3, 3.333e-4, 0.00169, 0.01, 99, 0, 0, -0.001, 1),

    #(3728.48, 'OII', 3650, 3800, 'OII3728', 1, 0.0, 0.0, 1e10, 1e-3, 3.333e-4, 0.00169, 0.01, 1, 1, 0, 0.001, 1),

    #(3426.84, 'NeV', 3380, 3480, 'NeV3426',    1, 0.0, 0.0, 1e10, 1e-3, 3.333e-4, 0.00169, 0.01, 0, 0, 0, 0.001, 1),
    #(3426.84, 'NeV', 3380, 3480, 'NeV3426_br', 1, 0.0, 0.0, 1e10, 5e-3, 0.0025,   0.02,   0.01, 0, 0, 0, 0.001, 1),

    (2798.75, 'MgII', 2700, 2900, 'MgII_br', 2, 0.0, 0.0, 1e10, 5e-3, 0.004, 0.05, 0.015, 0, 0, 0, 0.05, 1),
    (2798.75, 'MgII', 2700, 2900, 'MgII_na', 1, 0.0, 0.0, 1e10, 1e-3, 5e-4, 0.00169, 0.01, 1, 1, 0, 0.002, 1),

    (1908.73, 'CIII', 1700, 1970, 'CIII_br', 2, 0.0, 0.0, 1e10, 5e-3, 0.004, 0.05, 0.015, 99, 0, 0, 0.01, 1),
    #(1908.73, 'CIII', 1700, 1970, 'CIII_na',   1, 0.0, 0.0, 1e10, 1e-3, 5e-4,  0.00169, 0.01,  1, 1, 0, 0.002, 1),
    #(1892.03, 'CIII', 1700, 1970, 'SiIII1892', 1, 0.0, 0.0, 1e10, 2e-3, 0.001, 0.015,  0.003, 1, 1, 0, 0.005, 1),
    #(1857.40, 'CIII', 1700, 1970, 'AlIII1857', 1, 0.0, 0.0, 1e10, 2e-3, 0.001, 0.015,  0.003, 1, 1, 0, 0.005, 1),
    #(1816.98, 'CIII', 1700, 1970, 'SiII1816',  1, 0.0, 0.0, 1e10, 2e-3, 0.001, 0.015,  0.01,  1, 1, 0, 0.0002, 1),
    #(1786.7,  'CIII', 1700, 1970, 'FeII1787',  1, 0.0, 0.0, 1e10, 2e-3, 0.001, 0.015,  0.01,  1, 1, 0, 0.0002, 1),
    #(1750.26, 'CIII', 1700, 1970, 'NIII1750',  1, 0.0, 0.0, 1e10, 2e-3, 0.001, 0.015,  0.01,  1, 1, 0, 0.001, 1),
    #(1718.55, 'CIII', 1700, 1900, 'NIV1718',   1, 0.0, 0.0, 1e10, 2e-3, 0.001, 0.015,  0.01,  1, 1, 0, 0.001, 1),

    (1549.06, 'CIV', 1500, 1700, 'CIV_br', 2, 0.0, 0.0, 1e10, 5e-3, 0.004, 0.05, 0.015, 0, 0, 0, 0.05, 1),
    # (1549.06, 'CIV', 1500, 1700, 'CIV_na', 1, 0.0, 0.0, 1e10, 1e-3, 5e-4, 0.00169, 0.01, 1, 1, 0, 0.002, 1),
    #(1640.42, 'CIV', 1500, 1700, 'HeII1640',    1, 0.0, 0.0, 1e10, 1e-3, 5e-4,   0.00169, 0.008, 1, 1, 0, 0.002, 1),
    #(1663.48, 'CIV', 1500, 1700, 'OIII1663',    1, 0.0, 0.0, 1e10, 1e-3, 5e-4,   0.00169, 0.008, 1, 1, 0, 0.002, 1),
    #(1640.42, 'CIV', 1500, 1700, 'HeII1640_br', 1, 0.0, 0.0, 1e10, 5e-3, 0.0025, 0.02,   0.008, 1, 1, 0, 0.002, 1),
    #(1663.48, 'CIV', 1500, 1700, 'OIII1663_br', 1, 0.0, 0.0, 1e10, 5e-3, 0.0025, 0.02,   0.008, 1, 1, 0, 0.002, 1),

    #(1402.06, 'SiIV', 1290, 1450, 'SiIV_OIV1', 1, 0.0, 0.0, 1e10, 5e-3, 0.002, 0.05,  0.015, 1, 1, 0, 0.05, 1),
    #(1396.76, 'SiIV', 1290, 1450, 'SiIV_OIV2', 1, 0.0, 0.0, 1e10, 5e-3, 0.002, 0.05,  0.015, 1, 1, 0, 0.05, 1),
    #(1335.30, 'SiIV', 1290, 1450, 'CII1335',   1, 0.0, 0.0, 1e10, 2e-3, 0.001, 0.015, 0.01,  1, 1, 0, 0.001, 1),
    #(1304.35, 'SiIV', 1290, 1450, 'OI1304',    1, 0.0, 0.0, 1e10, 2e-3, 0.001, 0.015, 0.01,  1, 1, 0, 0.001, 1),

    (1215.67, 'Lya', 1150, 1290, 'Lya_br', 3, 0.0, 0.0, 1e10, 5e-3, 0.002, 0.05, 0.02, 0, 0, 0, 0.05, 1),
    (1240.14, 'Lya', 1150, 1290, 'NV1240', 1, 0.0, 0.0, 1e10, 2e-3, 0.001, 0.01, 0.005, 0, 0, 0, 0.002, 1),
    # (1215.67, 'Lya', 1150, 1290, 'Lya_na', 1, 0.0, 0.0, 1e10, 1e-3, 5e-4, 0.00169, 0.01, 0, 0, 0, 0.002, 1),
    ],

    formats='float32,    a20,  float32, float32,      a20,  int32, float32, float32, float32, float32, float32, float32,
          int32,  int32,  int32, float32, int32',
    names=' lambda, compname,   minwav, maxwav, linename, ngauss, inisca, minsca, maxsca, inisig, minsig, maxsig,    voff,
         vindex, windex, findex,  fvalue,  vary')

# Header
hdr1 = fits.Header()
hdr1['lambda'] = 'Vacuum Wavelength in Ang'
hdr1['minwav'] = 'Lower complex fitting wavelength range'
```

```
hdr1['maxwav'] = 'Upper complex fitting wavelength range'
hdr1['ngauss'] = 'Number of Gaussians for the line'

# Can be set to negative for absorption lines if you want
hdr1['inisca'] = 'Initial guess of line scale [flux]'
hdr1['minsca'] = 'Lower range of line scale [flux]'
hdr1['maxsca'] = 'Upper range of line scale [flux]'

hdr1['inisig'] = 'Initial guess of linesigma [lnlambda]'
hdr1['minsig'] = 'Lower range of line sigma [lnlambda]'
hdr1['maxsig'] = 'Upper range of line sigma [lnlambda]'

hdr1['voff  '] = 'Limits on velocity offset from the central wavelength [lnlambda]'
hdr1['vindex'] = 'Entries w/ same NONZERO vindex constrained to have same velocity'
hdr1['windex'] = 'Entries w/ same NONZERO windex constrained to have same width'
hdr1['findex'] = 'Entries w/ same NONZERO findex have constrained flux ratios'
hdr1['fvalue'] = 'Relative scale factor for entries w/ same findex'

hdr1['vary'] = 'Whether or not to vary the parameter (set to 0 to fix the line parameter to initial values)'

# Save line info
hdu1 = fits.BinTableHDU(data=line_priors, header=hdr1, name='line_priors')

"""
In this table, we specify the windows and priors / initial conditions and boundaries for the continuum fitting parameters.
"""

conti_windows = np.rec.array([
    (1150., 1170.),
    (1275., 1290.),
    (1350., 1360.),
    (1445., 1465.),
    (1690., 1705.),
    (1770., 1810.),
    (1970., 2400.),
    (2480., 2675.),
    (2925., 3400.),
    (3775., 3832.),
    (4000., 4050.),
    (4200., 4230.),
    (4435., 4640.),
    (5100., 5535.),
    (6005., 6035.),
    (6110., 6250.),
    (6800., 7000.),
    (7160., 7180.),
    (7500., 7800.),
    (8050., 8150.), # Continuum fitting windows (to avoid emission line, etc.)  [AA]
    ],
    formats = 'float32, float32',
    names =  'min, max')

hdu2 = fits.BinTableHDU(data=conti_windows, name='conti_windows')

conti_priors = np.rec.array([
    ('Fe_uv_norm', 0.0, 0.0, 1e10, 1), # Normalization of the MgII Fe template [flux]
```

```
    ('Fe_uv_FWHM', 3000, 1200, 18000, 1), # FWHM of the MgII Fe template [AA]
    ('Fe_uv_shift', 0.0, -0.01, 0.01, 1), # Wavelength shift of the MgII Fe template [lnlambda]
    ('Fe_op_norm', 0.0, 0.0, 1e10, 1), # Normalization of the Hbeta/Halpha Fe template [flux]
    ('Fe_op_FWHM', 3000, 1200, 18000, 1), # FWHM of the Hbeta/Halpha Fe template [AA]
    ('Fe_op_shift', 0.0, -0.01, 0.01, 1), # Wavelength shift of the Hbeta/Halpha Fe template [lnlambda]
    ('PL_norm', 1.0, 0.0, 1e10, 1), # Normalization of the power-law (PL) continuum f_lambda = (lambda/3000)^-alpha
    ('PL_slope', -1.5, -5.0, 3.0, 1), # Slope of the power-law (PL) continuum
    ('Blamer_norm', 0.0, 0.0, 1e10, 1), # Normalization of the Balmer continuum at < 3646 AA [flux] (Dietrich et al. 2002)
    ('Balmer_Te', 15000, 10000, 50000, 1), # Te of the Balmer continuum at < 3646 AA [K?]
    ('Balmer_Tau', 0.5, 0.1, 2.0, 1), # Tau of the Balmer continuum at < 3646 AA
    ('conti_a_0', 0.0, None, None, 1), # 1st coefficient of the polynomial continuum
    ('conti_a_1', 0.0, None, None, 1), # 2nd coefficient of the polynomial continuum
    ('conti_a_2', 0.0, None, None, 1), # 3rd coefficient of the polynomial continuum
    # Note: The min/max bounds on the conti_a_0 coefficients are ignored by the code,
    # so they can be determined automatically for numerical stability.
    ],

    formats = 'a20, float32, float32, float32, int32',
    names = 'parname, initial, min, max, vary')

hdr3 = fits.Header()
hdr3['ini'] = 'Initial guess of line scale [flux]'
hdr3['min'] = 'FWHM of the MgII Fe template'
hdr3['max'] = 'Wavelength shift of the MgII Fe template'

hdr3['vary'] = 'Whether or not to vary the parameter (set to 0 to fix the continuum parameter to initial values)'


hdu3 = fits.BinTableHDU(data=conti_priors, header=hdr3, name='conti_priors')

"""
In this table, we allow user to customized some key parameters in our result measurements.
"""

measure_info = Table(
    [
        [[1350, 1450, 3000, 4200, 5100]],
        [[
            # [2240, 2650],
            [4435, 4685],
        ]]
    ],
    names=([
        'cont_loc',
        'Fe_flux_range'
    ]),
    dtype=([
        'float32',
        'float32'
    ])
)
hdr4 = fits.Header()
hdr4['cont_loc'] = 'The wavelength of continuum luminosity in results'
hdr4['Fe_flux_range'] = 'Fe emission wavelength range calculated in results'

hdu4 = fits.BinTableHDU(data=measure_info, header=hdr4, name='measure_info')
```

```
hdu_list = fits.HDUList([primary_hdu, hdu1, hdu2, hdu3, hdu4])
hdu_list.writeto(os.path.join(path_ex, 'qsopar.fits'), overwrite=True)

Table(line_priors)
```

Fitting:

```
import os
import glob, os, sys, timeit
import matplotlib
import numpy as np

sys.path.append('../')
from pyqsofit.PyQSOFit import QSOFit
from astropy.io import fits
from astropy.table import Table
import matplotlib.pyplot as plt
import warnings

warnings.filterwarnings("ignore")

QSOFit.set_mpl_style()

# Show the versions so we know what works
import astropy
import lmfit
import pyqsofit

print(astropy.__version__)
print(lmfit.__version__)
print(pyqsofit.__version__)

import emcee  # optional, for MCMC

print(emcee.__version__)

print(pyqsofit.__path__)

path_ex = '.' #os.path.join(pyqsofit.__path__[0], '..', 'example')



# Assign directory
directory = sys.argv[1]

# Iterate over files in directory
for name in os.listdir(directory):
    try:
        # Open file
        with open(os.path.join(directory, name)) as f:
            print(f"{name}")
            vname= name.rstrip('.fits')
            os.chdir('C:\\Users\\robos\\Downloads\\Anaconda Stuff\\Honors thesis programs')
            path_out = os.path.join('C:\\Users','robos','Honors Thesis Stuff','Graphing folder (initial testing and stuff)','Test
                staging folder')
```

```
# Requried
data = fits.open(os.path.join(path_ex,directory,name))
lam = 10 ** data[1].data['loglam']  # OBS wavelength [A]
flux = data[1].data['flux']  # OBS flux [erg/s/cm^2/A]
err = 1 / np.sqrt(data[1].data['ivar'])  # 1 sigma error
z = data[2].data['z'][0]  # Redshift
print(z)
data.info()
# Optional
ra = data[0].header['plug_ra']  # RA
dec = data[0].header['plug_dec']  # DEC
plateid = data[0].header['plateid']  # SDSS plate ID
mjd = data[0].header['mjd']  # SDSS MJD
fiberid = data[0].header['fiberid']  # SDSS fiber ID
# Prepare data
q_mle = QSOFit(lam, flux, err, z, ra=ra, dec=dec, plateid=plateid, mjd=mjd, fiberid=fiberid, path=path_ex)

# Double check the installation path with the PCA / Fe template files
# print('install path:', q_mle.install_path)

# Change it if you installed them somewhere else
#q_mle.install_path = '...'

start = timeit.default_timer()
# Do the fitting

q_mle.Fit(name=vname,  # customize the name of given targets. Default: plate-mjd-fiber
          # prepocessing parameters
          nsmooth=1,  # do n-pixel smoothing to the raw input flux and err spectra
          and_mask=False,  # delete the and masked pixels
          or_mask=False,  # delete the or masked pixels
          reject_badpix=True,  # reject 10 most possible outliers by the test of pointDistGESD
          deredden=True,  # correct the Galactic extinction
          wave_range=None,  # trim input wavelength
          wave_mask=None,  # 2-D array, mask the given range(s)

          # host decomposition parameters
          decompose_host=True, # If True, the host galaxy-QSO decomposition will be applied
          host_prior=False, # If True, the code will adopt prior-informed method to assist decomposition. Currently,
              only 'CZBIN1' and 'DZBIN1' model for QSO PCA are available. And the model for galaxy must be PCA too.
          host_prior_scale=0.2, # scale of prior panelty. Usually, 0.2 works fine for SDSS spectra. Adjust it smaller
              if you find the prior affect the fitting results too much.

          host_line_mask=True, # If True, the line region of galaxy will be masked when subtracted from original
              spectra.
          decomp_na_mask=True, # If True, the narrow line region will be masked when perform decomposition
          qso_type='CZBIN1', # PCA template name for quasar
          npca_qso=10, # numebr of quasar templates
          host_type='PCA', # template name for galaxy
          npca_gal=5, # number of galaxy templates

          # continuum model fit parameters
          Fe_uv_op=True,  # If True, fit continuum with UV and optical FeII template
          poly=True,  # If True, fit continuum with the polynomial component to account for the dust reddening
          BC=False,  # If True, fit continuum with Balmer continua from 1000 to 3646A
```

```
            initial_guess=None,  # Initial parameters for continuum model, read the annotation of this function for
                detail
            rej_abs_conti=False,  # If True, it will iterately reject 3 sigma outlier absorption pixels in the continuum
            n_pix_min_conti=100,  # Minimum number of negative pixels for host continuuum fit to be rejected.

            # emission line fit parameters
            linefit=True,  # If True, the emission line will be fitted
            rej_abs_line=False,
            # If True, it will iterately reject 3 sigma outlier absorption pixels in the emission lines

            # fitting method selection
            MC=False,
            # If True, do Monte Carlo resampling of the spectrum based on the input error array to produce the MC error
                array
            MCMC=False,
            # If True, do Markov Chain Monte Carlo sampling of the posterior probability densities to produce the error
                array
            nsamp=200,
            # The number of trials of the MC process (if MC=True) or number samples to run MCMC chain (if MCMC=True)

            # advanced fitting parameters
            param_file_name='qsopar.fits',  # Name of the qso fitting parameter FITS file.
            nburn=20,  # The number of burn-in samples to run MCMC chain
            nthin=10,  # To set the MCMC chain returns every n samples
            epsilon_jitter=0.,
            # Initial jitter for every initial guass to avoid local minimum. (Under test, not recommanded to change)


            # customize the results
            save_result=True,  # If True, all the fitting results will be saved to a fits file
            save_fits_name=vname ,  # The output name of the result fits
            save_fits_path=path_out,  # The output path of the result fits
            plot_fig=True,  # If True, the fitting results will be plotted
            save_fig=True,  # If True, the figure will be saved
            plot_corner=True,  # Whether or not to plot the corner plot results if MCMC=True

            # debugging mode
            verbose=True,  # turn on (True) or off (False) debugging output

            # sublevel parameters for figure plot and emcee
            kwargs_plot={
            #     'save_fig_path': '.',  # The output path of the figure
                'save_fig_path': path_out,  # The output path of the figure
                'broad_fwhm'   : 1200  # km/s, lower limit that code decide if a line component belongs to broad
                    component
            },
            kwargs_conti_emcee={},
            kwargs_line_emcee={})

end = timeit.default_timer()
os.chdir(path_out)
print(f'Fitting finished in {np.round(end - start, 1)}s')
data = fits.open(os.path.join(path_out,f'{name}'))
import shutil as sl
file_to_copy = f'{name}'
file_to_copy2 = f'{vname}.pdf'
```

```
        destination_directory = 'C:\\Users\\robos\\Honors Thesis Stuff\\Graphing folder (initial testing and stuff)\\demo
            saving folder'

        if (np.abs((data[1].data['Hb_whole_br_peak'][0]/4861)-1)) > 0.00133333333 and data[1].data['Hb_whole_br_snr'] > 5 :
            sl.copy(file_to_copy, destination_directory)
            sl.copy(file_to_copy2, destination_directory)

except (ValueError,KeyError):
    print("No Usable Data")
    os.chdir(path_out)
    data = fits.open(os.path.join(path_out,f'{name}'))
    import shutil as sl
    file_to_copy = f'{name}'
    destination_directory = 'C:\\Users\\robos\\Honors Thesis Stuff\\Graphing folder (initial testing and stuff)\\Unusable Data
        '
    sl.copy(file_to_copy, destination_directory)
    try:
        file_to_copy2 = f'{vname}.pdf'
        sl.copy(file_to_copy2, destination_directory)
    except (ValueError,KeyError,FileNotFoundError):
```

Appendix C: Histogram Code

```
import os
import numpy as np
from astropy.io import fits

#arrays for redshift
OIIIv=np.empty(0)
names=np.empty(0)

# Assign directory
directory = 'C:\\Users\\robos\\Downloads\\Spectrafull'

# Iterate over files in directory
for name in os.listdir(directory):
    data = fits.open(os.path.join(directory,name))
    z = data[1].data['redshift'][0]  # Redshift

    try:
        snr=data[1].data['Hb_whole_br_snr']
        if snr>2 and data[1].data['Hb_br_1_scale'] > data[1].data['Hb_br_2_scale']:
            OIIIv=np.append(OIIIv,(((np.exp(data[1].data['Hb_br_1_centerwave'][0])*5007)/(4861*np.exp(data[1].data['
                OIII5007c_1_centerwave'][0])))-1)*299792.458)
            names=np.append(names,f'{name}')
        elif snr>2 and data[1].data['Hb_br_1_scale'] < data[1].data['Hb_br_2_scale']:
            OIIIv=np.append(OIIIv,(((np.exp(data[1].data['Hb_br_2_centerwave'][0])*5007)/(4861*np.exp(data[1].data['
                OIII5007c_1_centerwave'][0])))-1)*299792.458)
            names=np.append(names,f'{name}')
    except (ValueError,KeyError):
np.savetxt('OIIIv.csv', OIIIv, delimiter=',')
np.savetxt('names.csv',names,fmt="%s")
```

References

[1] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, and F. Acernese et al. Observation of gravitational waves from a binary black hole merger. Phys. Rev. Lett., 116:061102, Feb 2016.

[2] Volker Beckmann and Chris R. Shrader. Active Galactic Nuclei. 2012.

[3] EW Bonning, GA Shields, and S Salviander. Recoiling black holes in quasars. The Astrophysical Journal, 666(1):L13, 2007.

[4] Manuela Campanelli, Carlos Lousto, Yosef Zlochower, and David Merritt. Large Merger Recoils and Spin Flips from Generic Black Hole Binaries. , 659(1):L5–L8, April 2007.

[5] Manuela Campanelli, Carlos O. Lousto, Yosef Zlochower, and David Merritt. Maximum Gravitational Recoil. , 98(23):231102, June 2007.

[6] Bradley W. Carroll and Dale A. Ostlie. An introduction to modern astrophysics and cosmology. Caimbridge University Press, 2006.

[7] Michael Eracleous, Karen T. Lewis, and Hélène M.L.G. Flohic. Double-peaked emission lines as a probe of the broad-line regions of active galactic nuclei. New Astronomy Reviews, 53(7):133–139, 2009. Proceedings of the VII Serbian Conference on Spectral Line Shapes (VII SCSLSA) held in Zrenjanin, Serbia June 15th-19th 2009.

[8] Hengxiao Guo, Yue Shen, and Shu Wang. Pyqsofit: Python code to fit the spectrum of quasars. Astrophysics Source Code Library, pages ascl–1809, 2018.

[9] Patrick Hall. Sdss works. `https://www.yorku.ca/science/phall/ssds-work/`.

[10] Julian H. Krolik. Active Galactic Nuclei: From the Central Black Hole to the Galactic Environment, volume 60. Princeton University Press, 1999.

[11] T.A. Moore. A General Relativity Workbook. University Science Books, 2012.

[12] Hagai Netzer. Revisiting the Unified Model of Active Galactic Nuclei. , 53:365–408, August 2015.

[13] L. Pei, M. M. Fausnaugh, A. J. Barth, B. M. Peterson, and Bentz et al. Space Telescope and Optical Reverberation Mapping Project. V. Optical Spectroscopic Campaign and Emission-line Analysis for NGC 5548. , 837(2):131, March 2017.

[14] Jeremy D. Schnittman and Alessandra Buonanno. The Distribution of Recoil Velocities from Merging Black Holes. , 662(2):L63–L66, June 2007.

[15] SDSS. `https://www.sdss4.org/dr17/`.

[16] Yue Shen, Patrick B. Hall, Keith Horne, Guangtun Zhu, and Ian McGreer et al. The Sloan Digital Sky Survey Reverberation Mapping Project: Sample Characterization. , 241(2):34, April 2019.

[17] Daniel E. Vanden Berk, Gordon T. Richards, Amanda Bauer, Michael A. Strauss, and Donald P. Schneider et al. Composite Quasar Spectra from the Sloan Digital Sky Survey. , 122(2):549–564, August 2001.

[18] Qiaoya Wu and Yue Shen. A catalog of quasar properties from sloan digital sky survey data release 16. The Astrophysical Journal Supplement Series, 263(2):42, 2022.