

# Reverberation Mapping of Active Galactic Nuclei

Matthew A. Bellardini\*

*Department of Physics and Astronomy, University of California, Davis, California 95616*

(Dated: January 31, 2017)

The masses of active galactic nuclei (AGN) are closely correlated with characteristics of their host galaxy, such as stellar concentration and luminosity. Reverberation mapping is one technique employed to measure the mass of AGN, and it is the technique used in this campaign. Reverberation mapping utilizes variations in the flux of the broad emission lines in the spectra of AGN. These variations arise from fluctuations in the central continuum source with a delay resulting from light travel time. The goal of this project is to reduce data of AGN1435 and AGN1438, taken with the Very Large Telescope at the Paranal observatory in Chile, such that reverberation mapping can be performed. The data reduction pipeline is written using the Python programming language and includes external calls to methods written using the IRAF programming language. The reduction pipeline converts raw data to science frames, but currently has an issue dealing with cosmic ray artifacts and proper wavelength calibration. This paper provides a detailed description of the methods used to reduce data.

## I. INTRODUCTION

### A. The Relevance of Active Galactic Nuclei

Active galactic nuclei, AGN, are a topic of extreme importance to extragalactic astronomers interested in galaxy evolution. This is because there are galaxies in the universe that no longer form stars and galaxies that are still forming stars. The mechanism that causes the cessation of star formation is not completely understood by astronomers, but one hypothesis is a sort of AGN and stellar regulatory feedback<sup>1</sup>. The thought process is essentially that the materials used to form stars within a galaxy are expelled from the galaxy via jets at the poles of the AGN and through stellar outflow. The jets of an AGN can be seen in Fig. 1; when seen edge on we refer to the AGN as a blazar. Stellar outflow results from high velocity stellar winds produced by massive young stars and the supernovae explosions of dying stars.

The goal of this project is to look at the relationship between an AGN and its host galaxy. Specifically, we know that the mass of an AGN correlates with a multitude of properties of its host galaxy, such as luminosity and stellar concentration<sup>2,3</sup>. These correlations have been examined for nearby galaxies, but to strengthen the correlation we are looking at galaxies that are at higher redshifts<sup>4</sup>. The technique used for this project is reverberation mapping. This technique looks at variations in atomic emission lines within a region of the AGN to calculate size and mass.

### B. The Structure of Active Galactic Nuclei

AGN are compact objects present in the centers of active galaxies which emit the excess luminosity observed in these galaxies. A galaxy is considered active when its total luminosity exceeds its expected luminosity based on the number of stars present. This excess luminosity typically comes in the form of less common stel-

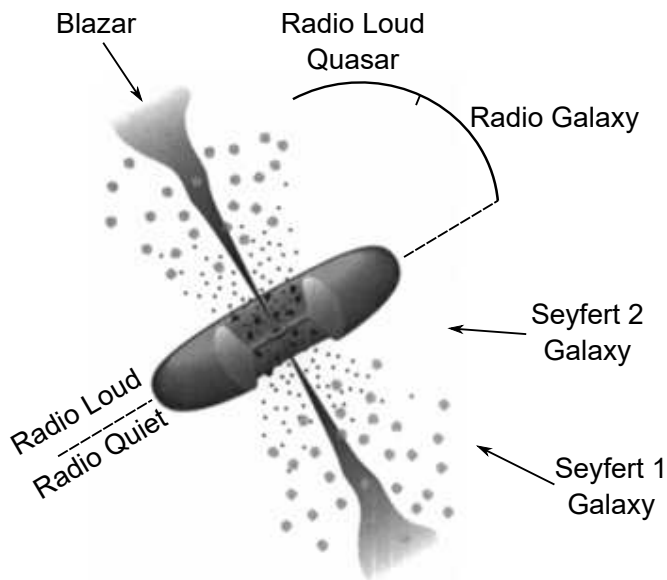


FIG. 1. The type of AGN seen by the observer is dependent on the angle the observer makes with the AGN. Each of the AGN types are the same structure; they are categorized to give the angle of observation. Original source: C.M. Urry and P. Padovani<sup>7</sup>

lar wavelengths such as radio or x-ray<sup>5</sup>. Active galaxies were first discovered in 1908 by Edward Fath when he observed very high intensity emission lines radiating from NGC 1068<sup>5</sup>. However, the first classification of an active galaxy wasn't until the 1940's by Carl Seyfert<sup>6</sup>. There are a multitude of AGN classifications; such as blazars, quasars, radio galaxies, Seyfert 1 galaxies, and Seyfert 2 galaxies. The classification of the AGN hinges entirely on the angle it is observed from, as seen in Fig. 1.

All AGN regardless of classification have the same sort of structure which can be seen in Fig. 2. The central region of an AGN is a super massive black hole (SMBH); this provides nearly all of the object's gravitational force.

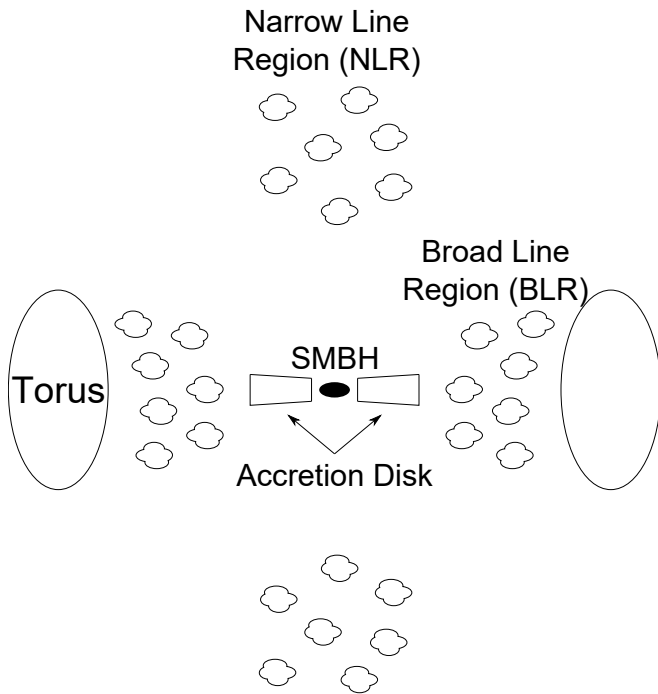


FIG. 2. Cross sectional view of an AGN. At the center of the AGN is a super massive black hole (SMBH) which is the primary source of mass. Surrounding the SMBH is an accretion disk composed of dust. The temperature of the accretion disk is responsible for the shape of black body radiation curve, given off by the accretion disk, which makes up the continuum of the AGN. Surrounding the accretion disk is the broad line region (BLR). This is a region of gas whose motion is governed by the gravitational force of the SMBH. Gas in this region is excited by radiation from the accretion disk, and emits spectral lines which are Doppler broadened. Around the BLR is an obscuring torus made of dust, this is noticeable in Seyfert 2 galaxies, but all observations for this project were of Seyfert 1 galaxies. Far from the BLR is a narrow line region (NLR). The NLR is far enough from the SMBH that its motion isn't Doppler broadened. Also, it is far from the accretion disk so variations in continuum radiation take a long time to propagate to this region in comparison to the BLR.

The SMBH is surrounded by an accretion disk composed of dust, some of which is infalling towards the SMBH. As matter infalls the gravitational energy heats the accretion disk, through viscous heating<sup>8</sup>, causing black body radiation, which is responsible for the continuum emission of the AGN. The continuum emission of the accretion disk varies with time depending on the dynamics of the infalling matter. Surrounding the accretion disk is the broad line region (BLR). This region has a radius that ranges in size from several light days to over 100 light days<sup>9</sup>. It is a region of gas whose motion is governed by the gravitational attraction of the SMBH. Continuum radiation from the accretion disk excites gas in the BLR which causes spectral emission. The motion of the BLR causes a Doppler broadening of the emission features. This means there is some material moving away from the

observer, which is red shifted, and some material moving towards the observer, which is blue shifted, leading to wider spectral lines. Surrounding the BLR is an obscuring torus of dust. However, it only obscures the AGN for Seyfert 2 galaxies, all observations for this project were of Seyfert 1 galaxies, so the torus can be ignored.

The outermost region of the AGN is the narrow line region (NLR). The NLR is a region of gas that is far enough from the SMBH that its motion isn't dominated by the gravitational attraction of the AGN. As a result, when gas in the region is excited by radiation from the continuum the emission lines are not Doppler shifted, so they don't show the same broadening as features in the BLR. Also, the NLR is far enough from the accretion disk that the flux of the features varies on a much larger timescale than the flux of the BLR and of the continuum.

## II. THEORY

### A. Reverberation Mapping Technique

Reverberation mapping (RM) is a technique utilized on AGN to measure the geometry and kinematics of the BLR and the accretion disk<sup>8</sup>. It is also used to make an indirect measurement of AGN mass through calculation of a lag time. To understand how RM works, one must first be aware of how radiation propagates within the AGN. Recall that the continuum radiation of the AGN comes from viscous heating of the accretion disk as matter infalls to the SMBH. Different wavelength ranges of the continuum come from different radial depths of the accretion disk. This is because the higher energy radiation of the continuum must come from the hotter portion of the accretion disk, whereas the lower energy comes from the cooler portion of the accretion disk. The hotter portion of the accretion disk is closest to the SMBH, and the temperature decreases with distance away from the SMBH. This relationship between emitted wavelength and temperature comes from the very basic relationship in Wien's displacement law:

$$\lambda \propto \frac{1}{T} \quad (1)$$

The same relationship can be applied to the broad line emissions that come out of the BLR. The higher energy emission lines come from regions closer to the accretion disk and the lower energy emission lines come from regions in the BLR further from the accretion disk. This is because gas in the BLR scatters the continuum radiation as the gas is excited.

RM is effectively a measurement of the lag time,  $\tau$ , for different energy emissions to vary in flux. It is most simple to think of how this applies to the BLR. Imagine there is an increase in continuum emission from the accretion disk; this leads to an increase in the measured flux of the continuum, which in turn will lead to an increase in the measured flux of the broad line emissions.

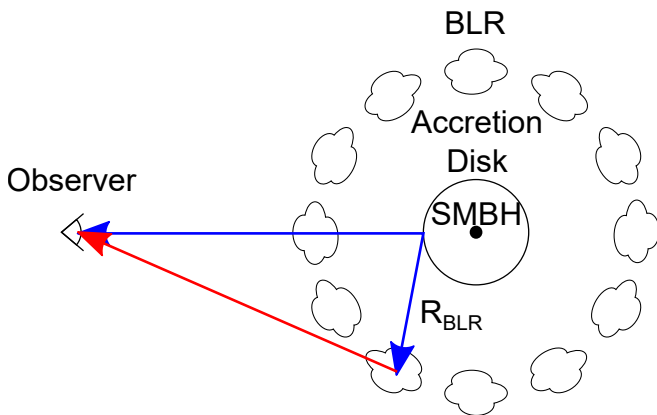


FIG. 3. The light from the continuum (blue) is the light that travels directly from the accretion disk to the observer. The light from the BLR (red) reaches the observer with a time delay  $\tau$  because the continuum light must first travel some radial depth into the BLR,  $R_{BLR}$ , before exciting the gas. Then the emissions of the excited gas travel to the observer.

However, there is a delay time associated with these increases in flux. This delay time is a direct consequence of the speed of light in a vacuum,  $c$ . The BLR has a radius in the range of light days to light months, so the light must take this amount of time to get to the region before it can excite the gas which emits the observed feature lines. A diagram of this lag time can be seen in Fig. 3. The continuum emission comes directly from the accretion disk of the AGN. However, the light from the BLR is only emitted after light from the continuum travels some radial depth into the BLR,  $R_{BLR}$ . This leads to a lag time of:

$$\tau = \frac{R_{BLR}}{c} \quad (2)$$

Now this may seem like an oversimplification of the process going on, because the light from the continuum travels radially outward different portions of the BLR will lag at different times. This is why the lag time that is calculated is an average lag time, not the lag time for a specific segment of the BLR.

Similarly,  $\tau$  can be calculated for different bands of the continuum. Bands are just a means of dividing wavelength into various ranges. A plot of sensitivity as a function of wavelength for several bands can be seen in Fig. 4. As stated previously, different bands of continuum emission come from different depths of the continuum. Because of this, the lag time can be calculated between bands, which gives valuable insight about the physical properties of the accretion disk, but that is beyond the scope of this project.

### B. Calculating Lag Time $\tau$

To calculate the lag time of a given emission feature, or continuum band, light curves are generated. A light

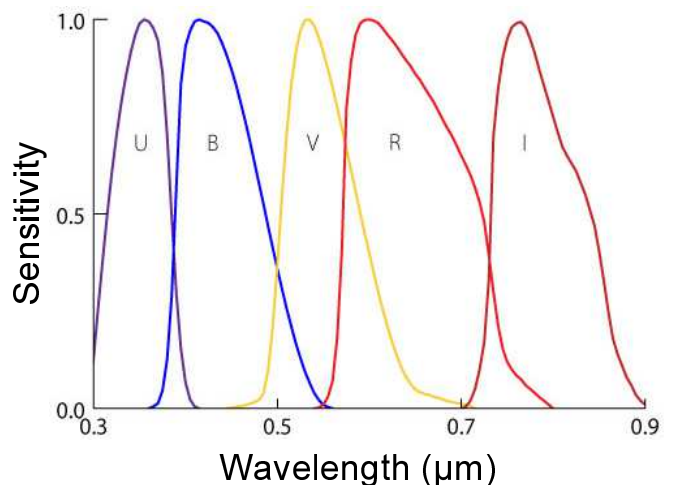


FIG. 4. A plot showing sensitivity as a function of wavelength for bands U,B,V,R, and I. Each band is effective over only a specific wavelength. Data from Johnson and Morgan, Cousins, and Matthews and Sandage<sup>10–12</sup>. Plot from Brainerd<sup>13</sup>.

curve is a plot of flux vs time. For the BLR the light curves of different features are compared to the reference light curve of the continuum of emission. If there is a peak in both light curves then a Markov Chain Monte Carlo in conjunction with a cross correlation algorithm is used to calculate the necessary time shift to align the peaks of the data<sup>14</sup>. This time shift is the lag time,  $\tau$ , and it is negative for emission lines at higher energies than the reference, and positive for emission lines lower energies than the reference. Because of a lack of periodicity for the intensity peaks in of the continuum emission, RM campaigns need to collect data over very large periods of time; ideally data are collected weekly for months at a time.

To generate the light curve of the BLR emissions, raw data are first reduced. After, the spectrum of each observation is scaled using the NLR emissions. These can be used for scaling because, as previously stated, the flux of them varies on a substantially larger time scale than that of the continuum or BLR. The scaling technique is outlined extensively in Ernst van Groningen's, *An Algorithm for the Relative Scaling of Spectra*<sup>15</sup>. Once the spectra are scaled, the flux is integrated across each BLR emissions. This flux is then plotted against the time of observation, for each observation, to create the light curve for any given spectral line. The continuum light curve is created using imaging data taken remotely by the Las Cumbres Observatory Global Telescope Network (LCOGT). The robotic telescopes that form the LCOGT are located around the world and take imaging data each night that conditions permit. The continuum light curve made by these data is updated daily with new observations.

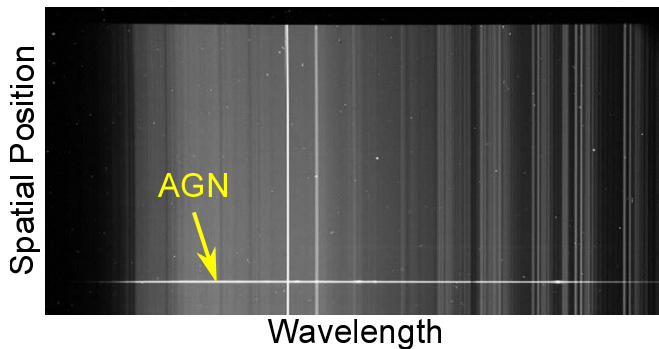


FIG. 5. Raw datum of AGN1435 taken on April 27, 2016. The x axis is wavelength, the y axis is spatial position in the sky, and the brightness of each pixel signifies relative flux. The AGN, identified in yellow, can be seen in the lower half of the data.

### III. DATA REDUCTION PROCESS

Data reduction is called thus because it is the process of taking raw data and reducing it down to only the science frame. This process involves several steps, which I wrote a Python program to implement. The Python program calls on functions from an old programming language, IRAF, to do much of the reduction. This is because IRAF has many predefined functions, which Python is lacking, that can do different data reduction techniques. These techniques are only effective at data reduction provided the parameters for the functions are properly defined, which requires first manually reducing data. The reduction process is a crucial step for analyzing the data; the slightest error in data reduction can completely skew the results of the project. This is because for some AGN the fluctuations in light curves are so minor,  $\sim 1\%$ , that peaks can be improperly identified, or noise can prevent peaks from being resolved.

#### A. Raw Data

The data for this project come from the Very Large Telescope (VLT) at the Paranal observatory in Chile. There were two AGN studied in this project, AGN1435 and AGN1438. As of the time of writing this paper, there have been 14 nights of observations for AGN1435 and 15 nights for AGN1438. However, one of the nights for AGN1438 produced poor data that were ignored. During each observing night the AGN were observed 2 times with exposure times between 400s and 500s. The raw data are the output of the ccd camera on the VLT, a typical raw datum can be seen in Fig. 5. The x axis of the image is wavelength, the y axis is spatial position in the sky, and the brightness of each pixel illustrates its relative flux. The AGN can be seen as the bright horizontal line in the bottom half of the image.

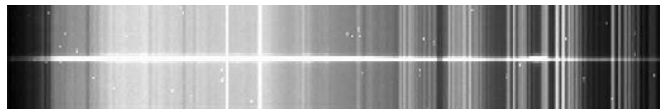


FIG. 6. Datum from AGN1435 that has been trimmed and subject to bias and flat correction.

#### B. Trimming, Flat Correction, and Bias Correction

The next step in data reduction is to trim the data and then apply a bias correction and a flat correction. Trimming the data simply cuts away all data not within a certain specified rectangle of the image. This is done to isolate the AGN and make the data faster to manipulate. Bias is essentially *dark current* in the ccd. This is just noise that is intrinsic to the instrument. To correct for bias, bias frames are taken each observing night. These are 0s exposure frames; with no exposure time only instrumental noise is captured in the frame. All bias frames for each night are averaged into “master” bias images which are subtracted from the raw data. Finally, flat correction is applied. This is a calibration that scales pixel sensitivity because each pixel of the ccd has a different sensitivity. Flat correction is done using flat frames that are taken each observing night. The flat frames are observations of a flat lamp; this is a light source of constant intensity. Similar to the bias frames, the flat frames for each night are combined into “master” flat images. The program `ccdproc` in IRAF then scales the flux of the science data using the “master” flat images. The output after performing each of these steps can be seen in Fig. 6.

#### C. Extracting a Spectrum

The next step in the data reduction is extracting spectra. The term extraction really just means the plane of data points is being further reduced to a single row of data points. The extraction process provides a spectrum of the AGN that is the ccd count of photons as a function of the ccd pixels. To perform the extraction, an individual column of data is plotted using counts as a function of spatial position. The position where the AGN lies on the plot is a bright peak, as the rest of the background is fairly uniform. Upper and lower positional bounds are placed on the AGN location; and then upper and lower ranges of background data are denoted. The IRAF program `apsum` then takes these bounds and traces the AGN all along the image; it looks for the brightest pixels in the range specified. Once a function representing the pixel location of the AGN is generated by `apsum`, the spectrum, i.e. counts as a function of pixel, is extracted along the AGN. From this spectrum the average value of the designated background range specified is subtracted to remove background noise. The output spectrum looks

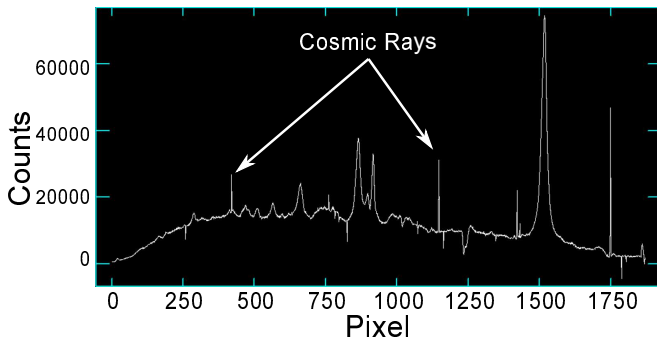


FIG. 7. Extracted spectrum of AGN1435, the spectrum is a plot of counts registered by the ccd as a function of pixel location on the ccd. The bright narrow peaks in the data (labeled in white) are cosmic rays. These narrow emissions come from high energy particles and serve as another source of noise in this data set.

like Fig. 7. The bright narrow peaks in the spectrum are cosmic rays.

Cosmic rays come from high energy particles traveling through space at relativistic speeds<sup>16</sup>. These rays function as noise in this data set and aren't removed through the basic reduction process. The techniques for their removal are employed just after the spectral extraction stage. One technique is to use a cleaning algorithm built into the IRAF program *apall*. Another method for removing cosmic rays is to use the program *Laplacian Cosmic Ray Identification (L.A. Cosmic)* by Yale professor Pieter G. van Dokkum. More on the efficacy of these in the results section.

#### D. Wavelength Calibration

The next step in the data reduction process is wavelength calibration. As seen in Fig. 7 the x axis is in units of pixels; a true spectrum needs to be flux as a function of wavelength. A wavelength calibration is a means to convert the spectrum from pixels to Angstroms ( $10^{-10}\text{m}$ ). The IRAF program *dispcor* is used to perform a wavelength calibration.

The first step in wavelength calibration is taking a spectrum of an arc lamp, thereby generating a wave file. This is just a lamp with an element inside it which will give off emission lines at certain wavelengths that are well documented. Next, the emission lines in this wave file are manually assigned wavelengths, and an IRAF program, *identify*, fits a polynomial function of wavelength as a function of pixel to the features. Because of the high caliber of the VLT, the function for wavelength calibration varies almost imperceptibly from night to night, but these variations are still accommodated for. Each observing night has an arc lamp exposure taken, and the wavelength calibration function is applied to each of these functions using the IRAF program *reidentify*. This program assumes that the function generated on one night

is approximately correct for all nights and makes slight modifications to it to more accurately align it with the emission lines observed for a given night. Finally *dispcor* is used; it takes a spectrum from a given night and the corresponding wavelength calibration data and applies the corrective function.

#### E. Flux Calibration

The final step in the reduction process is to perform a flux calibration. As mention before, a spectrum is flux as a function of wavelength. The previous steps only got the data to counts as a function of wavelength.

Flux calibration is performed using an observation of a standard star. When astronomers reference a standard star what is meant is a star that has been observed frequently over a long period of time, has been observed in at least the wavelength range over which the data was taken, and has a well known spectrum that is consistent each time it is observed. There are a multitude of standards stars used by astronomers; this is important because ideally standard stars should be taken at the same airmass as a corresponding science observation. Airmass is a measure of the amount of atmosphere the telescope is looking through. As a telescope points more towards the horizon the amount of atmosphere it has to look through increases. Consistent airmass is important for observations because particles in the air cause scattering of light, as does the turbulence of the atmosphere. This distorting of light can clearly be seen in the twinkling of stars. When observations are made at the same airmass, that just means that the angle of the telescope is approximately the same, the location in the sky is similar, and the time of the two observations are close together.

To actually calibrate the science data the data of the standard star are used to make a sensitivity function. This is done using the *sens* program in IRAF. The sensitivity function is ccd sensitivity as a function of wavelength. This is generated by comparing the actual observed standard star data with the known standard star data. It is a function as opposed to a scalar multiple because the detector functions over a wavelength band, so the sensitivity varies with wavelength, as seen in Fig. 4. The sensitivity function is then applied to the science data to generate the final reduced data, an example of which can be seen in Fig. 8.

## IV. RESULTS

My role in this project did not involve analysis of the data. This is primarily because we did not have all of the data necessary to see a peak in the light curves of the AGN I was working on. As such, this results section will be looking into the efficacy of my data reduction pipeline.

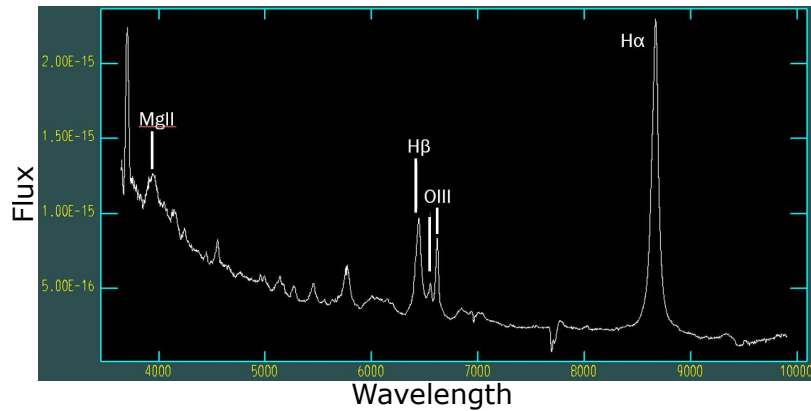


FIG. 8. Flux calibrated spectrum of AGN1435. The narrow band OIII emission lines are designated as are several broad line emissions: H $\alpha$ , H $\beta$ , and MgII. The x axis is wavelength in Angstroms, and the y axis is flux in units of ergs per second cm<sup>2</sup> Hz.

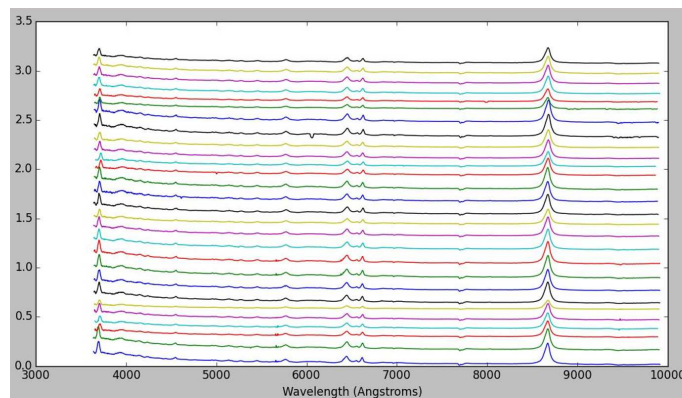


FIG. 9. All reduced spectra for AGN1435. The x axis is wavelength. Each spectrum is displaced on the y axis in chronological order. The bottommost spectrum representing the first night of observations, and the topmost spectrum representing the most recent night of observations. The H $\alpha$  line is the most prominent feature around 8750 Angstroms. Around 6500 Angstroms the H $\beta$  and the OIII lines can be observed.

The first note to make about the data reduction pipeline is the impact on time of reduction. My first pass of manually reducing data for a single night took me the entirety of a day. After having done the reductions multiple times, to figure out how to best optimize function parameters, I got reduction time down to around 15 minutes for a night of data. For the 15 nights of data I have, it would take close to 4 hours to do the reductions by hand. The reduction pipeline I wrote is able to reduce the data for all 15 nights in less than 5 minutes. This is an increase in efficiency by more than a factor of 45. Obviously, the pipeline optimizes reduction time.

The next aspect of comparison is the quality of the reduced data. The pipeline actually produced more reliable data than the manual reductions because it eliminated human bias from the reduction process. While doing reductions, I was in charge of setting parameters, and could modify them based on what looked to be the most accurate. The pipeline standardized the reduction parameters, taking the optimum parameters from the user and making slight modifications to them to best reduce

each night.

The pipeline is an improvement because it allows for more adaptability, for instance, dealing with the cosmic rays. It was previously mentioned that there are different techniques for dealing with cosmic rays, L.A. Cosmic and the built-in IRAF cleaning program. The pipeline has the functionality to run data sets with either cleaning method. The only requirement is changing a keyword in the source code. So two different cleaning techniques can be applied to all of the data in under 10 minutes. Doing this manually requires additions to the reduction process, at two different steps. Due to the timing inefficiency, it's only practical to compare a couple data sets before making a decision on which technique to use.

Regarding which cosmic ray rejection technique was better. The determination is still uncertain. Preliminary testing of the two indicates that L.A. Cosmic removes more of the cosmic rays. However, because of how L.A. Cosmic functions, it may alter the ratio of the flux of the narrow line emissions. This would make reverberation mapping unreliable.

A plot of all data for AGN1435 can be seen in Fig. 9. The data for AGN1435 are the only data shown because at this scale the spectra of AGN1438 and AGN1435 are nearly indistinguishable.

The only issue, currently experienced with both the manually reduced data and the pipeline reduced data, is an improper wavelength calibration. This most likely stems from a misidentification of features in the wave file.

## V. CONCLUSION

This project to construct a data reduction pipeline for AGN data from the VLT should be thought of as a success. A pipeline was written which can successfully reduce data with the same, or better, quality as manual reductions. The time required to reduce data has been cut by a factor of 45. Also, L.A. Cosmic shows promise as being a means to eliminate cosmic rays.

The next steps to undergo with this project are to more accurately identify emission features in the arc lamp used for wavelength calibration. This should eliminate the

wavelength shift seen in the data. Another step is to test the effects of L.A. Cosmic on the narrow line emission features. I have already started this process. I have started writing a Python program which integrates over the wavelengths encompassed by the narrow line features and records the total flux and peak in the feature. These measurements need to be made for both cosmic ray corrected data and non-corrected data and compared. If they are the same or off by a negligible degree then L.A. Cosmic will be deemed viable to use.

## ACKNOWLEDGMENTS

The author would like to thank Prof. S. Valenti for his tremendous guidance on this project, UC Davis for the providing the research opportunity, housing, and a stipend. The author would also like to thank Rena Zieve for coordinating the program. Lastly, the author would like to thank the other students participating in the research program for the camaraderie and support.

---

\* mbellar1@ithaca.edu

<sup>1</sup> B. Darvish, B. Mobasher, D. Sobral, A. Rettura, N. Scoville, A. Faisst, and P. Capak, arXiv preprint arXiv:1605.03182 (2016).

<sup>2</sup> D. Richstone, E. Ajhar, R. Bender, G. Bower, A. Dressler, S. Faber, A. Filippenko, K. Gebhardt, R. Green, L. Ho, *et al.*, arXiv preprint astro-ph/9810378 (1998).

<sup>3</sup> A. W. Graham, P. Erwin, N. Caon, and I. Trujillo, *The Astrophysical Journal Letters* **563**, L11 (2001).

<sup>4</sup> R. Blandford and C. McKee, *The Astrophysical Journal* **255**, 419 (1982).

<sup>5</sup> C. Impey, “Active galaxies,” (2012).

<sup>6</sup> C. K. Seyfert, *The Astrophysical Journal* **97**, 28 (1943).

<sup>7</sup> C. M. Urry and P. Padovani, *Publications of the Astronomical Society of the Pacific* **107**, 803 (1995).

<sup>8</sup> B. M. Peterson, *Publications of the Astronomical Society of the Pacific* **105**, 247 (1993).

<sup>9</sup> D. E. Osterbrock and G. J. Ferland, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*, 2nd ed. (Uni-

versity Science Books, 2006).

<sup>10</sup> H. Johnson and W. Morgan, *The Astrophysical Journal* **117**, 313 (1953).

<sup>11</sup> A. Cousins, *Memoirs of the Royal Astronomical Society* **81**, 25 (1976).

<sup>12</sup> T. A. Matthews and A. R. Sandage, *The Astrophysical Journal* **138**, 30 (1963).

<sup>13</sup> J. J. Brainerd, “Observational astronomy,” (2008).

<sup>14</sup> A. J. Barth, M. L. Nguyen, M. A. Malkan, A. V. Filippenko, W. Li, V. Gorjian, M. D. Joner, V. N. Bennert, J. Botyanszki, S. B. Cenko, *et al.*, *The Astrophysical Journal* **732**, 121 (2011).

<sup>15</sup> E. van Groningen and I. Wanders, *Publications of the Astronomical Society of the Pacific*, 700 (1992).

<sup>16</sup> V. Berezinskii, S. Bulanov, V. Ginzburg, V. Dogel, and V. Ptuskin, Moscow, Izdatel’stvo Nauka, 1984, 360 p. In Russian. **1** (1984).