IMPROVING ALGORITHM AND MODULAR PROGRAMMING IN THE SEARCH OF EHVO IN SDSS QUASAR SPECTRA

Mikel Charles, Dakota Bunger, Cora DeFrancesco, Wendy García Naranjo, Nathnael Kahassai, Michael Parker, & Anish Saurav Rijal

ABSTRACT: Quasars, the most luminous and energetic of Active Galactic Nuclei, are powered by a supermassive black hole with an accretion disk that creates wind-like outflows. These outflows provide us with key information that can contribute to understanding how the matter from the central region interacts with the surrounding galaxy. We observe quasar outflows as blue-shifted absorption features in the spectra. An outflow with speeds between 10-20% of the speed of light is classified as an extremely high-velocity outflow (EHVO). EHVOs have yet to be extensively studied, though their powerful outflows show the most potential for theoretical studies as they might be the most challenging for theoretical simulations.

We searched for EHVOs by developing automated algorithms to normalize spectra and search for broad Carbon IV absorption lines at 10-20% of the speed of light in quasar spectra from the Sloan Digital Sky Survey Data Release 16. The number of spectra available is rapidly growing with each data release which makes it vital to have accurate algorithms to limit the amount of work done manually. Here we discuss improvements that have been made to our algorithms and how we designed the code to be easily accessible to other researchers utilizing our code, as well as future undergraduate students that will continue the work on this project.

Introduction

Active Galactic Nuclei (AGN) are found at the center of galaxies containing an active supermassive black hole surrounded by matter (Carroll et al., 2017, p.1121). The gravitational pull of the black hole attracts matter which begins to spiral around the black hole, forming an accretion disk (Carroll et al., 2017, p.1111). Friction between the spiraling matter causes the matter to heat up and emit light (Lynden-Bell 1969), which allows us to study these distant, otherwise invisible, objects. The friction decreases the energy and angular momentum of the matter which allows the matter to fall inward toward the black hole (Montesinos, 2012). The accretion disk produces outflows, or winds (Carroll et al., 2017, p.1113), that might interact with the surrounding galaxy.

Quasars, short for "quasi-stellar radio source" (Peterson, 1997, p.7), are the most luminous of the AGN and emit massive amounts of energy (Peterson, 1997, p.23). Extremely high-velocity outflows (EHVOs) are quasar outflows that show absorption features in their spectra at 10-20% of the speed of light (Rodríguez Hidalgo et al. 2011 Rodríguez Hidalgo et al. 2020). Studying quasars can provide information about the chemical composition in the surroundings of the quasar. Currently, the most distant quasar - redshift z = 7.642; see below - (Wang et al. 2021) is approximately 13 billion light years away from Earth; quasars with high redshifts such as this allow us to study properties of the early universe.

Spectroscopy is a tool used to study the spectrum of electromagnetic radiation. Light from the source is split, using a prism or a diffraction grating, into wavelengths to create a spectrum (Knight, 2016, p.1064). Spectra can contain emission and/or absorption features over a continuum; these features give us information

about the chemical composition of the object, as well as of the space between the object and the observer. As explained in Carroll et al. (2017, p.111-112), hot, dense gas produces a continuous spectrum; the accretion disk gives rise to the continuum part of the spectrum. Emission features in the spectrum are produced by hot, diffuse gas at the source. Absorption features are produced when the light from the accretion disk travels through gas in the line of sight to the observer. These emission and absorption features are unique to the type of gas that is interacting with the light. We can determine what elements make up the gas by analyzing the wavelength at which these features are found in the spectra. We use spectroscopy in our study to identify EHVO in quasars by searching for spectra containing Carbon IV (CIV; Carbon ionized 3 times) absorption at 10-20% of the speed of light.

Cosmological redshift is a measurement that accounts for the expansion of the universe when observing an object's spectra (Carroll et al., 2017, p.1099). The redshift affects the wavelength that we observe here on Earth (observed wavelength; λ observed); a larger cosmological redshift means that the wavelengths we observe will be shifted more towards the red part of the spectrum. The actual wavelength emitted from the quasar is referred to as the rest frame wavelength (λ_(rest frame)). Cosmological redshift is defined as:

$$z=(\lambda\text{observed} - \lambda(\text{rest frame}))/\lambda(\text{rest frame})$$

(Carroll et al., 2017, p.1100). If we know the redshift of a quasar and the observed wavelength, we are able to calculate the rest frame wavelength and vice versa. Figure 1 shows two quasar spectra from our sample with different redshifts.

We implemented a test to check that the power law is going closely through each anchor point as a first step to determining if the power law is going through the continuum of the spectra.

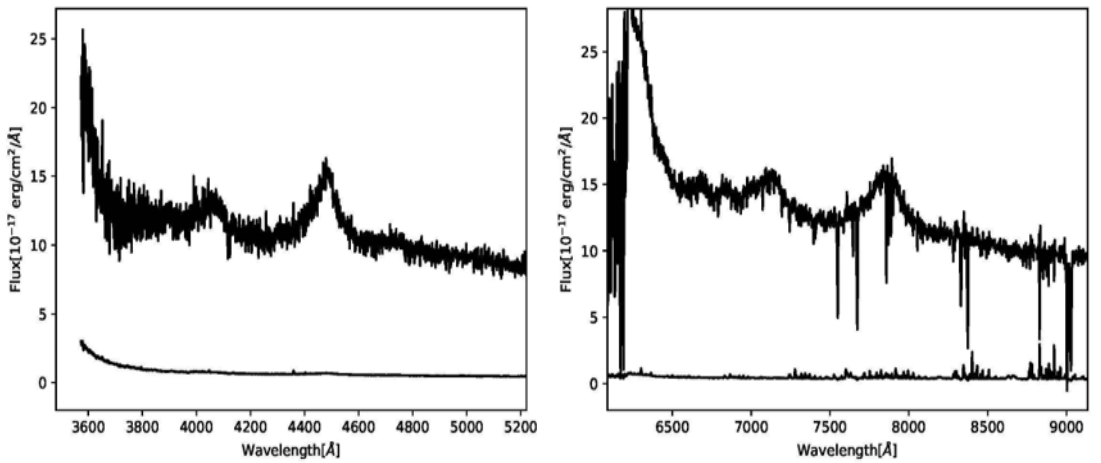


Figure 1:
Examples of quasar spectra at different redshifts. Both spectra are displaying the rest frame wavelength range of (1200-1800 Å). The spectra on the left has an observed wavelength range of approximately (3600-5200 Å) and a redshift of z=1.9008. The spectra on the right has an observed wavelength range of approximately (6000-9000 Å) and a redshift of z=4.074. The quasar with the largest redshift shows the largest observed wavelengths (right).

The expansion of the universe causes the quasars we observe to move away from us, which results in the cosmological redshift. Some of the outflowing gas from the quasars, however, is moving towards us relative to the quasar. Similar to the cosmological redshift is the Doppler shift; an object moving away from us will have a longer wavelength that is

redshifted, or shifted towards the red part of the spectrum, while an object moving towards us will have a shorter, more condensed wavelength that is blueshifted, or shifted towards the blue part of the spectrum (Carroll et al., p.98). The CIV absorption we are searching for to identify EHVOs is blue shifted in the spectrum because of its motion towards us relative to the motion of the quasar.

In order to search for quasars showing EHVO in their spectra, we need large quasar spectra samples to begin with. We use archival, publicly available Sloan Digital Sky Survey (SDSS) quasar spectra. The number of quasars has increased significantly with each of their data releases, which has motivated us to improve our methods for normalizing and searching for absorption within the SDSS spectra. We have developed automated algorithms to normalize spectra and search for absorption features, with a focus on improving the accuracy of these algorithms to reduce the amount of work that needs to be done manually.

An additional goal of the entire project was to make the results, as well as the code, accessible to the public. Thus, it was important to us to write the code in a way that students joining the research group with little coding experience can understand and utilize the code more easily. To do so, constructing a clear and readable code as well as implementing proper documentation was necessary. Many of the programs we started with were not written with the intention of anyone other than the author of the code using them. We have implemented modular programming methods to contribute to the reusability of the code by giving the ability to utilize the same functions in many different programs. To ensure these functions are usable between codes, we rewrote the functions to be as general and accessible as possible.

We are using data from the SDSS Data Release 16 Quasar (DR16Q) catalog. Each data release contains more quasar spectra than the previous; the DR9 observed a total of 87,822 quasars, of which 78,086 were newly observed (Pâris et al. 2012) and the DR16 observed a total of 750,414 with 225,082 quasars being newly observed (Lyke et al. 2020). To ensure we can search for absorption easily within the data, we chose data with a signal to noise ratio (SNR) greater than 10. When searching for EHVOs, we are looking for continuous, broad CIV absorption at 10-20% of the speed of light. CIV outflowing at this speed would appear in the rest frame wavelength range of 1250-1400 Å. We made a secondary cutoff that the redshift must be $1.9 < z_{em} < 6.3$ to ensure that the region we are searching for CIV absorption in is within the data obtained from SDSS. Quasars with a redshift outside this range will not show the region of interest, 1250-1400 Å in rest frame wavelength, in the SDSS spectra since these cover approximately 3600-10400 Å (Lyke et al. 2020) in observed wavelength. After these cutoffs, along with another calculation for SNR in the rest frame wavelength range of 1250-1400 Å, we were left with a parent sample of 18181 quasars for the DR16.

Methods

Accessibility

One of the goals of this project was to make our code accessible to the public. With this in mind, we developed the code in a way that ensured others would be able to use it with ease. Documentation is key when code is being utilized by more than just the authors of the code. We implemented Sphinx documentation - Developed by Georg Brandl in 2008 - to give the necessary information to use the functions within our code. This entailed documenting the required inputs, a brief overview of what each function does, as well as the outputs of the function. Sphinx automatically compiles these into an HTML page containing the documentation in an organized, easy to read format.

Documentation was not the only improvement we made to make our programs more accessible. We organized the code so that all variables and

parameters that may need to be changed are located at the top of the code. This includes, but is not limited to, variables that define the wavelength range for SNR, locations of anchor points, locations of directories to read files or save files, and what spectra to run through the code. Organizing the code in this way makes it easier for future programmers to find the variables that need to be adjusted without needing to sort through the entire code. To ensure others are able to understand what these variables represent, we named the variables to clearly define what they represent. For example, rather than variables x, y, and z, we chose names such as wavelength, flux, and error to be clearer about what each variable represents.

With each new data release, it is likely that the code will need to be updated to account for any changes to the format of input files or future project needs. It is important that this code is written in a way that future undergraduate students working on the project will be able to understand the code to be able to make these changes. In conjunction with the improvements described above, we chose to write the code in order from top to bottom which is more natural to read through. Writing the code so it is easy for others to interpret will expedite the learning curve for future students and will allow them to begin working on the project sooner.

Modularization

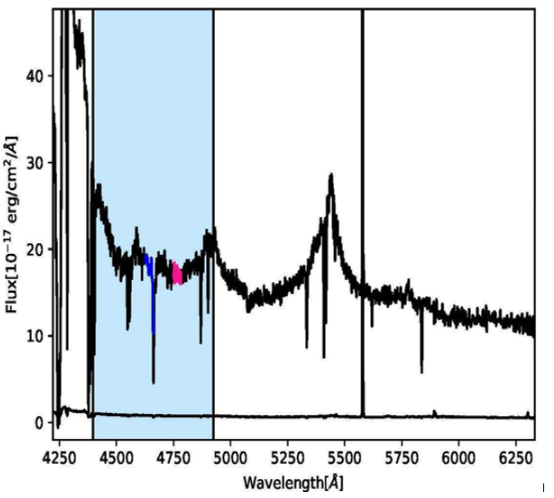
When writing code, multiple functions are often used in other sections of the program or in another program completely. Rather than rewrite that function in its entirety you can 'call' it from a module and use it in your program. Modular programming can turn 30 lines of a code function into one line because you only need to 'call' the name of the module or function that you want to use. Calling the function consists of typing the name of the module and any other parameters you have implemented. This also allows for cleaner code because you have a clear and concise visual representation of what the algorithm is doing. Aside from that, we are also saving time by not having to rewrite any of the modules we created since they have already been written and we only need to call them. It is also very useful because if you need to update a section of your module you can go into it and it will change it in every single spot where it is 'called.' This allows you to change one section of the code as opposed to having to manually change it in the numerous spots you utilized it. Modular programming can be a very powerful tool and is often what is most exercised by programmers. Ultimately, this is why we were sure to write code using modularization in order to streamline the programming needed for this research.

Algorithm Improvements

We developed automated algorithms to assist in the search for CIV absorption at 10-20% of the speed of light, which corresponds to the rest frame wavelength range of 1250-1400 Å as shown in Figure 2. We need to first normalize each spectrum and then search for EHVO CIV absorption in this region; we were most concerned with normalizing this region well since this is where we are searching for EHVO CIV absorption. In the sections below we describe the algorithms and the improvements

Figure 2

Quasar spectrum showing highlighted in blue the region of 1250-1400 Å in rest frame



we carried out from our previous versions in our current work. As mentioned above, with each data release our data sample has increased significantly which motivated us to develop accurate, automated algorithms to reduce the amount of work needed to be done manually.

Normalization Algorithm

Figure 3 shows an example of a spectra that has been normalized by the updated normalization algorithm. Similar to the process described in Rodríguez Hidalgo et al. 2020, we fit each spectra with a power law to normalize the spectra. We chose to fit each spectra using a power law to normalize the spectra since the continuum of a quasar spectra in the wavelength range contained in our data closely follows the shape of a power law (Vanden Berk et al. 2001). We anchored the power law at three locations on each spectra that were typically void of emission and absorption features. The regions we defined for the normalization code to select these anchor points in rest frame wavelengths are: Region A (1280-1290 Å), Region B (1440-1450 Å), and Region C (1690-1710 Å). Each anchor point is calculated by taking the average wavelength, the median flux, and the median error within the defined regions (Regions A, B, and C) for each spectra. Once the spectra is fit with a power law, it is normalized by dividing the flux by it.

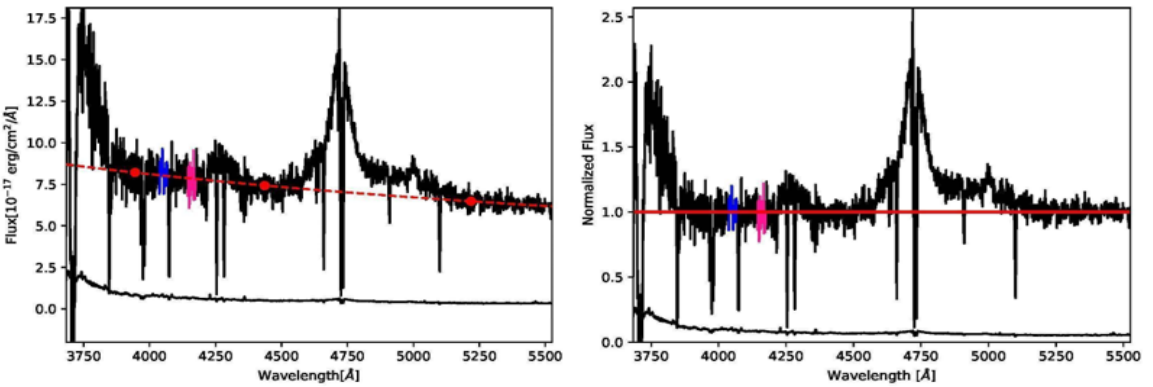


Figure 3:
Example of a successful normalization of one of the quasar spectra in our sample. The left figure shows the original spectrum that has been successfully fit by our automated normalization code with a power law (dashed line) using the three selected anchor points (red dots). The power law is going through the center of the continuum of the spectra, allowing for a proper normalization which is shown on the right, where we have divided the flux values by the power law to obtain the normalized flux values. Blue and pink part of each spectrum indicate our testing regions, as described below.

The locations of the anchor points are instrumental in fitting the spectra well. When an anchor point falls in a region of absorption or emission, the power law averages between the points typically causing a poor fit. The location for the anchor point defined by Region B, previously 1420-1430 Å as described in Rodríguez et al. (2020) was updated to avoid the edge of the SiIV+OIV emission line. Moving this anchor point slightly towards longer wavelengths resulted in a better fit. The region of 1250-1400 Å tends to be difficult to fit due to weak emission lines in the spectra that may or may not be present; some examples of the variability of these emission lines in this region are shown in Figure 4.

Due to the fluctuating ranges of flux values in the sample, we defined "closely" as follows: we calculated the median normalized flux in the rest frame wavelength range 1650-1700 Å then allowed for the anchor points to each be within 10% of this value. This is a region of

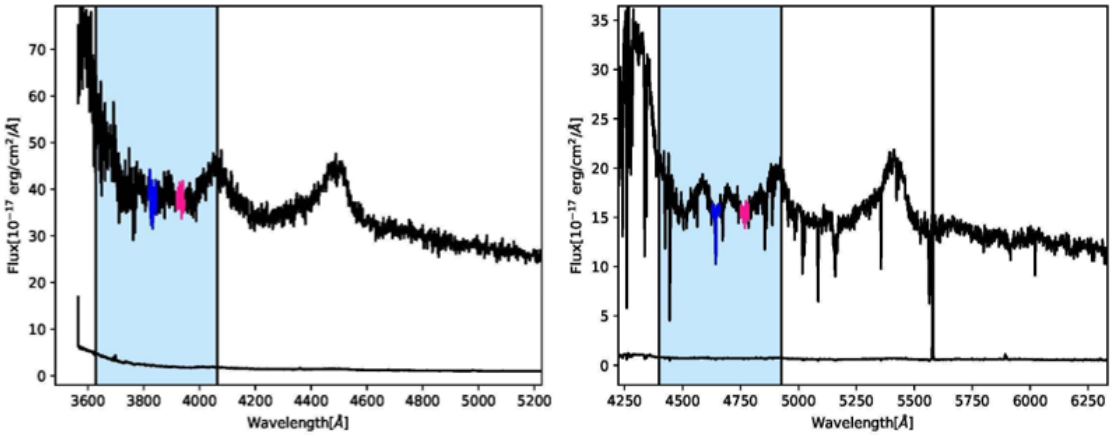


Figure 4:
Examples of the varying emission features in the region of 1250-1400 Å in quasar spectra in our sample. Within this range, the spectrum on the left shows almost no emission within the region of interest (in other words, the spectrum is featureless and flat), while the spectrum on the right shows weak emission lines due to OI (left of blue part) and CII (left of pink part) that can affect our normalization. Blue and pink parts of each spectrum indicate our testing regions, as described below.

the spectra that is typically void of absorption and emission features and allows us to get an accurate scale of the flux values in the spectra. The region surrounding the middle anchor point consistently had unpredictable absorption features causing the power law to fit below the continuum. To account for this, if the initial points are not within the allowed range, the algorithm attempted two other locations for the middle anchor point, in rest frame wavelengths: (1420-1430 Å) and (1460-1470 Å), as shown in Figure 5. If all three locations for the middle anchor point fail to be within the required range from the power law, the spectra is flagged as a bad fit and does not proceed to any other tests.

The region of the spectrum that we were most concerned with normalizing correctly is the region we searched for EHVOs, in rest frame wavelength: (1250-1400 Å). In order to confidently identify absorption there, the power law must follow the center of the continuum in this region. We defined two test regions (see Figure 6), in rest frame wavelengths: the blue test region (1315-1325 Å) and the pink test region (1350-1360 Å).

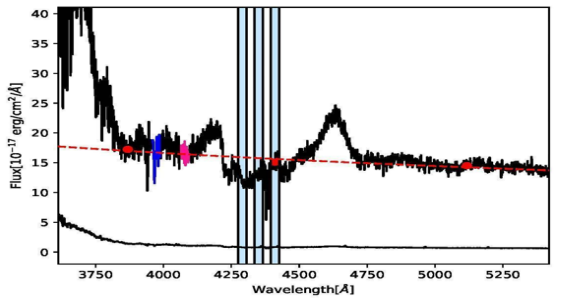


Figure 5:
Example of quasar spectrum normalization showing the three possible locations for the anchor points. The two leftmost locations fall in a region of absorption which would cause a poor fit. This is an example of a spectra that has been fit conservatively but it is still fitting well. The power law is going through the lower portion of the spectrum, but still follows the shape of the continuum.

Figure 6 shows how we use these test regions to check the fit of the power law. In the normalized spectra, we allow for the power law to be within 0.05 above and below the median of each test

region. This test ensures that the powerlaw is going through the centermost region of the spectra. Since quasar spectra are so variable in absorption and emission features, they can be challenging to fit well. The test shown in Figure 6 uses a constant value as the range that the power law can be within; this test ensures that the fit is through the center of the continuum, but resulted in flagging many cases that were fit well as a bad fit. The goal of normalizing the spectra is to be able to later search for absorption below the continuum of 1, so it is important that the power law goes through the continuum; however, the power law does not need to go through the exact center of the continuum to successfully identify absorption.

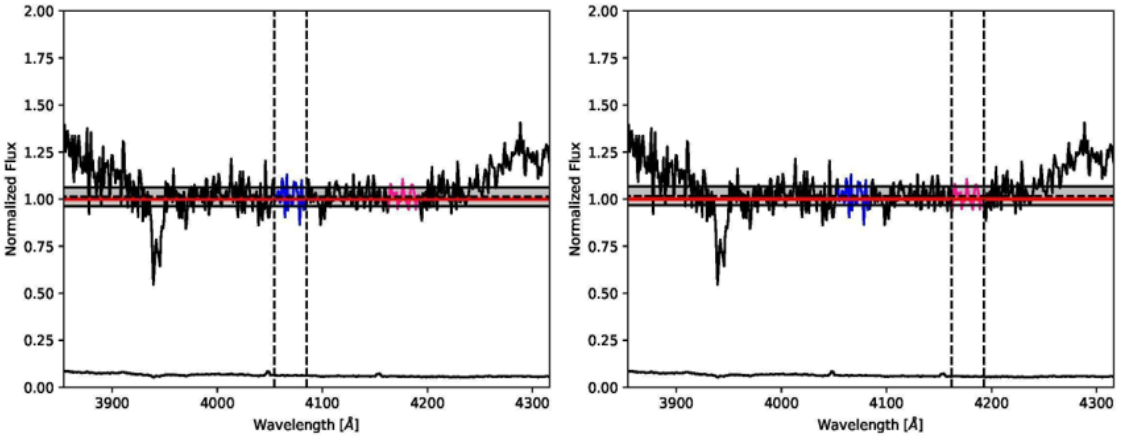


Figure 6:
Zoomed-in version of an example to show the regions that we use as a secondary test for a good normalization. The figure on the left is showing the range for the blue region; the figure on the right shows the range for the pink region. the vertical dashed lines are included to guide the eye to the corresponding test region. The red line is the power law in the normalized graphs. The horizontal black dashed line is the median of the corresponding blue or pink test region. Since the power law is within the shaded region, the spectra is deemed a good fit.

Figure 7 shows the test we implemented to allow for more flexibility with the fit of the power law through the continuum of the spectra. This test expands the range that the power law can be within, keeping in mind that we want to fit the spectra conservatively. Fitting the spectra conservatively means we want to err on the side of fitting the spectra too low through the continuum rather than too high, which could cause false identification of absorption later on.

This test works by expanding the range of flux values that the power law is able to be within, expanding more below than above the median. After testing a variety of values and performing visual inspections, we chose to allow for the power law to be within the region that is 35% above the median normalized flux, and 45% below the median normalized flux (see Figure 7). Expanding the range in this way allows for a dynamic range of values that adjust with the varying flux ranges, rather than choosing a constant range for the power law to be within. This allows us to account for cases that fit well and will allow us to find absorption, but are not necessarily through the very center of the continuum. Spectra that the code deemed to be a good fit based on this test were all visually inspected to ensure they were fit well; the expanded range allowed for more cases to be considered a good fit, of which most were, but with the wider range it was no longer guaranteed that they were fit well.

Spectra that have failed all of the previously described tests are flagged as a bad fit by our code and require visual inspection to determine if they need to be manually normalized. We visually inspect these spectra to determine if they have any absorption in the rest frame wavelength region of (1250-1400 Å). If absorption is present in this region, we utilize the manual normalization feature of the code we developed which allows for the user to manually select the number of anchor points to use as well as the locations of the anchor points. Figure 8 shows an example of a spectra that failed all previous tests and was later normalized manually.

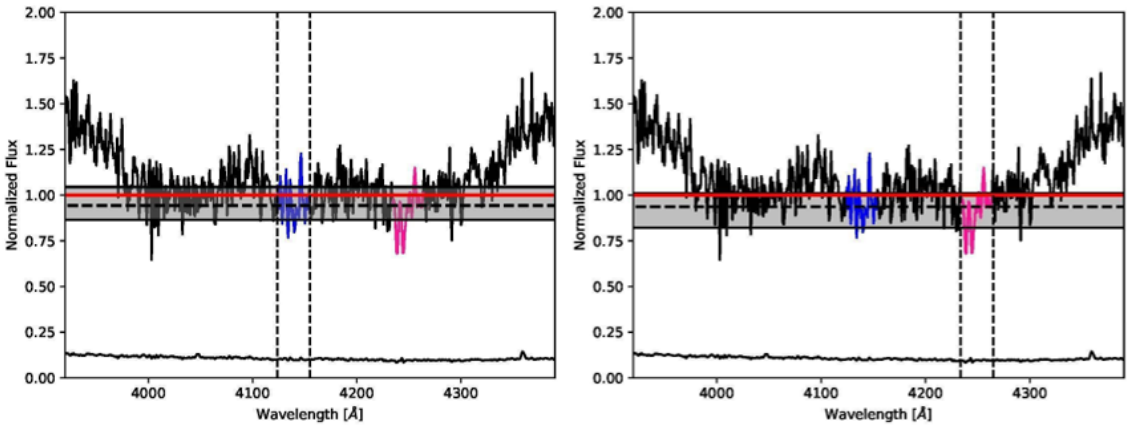


Figure 7:
Zoomed-in version of an example to show the regions that we use as a secondary test for a good normalization. On the left, the expanded range for the blue region is shown and the range for the pink region is shown on the right. The red line indicates the power law, the horizontal black dashed line shows the median of the corresponding region. This spectra was flagged as a bad fit by the test shown in Figure 6, but widening the range with this test has allowed for this spectra to be considered a good fit.

Figure 7 shows the test we implemented to allow for more flexibility with the fit of the power law through the continuum of the spectra. This test expands the range that the power law can be within, keeping in mind that we want to fit the spectra conservatively. Fitting the spectra conservatively means we want to err on the side of fitting the spectra too low through the continuum rather than too high, which could cause false identification of absorption later on.

This test works by expanding the range of flux values that the power law is able to be within, expanding more below than above the median. After testing a variety of values and performing visual inspections, we chose to allow for the power law to be within the region that is 35% above the median normalized flux, and 45% below the median normalized flux (see Figure 7). Expanding the range in this way allows for a dynamic range of values that adjust with the varying flux ranges, rather than choosing a constant range for the power law to be within. This allows us to account for cases that fit well and will allow us to find absorption, but are not necessarily through the very center of the continuum. Spectra that the code deemed to be a good fit based on this test were all visually inspected to ensure they were fit well; the expanded range allowed for more cases to be considered a good fit, of which most were, but with the wider range it was no longer guaranteed that they were fit well.

Spectra that have failed all of the previously described tests are flagged as a bad fit by our

code and require visual inspection to determine if they need to be manually normalized. We visually inspect these spectra to determine if they have any absorption in the rest frame wavelength region of (1250-1400 Å). If absorption is present in this region, we utilize the manual normalization feature of the code we developed which allows for the user to manually select the number of anchor points to use as well as the locations of the anchor points. Figure 8 shows an example of a spectra that failed all previous tests and was later normalized manually.

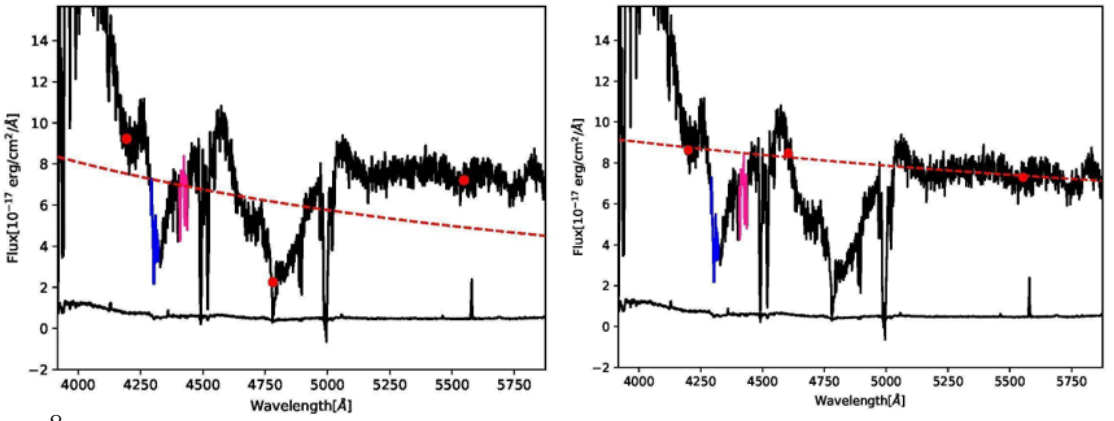


Figure 8:
(left) Example of a spectra with the middle anchor point falling in absorption, pulling the fit down. Cases such as this are flagged as a bad fit and require manual normalization, which is shown in the figure on the right.

In an effort to ensure we have all EHVO cases, another "test" was implemented to find cases in which either the blue or pink test regions were entirely beneath the power law, as shown in Figure 9. Aside from very poorly fit cases, this provides a list of cases in which there is possible absorption within these regions which corresponds to the region we are searching for CIV absorption at 10-20% of the speed of light. While this method does not flag all EHVO cases, it has shown to be a promising check for the absorption algorithm (see section 3.2.2) to ensure we have all cases that have some absorption within that region.

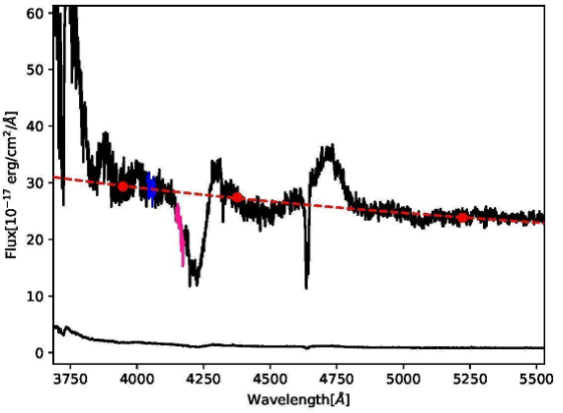


Figure 9:
This spectra was flagged by our test for absorption due to the pink test region being completely beneath the power law, even when the power law fit is good in the rest of the region.

Overall, these improvements have not only increased the number of spectra that fit well, but have also increased the accuracy of the flags within the code. With a data set that is growing with each data release, it is important for the code to be accurate. This reduces the number of spectra that need to be manually normalized in order to analyze and expedites the process of the search. A well normalized spectra is crucial in searching for absorption, so the fit and accuracy of flags within the algorithm is also important in obtaining accurate results from the absorption algorithm.

Absorption Algorithm

The next step is to search for absorption features in our normalized spectra. As mentioned in the introduction, we are searching for CIV absorption at 10% - 20% of the speed of light. With the continuum of our spectra being normalized to a value of one, we are searching for broad absorption stronger than 10% of the continuum, so we search for flux values in our spectra that drop below a value of 0.9. Observing broad absorption lines is a key signature of the presence of supermassive black holes, and thus AGN, as not any other phenomena can produce these features (Peterson,1997) .

Figure 10 shows the normalized spectra vs velocity of the outflow. When our conditions for deep broad absorption are met, this is signified by a red column which begins when the spectra goes below 10% of the continuum and ends where the spectra exits this 10% value. The velocity values are saved from these two instances. Assuming that our finding is due to CIV, we locate and show where Carbon II (CII) would be (blue column), where Oxygen I (OI) would be (yellow column), but most importantly, where CIV would be located if the absorption we found is due to Silicon IV (SiIV) and not in fact CIV (gray column).

The method we use to characterize broad absorption is by calculating the Balnicity index (BI) (Weymann et al. 1991). The equation to calculate the BI of extremely-high velocity absorption consists of an integral where we set the bounds to values of -60,000 km/s to -30,000 km/s (which is 0.1 - 0.2 the speed of light), the function inside the integral is one subtracted by the normalized flux (as a function of velocity) and this is all over the quantity of 0.9. There is also a leading scaler value of C: the C is initially set to zero but will be set to a value of 1.0 whenever there is 2,000 km/s of continuous absorption. It should also be noted that the BI does not begin to be calculated until this value of 2,000 km/s is met (Weymann et al. 1991):

$$BI = -\int_{30{,}000}^{60{,}000} [1 - f(v)/0.9]C\, dV$$

Due to this characteristic, we visually print a horizontal black line on the spectra plot to show where that threshold is met (the black horizontal line can be seen in figure 10 within the red column).

As mentioned previously the normalized flux is converted to a velocity for the BI calculation so we must do this in our code as well. To account for relativistic speeds, we use the equation:

$$\beta = \left[\frac{(1+z)^2 - (1+z_{abs})^2}{(1+z)^2 + (1+z_{abs})^2}\right]$$

(Perry et al. 1978), where β is the relative velocity between the outflow and the quasar, z is the redshift of the quasar, z_{abs} is the redshift of the outflow. We utilize the algorithm originally described in Rodríguez Hidalgo et al. (2020) to output key values such as the BI value, the velocity value where the spectra first goes below 10% of the continuum (v_{min}), the velocity value where the spectra goes back above 10% of the continuum (v_{max}), the depth of this absorption trough, and the equivalent width (EW).

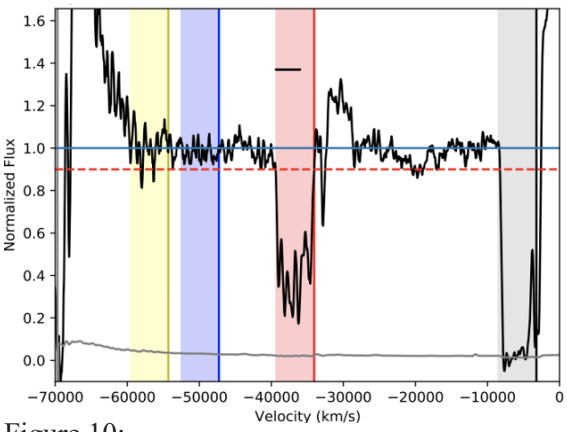


Figure 10:
Example of a quasar spectra where absorption is found in the region of interest. The light blue horizontal line at 1 signifies where the normalized continuum lies. The red horizontal dashed line shows when our spectra is at 10% of the continuum. The red column represents where EHVO CIV would be found; the blue column represents where Carbon II (CII) would appear; the yellow column represents Oxygen I (OI); and the gray column shows where CIV

would be if the red column was Silicon IV (SiIV) and not CIV. The horizontal black line within the red column signifies where the 2,000 km/s of broad absorption has been met and where we can begin to calculate the BI value.

As mentioned above and shown in Fig 10, some of the absorption found in the region of interest is not due to CIV but SiIV instead. We decide on the nature of the absorption upon visual inspection. If there is corresponding absorption in the gray column then we know we have found absorption due to SiIV and not CIV, as SiIV absorption is always accompanied by CIV as C is an element more common in the universe. Figure 10 shows an example of absorption due to SiIV. Figure 11 shows an example of a spectra with broad absorption due to CIV, as shown there is no corresponding absorption located in the gray column.

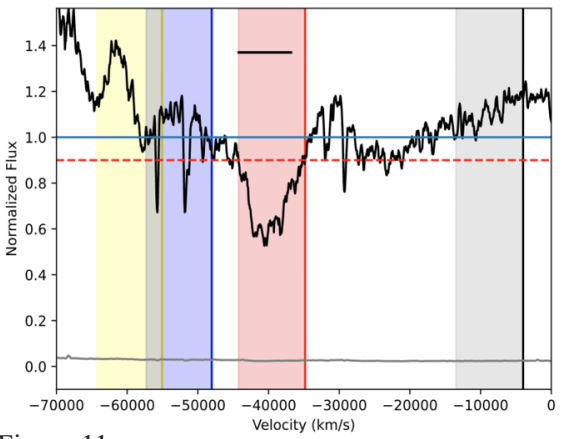


Figure 11:
Showing an example of a spectra with broad absorption due to CIV.

With our goal of adding accessibility and modularization to this code, we added several improvements. As mentioned in our above modularization section, we rewrote this code in the order the calculations and functions are consecutively executed in order to clearly see and read how the algorithm operates and we created several modules or functions that can be called from other programs to run many intermediate tasks. We were also sure to have descriptive variable names and any quantities we may want to change located at the top of the code so that they are easily accessible. The code is now more easy to read and new programmers can call it from other programs to calculate velocities and absorption properties in a simple way.

Preliminary Results

After all the improvements discussed above, we ran our codes over the 18181 quasar spectra. Our normalization code was able to automatically normalize 16511 quasar spectra correctly. We visually inspected all the spectra flagged as bad normalization and used our manual normalization over those that we suspect of showing absorption in the wavelength region of interest. Then, we ran our absorption searching code over these 16823 quasar spectra and found that 1127 showed potential CIV EHVO. As explained above, many of these identifications would correspond to SiIV as it is more common at these wavelengths, but we currently have ~150 cases that are promising EHVO candidates; we are currently finalizing our visual inspection. Once completed, we will have compiled the latest and largest list of EHVO in quasar spectra in the scientific literature. We will share this list with the community so follow-up studies, both theoretical and observational, can be carried out over these interesting objects.

Acknowledgements

Support for this work was provided by the Royalty Research Fund (A162630) at the University of Washington. Additionally, M.M.C and W.F.G.N would like to thank the Washington NASA Space Grant Consortium, funded by the NASA Office of STEM Engagement, NASA Award 80NSSC20M0104. All authors would like to thank the NSF-REU grant at UW Bothell for additional support (#2050928), as well as Pat Hall for fruitful discussions and initial processing of spectra.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy

Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics | Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

References

Carroll, B. W., Ostlie, D. A. (2017). *An Introduction to Modern Astrophysics*. Cambridge University Press.

Knight, R. D. (2017). *Physics for Scientists and Engineers: A Strategic Approach with Modern Physics*. Pearson Education Incorporated.

Lyke, B. W., Higley, A. N., McLane, J. N., Schurhammer, D. P., Myers, A. D., Ross, A. J., Dawson, K., Chabanier, S., Martini, P., Busca, N. G., du Mas des Bourboux, H., Salvato, M., Streblyanska, A., Zarrouk, P., Burton, E., Anderson, S. F., Bautista, J., Bizyaev, D., Brandt, W. N., …, Weaver, B. A. (2020). The Sloan Digital Sky Survey Quasar Catalog: Sixteenth Data Release. *ApJS*, 250(1), 8. https://doi.org/10.3847/1538-4365/aba623

Lynden-Bell, D. (1969). Galactic Nuclei as Collapsed Old Quasars. *Nature*, 223, 690-694. https://doi.org/10.1038/223690a0

Montesinos, M. (2012). Review: Accretion Disk Theory. *arXiv*. https://doi.org/10.48550/arXiv.1203.6851

Pâris, I., Petitjean, P., Aubourg, E., Bailey, S., Ross, N. P., Myers, A. D., Strauss, M. A., Anderson, S. F., Arnau, E., Bautista, J., Bizyaev, D., Bolton, A. S., Bovy, J., Brandt, W. N., Brewington, H., Browstein, J. R., Busca, N., Capellupo, D., Carithers, W., …, Yèche, C. (2012). The Sloan Digital Sky Survey quasar catalog: ninth data release. *Astronomy & Astrophysics*, 548, A66. https://doi.org/10.1051/0004-6361/201220142

Perry, J. J., Burbidge, E. M., & Burbidge, G. R. (1978). Absorption in the Spectra of Quasi-Stellar Objects and BL Lacertae Objects. *Publications of the Astronomical Society of the Pacific*, 90, 337–366. https://doi.org/10.1086/130340

Peterson, B. M. (1997). *An Introduction to Active Galactic Nuclei*. Cambridge University Press.

Rodríguez Hidalgo, P., Moiz Khatri, A., Hall, P. B., Haas, S., Quintero, C., Khatu, V., Kowash, G., & Murray, N. (2020). Survey of Extremely High-Velocity Outflows in Sloan Digital Sky Survey Quasars. *ApJ*, 896(2), 151. https://doi.org/10.3847/1538-4357/ab9198

Vanden Berk, D. E., Richards, G. T., Bauer, A., Strauss, M. A., Schneider, D. P., Heckman, T. M., York, D. G., Hall, P. B., Fan, X., Knapp, G. R., Anderson, S. F. , Annis, J., Bahcall, N. A., Bernardi, M., Briggs, J. W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., & Zheng, W. (2001). Composite Quasar Spectra from the Sloan Digital Sky Survey. *The Astronomical Journal*, 122, 549-564. https://doi.org/10.1086/321167

Wang, F., Yang, J., Fan, X., Hennawi, J. F., Barth, A. J., Banados, E., Bian, F., Boutsia, K., Connor, T., Davies, F. B., Decarli, R., Eilers, A. C., Paolo Farina, E., Green, R., Jiang, L., Li, J. T., Mazzucchelli, C., Nanni, R., Schindler, J. T., …, Yue, M. (2021). A Luminous Quasar at Redshift 7.642. *ApJL*, 907(1), L1. https://doi.org/10.3847/2041-8213/abd8c6

Weymann, R. J., Morris, S. L., Foltz, C. B., & Hewett, P. C. (1991). Comparisons of the Emission-Line and Continuum Properties of Broad Absorption Line and Normal Quasi-Stellar Objects. T*he Astrophysical Journal*, 373, 23-53. https://doi.org/10.1086/170020