Investigating the Effect of Active Galactic Nuclei on Their Host Galaxies

Micah J. Allard

Received November 13, 2024
Accepted August 29, 2025
Electronic access November 15, 2025

Active galactic nuclei (AGN) are a doorway for astronomers to understand the deep cosmos. This paper gathers evidence to better understand the correlation between AGN characteristics and their surrounding host galaxies. To investigate these correlations, three scientific papers from the past decade and 35 quasar-type Superspectra were analyzed in the form of fits files. The spectra were imported from the Sloan Data Sky Survey and four crucial emission lines were fitted for each of the 35 spectra using Gaussian functions to determine their amounts of hydrogen alpha and beta, nitrogen II, and doubly ionized oxygen (OIII). This investigation revealed that more massive AGN tend to correlate to an overall higher galactic mass, and that Intermediate Mass Black Holes are normally formed in the accretion disk of a Super Massive Black Hole (SMBH). As well, SMBH are not naturally occurring by the process of star life. Finally, increasingly larger AGN result in an increase in central kiloparsec intensity and a decrease in galactic activity radius. This paper provides insight into the intricacies of galactic activity, detailing the complex relationship AGN have with their host galaxies.

Introduction

In modern astronomy, an overwhelming amount of evidence suggests the presence of Super Massive Black Holes (SMBH) in the center of galaxies. High luminosity, compactness, time variability, and spectral energy distribution of AGN are best explained as effects of disk accretion into an SMBH. However, concrete evidence is needed to confirm that gas indeed accretes into a SMBH while inside the event horizon. It is, rather, commonly observed that gas outflows from AGN, causing massive relativistic jets traveling at extremely fast speeds in relation to the speed of light itself. This phenomenon coupled with the sheer amount of electromagnetic radiation it expels allows for astronomers to categorize the SMBH at the center of the galaxy and its corresponding accretion disk, narrow line region (NLR), broad line region (BLR), and relativistic jet as the active galactic nuclei: a nucleus usually only around 1 parsec in diameter which is more luminous than all of the surrounding kiloparsecs combined that form a host galaxy. In recent years, astronomers have posited a correlation between properties of black holes such as angular momentum and mass with the overall size, structure, and level of star formation in the host galaxy. The objective of this paper is to examine published documents and spectra to analyze and better understand the correlation between the characteristics of AGN and the evolution of their respective galaxies.

To answer this central question, this paper will consist of two major sections: a literature review and a data analysis. The literature review section covers the sources discussions on the modern understanding of AGN spectra, classification of AGN, and correlations between AGN and their host galaxies. This review also provides academic context to clarify the current understanding of the topic and a further discussion on the correlations that were drawn. The second section will be centered on data analysis with a step-by-step methodology, describing the process by which the results were synthesized.

Academic Context

Aside from dark matter, components of the known universe are present in AGN such as stars, gas, and dust which contribute to their spectra. This literature review and data analysis will focus on optical and broadband spectra as they are relevant tools that characterize AGN and can potentially explain a relationship between the properties of AGN and the evolution of their host galaxies. Below is a brief summary of the subject material that will be analyzed further in this work. This academic context was derived from reading the textbook, *An Introduction to Active Galaxies* (The Open University, 2016)[1].

Optical Spectra

Optical in this context refers to visible light. Optical spectra for galaxies typically consist of flux from stars and emissions from ionized gas clouds. Wavelengths of the optical spectra measure roughly between 380 and 700 nm. Elliptical galaxies have smaller hydrogen II (HII) regions and have an optical spectrum which resembles a stars spectrum but with fainter absorption lines. A spiral galaxy has both stars and star-forming regions, and its optical spectrum is the composite of its stars and its HII

© The National High School Journal of Science 2025

regions which show weaker emissions. Dust, being too cool to appear in optical spectra, often attenuates and dims galactic spectra. Optical light absorption lines also reveal chemical information about the host galaxy. In Figure 1, (below) the optical spectra only covers wavelengths specific to the visible portion of the electromagnetic spectrum. From this spectra, there are unique chemical signatures that can be recognized as each element or compound has a different fingerprint at a specific wavelength. This will be heavily utilized in the later data analysis.

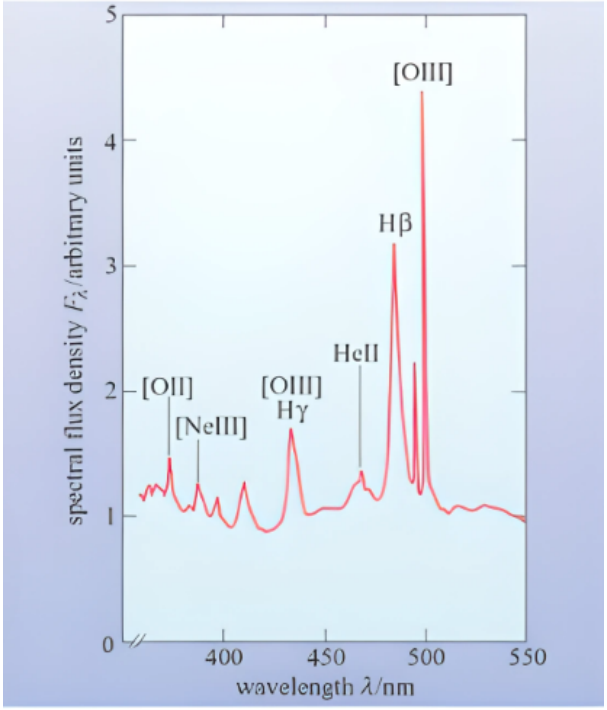

Fig. 1 Example of optical spectra measured by light absorption detailing an active galaxy - Introduction to Active Galaxies (Open University) [1]

Broadband Spectra

Broadband spectra fully encompass the electromagnetic spectrum and are often plotted using a logarithmic equation. The logarithms are often derived from wavelength measurements. Incorporating all forms of the electromagnetic spectrum into the characterization of a galaxy provides more information about the galactic nucleis effect on the host galaxy as well.
Broadband spectroscopy will be essential in the data analysis portion when using a BPT diagram. Broadband spectra typically compare the ratios of emission fluxes.

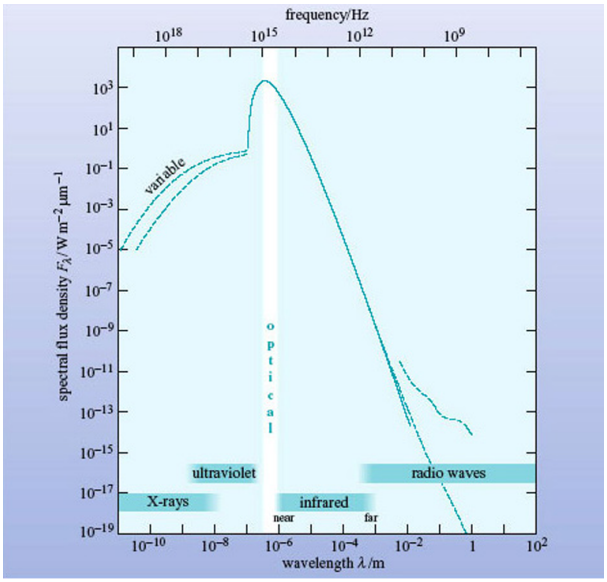

Fig. 2 Example of Broadband Spectra - Introduction to Active Galaxies (Open University) [1]

AGN Classification

Classification of AGN by mass and electromagnetic emissions allows connections between AGN and their host galaxies to be better understood. The largest subclasses of active galaxies are Seyfert and quasar galaxies which differ mainly in the radiation they emit. In a typical Seyfert galaxy, the nuclear source emits at visible wavelengths an amount of radiation comparable to that of the whole galaxy's constituent stars. In a quasar, however, the nuclear source is brighter than the constituent stars by at least a factor of 100. Seyfert galaxies are typically closer to Earth than quasars which in addition to their lower overall luminosity makes them harder to discover from more than several million parsecs away.

Seyfert Galaxies

Seyfert galaxies (Seyfert Type 1, Type 2, and LINERS) show an excess of radiation in the far infrared spectrum. Even more remarkably, at some wavelengths, including optical, this excess radiation is variable. Seyfert galaxies are also subclassified into two major types based on the spectra of their broad emission lines. Type 1 Seyferts have two sets of emission line regions, narrow and broad, whereas Type 2 Seyferts have only a noticeable narrow region. This indicates that Type 1 Seyferts have more mass and likely a corresponding higher angular velocity. Figure 3 compares the two types of Seyfert galaxies. There is a noticeable difference between the two where the spectra of the first graph, a Seyfert Type 1 galaxy, is heavier in shorter wavelengths at a higher spectral density. This seems to be consistent with previous ideas that Type 1 Seyferts are more massive and

© The National High School Journal of Science 2025

are rotating relatively faster than Type 2s.

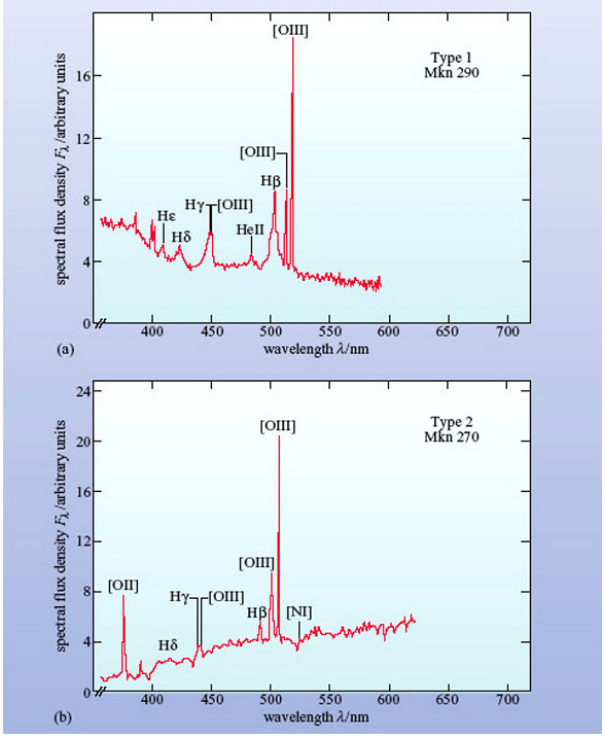

Fig. 3 Seyfert Type 1 contrasted with Seyfert Type 2 (optical spectra) - Introduction to Active Galaxies (Open University) [1]

Quasar galaxies (Quasars, Blazars, Radio Galaxies)

Thought to be stars when first discovered, quasars are the most luminous type of AGN. Quasars resemble distant Seyfert galaxies with variable luminous nuclei. About 10% of quasars are strong radio sources thought to be powered by jets of material moving at speeds close to the speed of light. Radio and blazar galaxies are two AGN archetypes similar to quasars. Radio galaxies are distinguished by having giant radio lobes fed by one or two emission jets. All three have compact nuclei like Seyfert Galaxies, but their nuclei are variable with broad or narrow emission lines. Blazars exhibit a continuous spectrum across a wide range of wavelengths and emission lines which, when present, are broad and weak.

Methodology

Review Paper

Search Strategy & Inclusion Criteria

In order to answer the central question on the existence of a correlation between AGN and the respective structure of their parent galaxy, the following three works were analyzed. Three

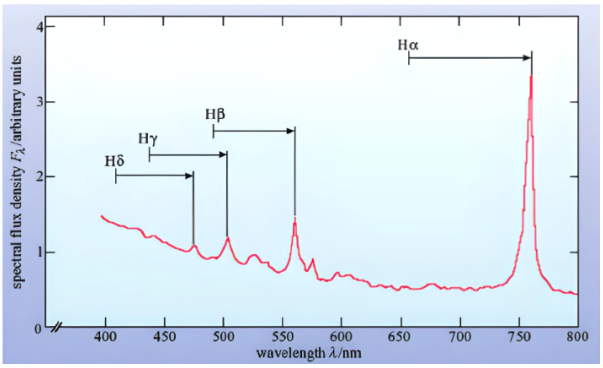

Fig. 4 Quasar galaxy optical spectra - Introduction to Active Galaxies (Open University) [1]

works were chosen for a pragmatic overview of the central question. The details of each work that regard the central question are documented in this review paper. Additionally, the following works were chosen because of their recency in astronomy and their high citation count.

Relations Between Central Black Hole Mass and Total Galaxy Stellar Mass in the Local Universe Reines et al. [2] studied 262 broad-line AGN in the nearby universe, using SDSS spectroscopy and searching for Seyfert-like narrow-line ratios and broad Hα emission, only using galaxies with a redshift z $<$ 0.055. The authors identify a correlation between SMBH and the size and total mass of the host galaxy. Investigation of this correlation assists in understanding the overall effect of AGN.

Intermediate Mass Black Holes in AGN Disks II. Model Predictions and Observational Constraints McKernan et al. [3] delves into the accretion and formation of Intermediate Black Holes (IMBH) inside the disk of a galactic nucleus. This paper lends insight on the formation of black holes of intermediate size and on the chemical compounds that formed them. It also provides an understanding of whether IMBHs contributed to galactic evolution.

Fueling Active Galactic Nuclei I. How the Global Characteristics of the Central Kiloparsec of Seyferts Differ from Quiescent Galaxies Hicks et al. [4] examines the inner parsecs (pc) of central SMBH, then raises a hypothesis about Seyfert galaxies being cold relative to other galaxies. This paper initially investigates the similarities and differences between inactive or quiescent galaxies and the most common type of AGN, a Seyfert galaxy. The paper then contrasts the two galaxy types within distances of 1, 50, 100, 200, and 500 parsecs. These three papers function as a good reference for investigating the correlation between AGN characteristics and galaxy evolution.

Information Extraction & Synthesis

To elaborate further on the contents of the analysis, the details of each work that regard the central question are documented in this review paper. This includes figures, equations, and discussions

that offer evidence of a correlation between Active Galactic Nuclei and their host galaxies. Takeaways will be discussed at the end of the paper analysis, compared with one another and correlations will be drawn if there are similarities between multiple works.

Paper Analysis

Relations Between Central Black Hole Mass and Total Galaxy Stellar Mass in the Local Universe Reines et al.[2]

Reines et al.[2] collected a sample of AGN emissions and then calculated the mass of black holes for comparison. A focus on other objects such as dwarf galaxies containing AGN,
reverberation-mapping of AGN, and galaxies with dynamically detected black holes form the remainder of this study. Reines et al.[2] constructed a sample of broad-line AGN by analyzing Sloan Digital Sky Survey (SDSS) spectra of approximately 67,000 emission-line galaxies, searching for objects exhibiting broad hydrogen alpha emissions and simultaneously comparing them to their hydrogen beta emissions. SDSS focuses on galaxies in the local universe: redshift z $\sim$ 0. Figure 5 below is a spectral analysis of the galaxies in the Sloan survey sampled in this paper.
With AGN emissions data, Reines et al.[2] shifted focus to black hole luminosity. The mass of the black hole is to be proportional to $\frac{[R\times(V2-V1)^2]}{G}$. It also states that the average gas velocity is inferred from the width of a broad emission line, usually looking at its levels of hydrogen β. Reines et al.[2] estimated black hole masses from their sample of broad-line AGN using the single-epoch virial mass estimator given by equation 5 in Reines et al.[2]. Equation 1 below was derived following the approach outlined in Greene & Ho (2005) for using the broad Hα line but incorporates the updated radius-luminosity relationship from Bentz et al. (2013).

$$\log\left(\frac{M_{\rm BH}}{M_\odot}\right) = \log\varepsilon + 6.57 + 0.47\log\left(\frac{L_{\rm H\alpha}}{10^{42}\ \rm erg\ s^{-1}}\right) + 2.06\log\left(\frac{\rm FWHM_{H\alpha}}{10^3\ \rm km\ s^{-1}}\right) \quad (1)$$

Reines et al.[2] utilized an equation to calculate the total stellar mass of the host galaxy without the AGN contribution, using this equation to compare the mass of the black hole and the host galaxy to uncover the desired correlation. The equation below removes the mass calculation of the AGN contribution. Host-only magnitudes (corrected for galactic reddening)
are used to estimate galaxy stellar masses with a color dependent mass-to-light ratio. Below is the dependent mass-to-light ratio from Zibetti et al. (2009).

$$\log\left(\frac{M}{L_i}\right) = 1.032\,(g-i) - 0.963 \quad (2)$$

Finally, Reines et al.[2] graphed the full width at half maximum (FWHM) of the galaxies broad hydrogen alpha (left panel), graphed their luminosity, and then calculated the distribution of said galaxies using the radius-luminosity relationship.
From these three models, Reines et al.[2] were able to produce a graph to showcase a line of best fit for the relation between the mass of the central SMBH compared to the stellar galaxy mass. This relationship is shown in the graph below using the sample from SDSS.
This graph communicates a line of best fit, suggesting a consistent relationship between the mass of a black hole and the mass of its surrounding galaxy. This shows that AGN with a higher mass are more likely to cause larger galaxies to evolve. This also infers that quasar-type galaxies (quasars, blazars, and radios) likely are more massive than Seyfert type galaxies (Seyfert 1, 2, and LINERS) and quiescent galaxies. From the collected data, however, there does not appear to be a correlation with dwarf galaxies as they are completely irregular and do not match well with the line of best fit given by Kormendy & Ho (2013).

Intermediate Mass Black Holes in AGN Disks II. Model Predictions and Observational Constraints McKernan et al.[3]

McKernan et al.[3] presents a study of Intermediate Mass Black Holes (IMBH). This work focuses on star activity in globular clusters which relates to a previous paper where the team mainly focused on building a model to summarize the accretion of IMBH in those clusters. This new paper focuses on IMBH formation caused by accretion in the disk of an AGN. Utilizing the Laser Interferometer Gravitational - Wave Observatory (LIGO) and the Laser Interferometer Space Antenna (LISA), they detected gravitational waves which they used to signify a black holes presence. The first several sections focus on establishing equations and explaining the formation of IMBHs in the protoplanetary disk of an AGN where star birth occurs. It seems that IMBHs forming in a protoplanetary disk tend to follow the same process as a
main sequence star, however, they gather enough mass at their inception to directly transform into an intermediate black hole. This indicates that more massive SMBHs present in AGN are primordial, not synthesized by natural processes. To understand differences between Seyfert and quiescent galaxies that can contribute insight into what mechanisms are fueling the Seyfert AGN, this team chose the observed sample to reduce any potential biases due to large-scale (>1 kpc) host galaxy properties. The remainder of this paper focuses on the characteristics of IMBH and their formation. The final graph is shown below:
This graph indicates that thermal emissions are not mass dependent. There is no line of best fit between the two characteristics, however, there does seem to be an emission limit. This graph shows that the objects are SMBHs considering the sheer mass from between 10^6 and 10^9 solar masses. This indicates that AGN mass is likely not attributable to thermal emissions.
Thus, for later data analysis, it is better to utilize an emission ratio delineated by the BPT diagram. The conclusion states that if

© The National High School Journal of Science 2025

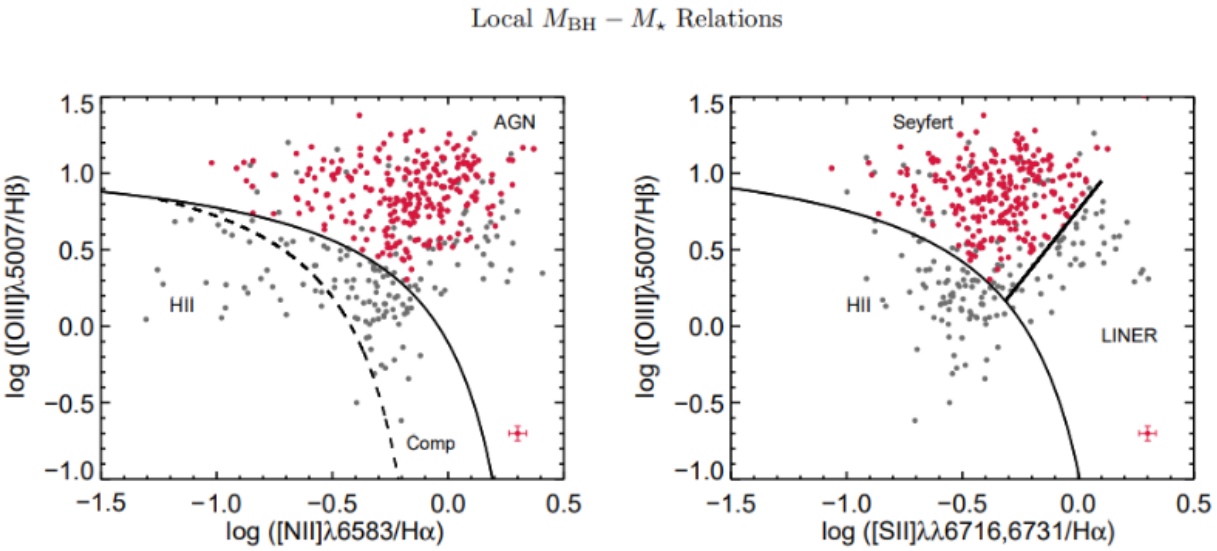

Fig. 5 Reines et al.[2] - Narrow-line diagnostic diagrams for sources with detectable broad Hα emission. These are produced using the classification scheme outlined in Kewley et al. (2006). The sample of broad-line AGN is restricted to objects with narrow line ratios placing them in both the AGN region of the [OIII]/Hβ vs. [NII]/Hα diagram and the Seyfert region of the [OIII]/Hβ vs. [SII]/Hα diagram (red points). The typical error for the red points is shown in the lower right corner of each plot.

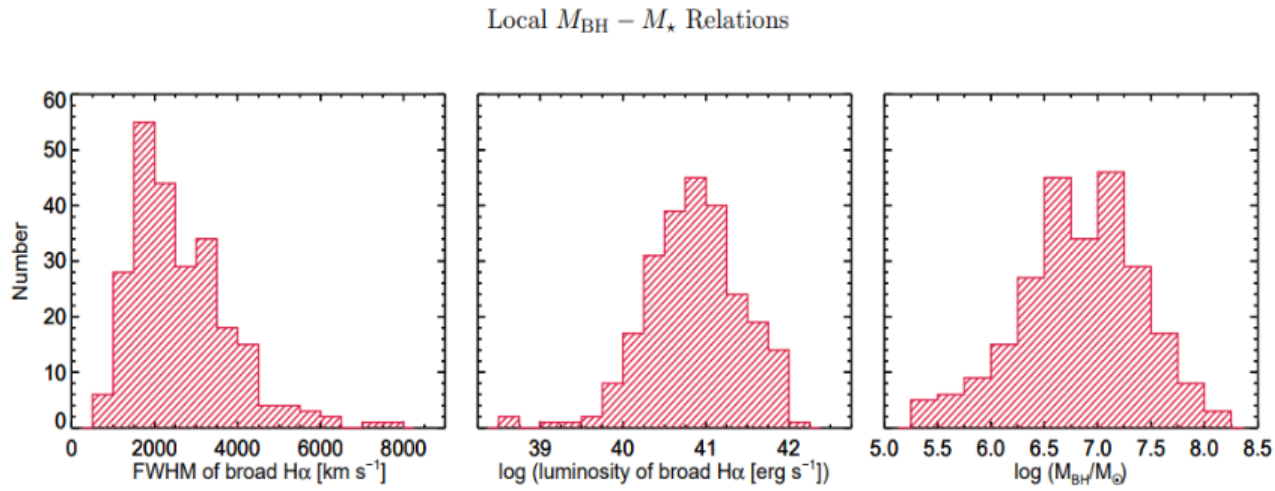

Fig. 6 Reines et al.[2] - Distribution of the FWHM (left panel) and luminosity (middle panel) of broad Hα emission for our sample of nearby broad-line AGN. The distribution of virial BH masses calculated from equation 1 is shown in the right panel.

IMBHs can form efficiently in AGN disks, the AGN host should exhibit a myriad of observational signatures. IMBHs that open gaps in AGN disks will exhibit strong observational parallels with gapped protoplanetary disks and may be detectable near mergers in the broad FeKα line. LINER activity may be due to weakly accreting MBH binaries in a large disk cavity. If IMBHs do not open a gap, detection depends on signatures of accretion onto the IMBH (including tidal disruption events).

Fueling Active Galactic Nuclei I. How the Global Characteristics of the Central Kiloparsec of Seyferts Differ from Quiescent Galaxies Hicks et al.[4]

The third paper details the current understanding of the AGNs role in galaxy evolution. Hicks et al.[4] states that it is widely accepted that BHs accumulate their mass through an AGN phase in which they accrete material from the host galaxy. Hicks et al.[4] finds that Seyfert AGN generally reside in host galaxies with a younger stellar population relative to that found in quiescent

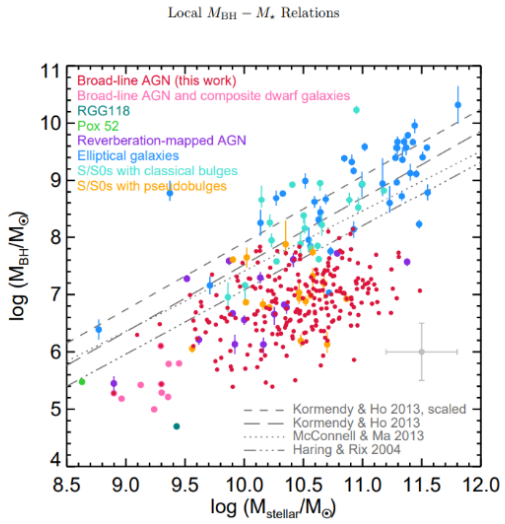

Fig. 7 Reines et al.[2] - Black hole mass versus total host galaxy stellar mass. All stellar masses are estimated using color dependent mass-to-light ratios presented in Zibetti et al. (2009)

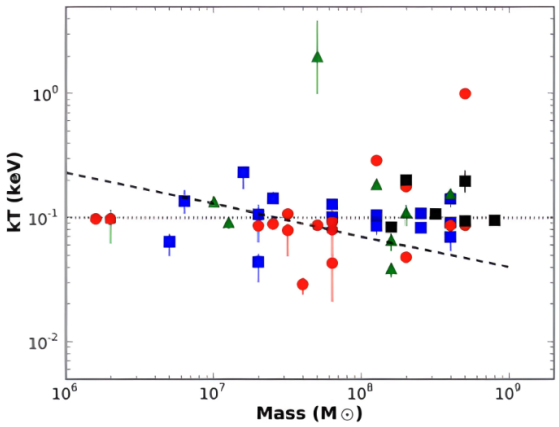

Fig. 8 McKernan et al.[3] - The peak energy (kT) of the soft X-ray excess found in AGN observed with Swift-BAT plotted against the mass of the central SMBH (Winter et al., 2012). Blue squares are Seyfert 1s, black squares are broad-line radio galaxies, green triangles are Seyfert 1.2s, and red circles are Seyfert 1.5s. The horizontal dotted line corresponds to a model of ~ 0.1 keV thermal emission due to bow-shocks from NCO/IMBH on retrograde orbits in the AGN disk, encountering "headwinds" centered on 100 km s^{-1}. The sloping dashed line corresponds to IMBH with a constant mass ratio of $q = 2 \times 10^{-4}$. Below this line, AGN disks with fiducial $H/r = 0.05$, $\alpha = 0.01$, are not expected to have soft excess due to accreting IMBH.

galaxies, confirming that an abundant fuel supply is available in these galaxies on scales of several kiloparsecs.

Hicks et al.[4] argues that by observing a sample of matched Seyfert-quiescent galaxy pairs and by constraining what processes are relevant to fueling local AGN, which are likely candidates for fueling at higher redshifts as well, they can identify properties unique to Seyferts that are potentially related to fueling of their AGN activity. Hicks et al.[4] also attempts to identify differences between Seyfert and quiescent galaxies that can provide insight into what mechanisms are fueling the Seyfert AGN. The observed sample was selected to reduce any potential biases due to large-scale (> 1 kpc) host galaxy properties.

Central kpc in Seyferts versus Quiescent Galaxies

Table 1
Galaxy Properties

I.D. (1)	Galaxy (2)	AGN (3)	T (4)	M_B (5)	v (6)	R_{25} (7)	$frac_{20}$ (8)	Nuc. Dust (9)
1a	NGC 3227	Sey 1.5	1.0	-20.4	1152	0.17	0.06	C
1q	IC 5267	no	0.0	-20.3	1713	0.13	0.06	C
2a	NGC 5643	Sey 2	5.0	-20.9	1199	0.06	0.07	GD
2q	NGC 4030	no	4.0	-20.9	1460	0.14	0.16	TW
3a	NGC 6300	Sey 2	3.0	-20.6	1110	0.18	0.07	C
3q	NGC 3368	no	2.0	-20.3	897	0.16	0.04	CS
4a	NGC 6814	Sey 1.5	4.0	-20.5	1563	0.03	0.11	GD
4q	NGC 628	no	5.0	-20.2	656	0.04	0.03	N
5a	NGC 7743	Sey 2	-1.0	-19.4	1658	0.07	0.11	LW
5q	NGC 357	no	0.0	-19.9	2541	0.14	0.14	N
Seyfert Subsample			2.4 ±2.4	-20.4 ±0.6	1336 ±254	0.10 ±0.07	0.05 ±0.02	
Quiescent Subsample			2.2 ±2.3	-20.3 ±0.4	1453 ±741	0.12 ±0.05	0.10 ±0.04	

Fig. 9 Hicks et al.[4] - The 10 Quasars reviewed in these papers and their unique base stats.

After presenting this table, Hicks et al.[4] goes on to analyze each galaxy. This analysis allows differences to be easily seen. Each AGN/quiescent galaxy is measured in intensity by flux, velocity, and dispersion. Shown here are the compared 2D maps of CO absorption for a quiescent (no galactic activity), Seyfert 1.5, and Seyfert 2 galaxy. From the samples in Figures 10 through 12 below, it is clear that in terms of flux, the more active an AGN becomes, the more its flux becomes dense. This is shown by the quiescent galaxy having a relatively spread-out flux with relative intensity to the extent of even 300 parsecs from the galactic nucleus while the more luminous Seyfert 2 galaxy has a much more concentrated flux. This is seen across all three graphs, rather than just the flux graph, and is supported by the figure at the end of Reines et al.[2] since a more dense AGN correlates with more central luminosity in the innermost kiloparsec, causing a disproportionate ratio between SMBH mass and stellar galactic mass. Another interesting feature is the velocity for both Seyfert galaxies. They both seem to be evenly split with one half being intense and the other relatively less intense. This seems counterintuitive as nuclei proximity normally correlates with galactic intensity which may be an area of future investigation.

Overall, Hicks et al.[4] portrays that the effect of an AGN on a host galaxy is strongly shown in the overall composition of the galaxy. More active or luminous AGN cause an increase in central intensity and a decrease in overall galactic radius.

© The National High School Journal of Science 2025

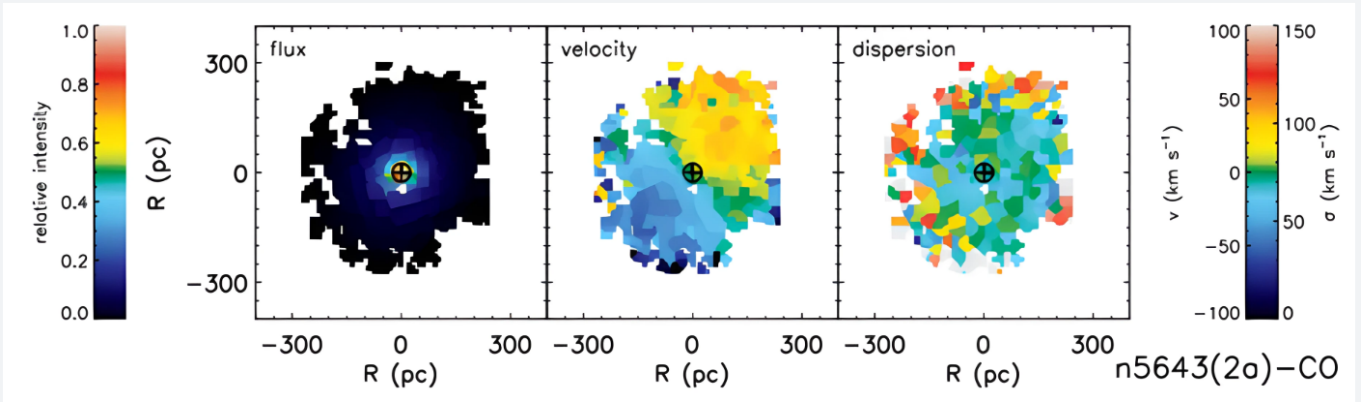

Fig. 10 Hicks et al.[4] - Seyfert 2 NGC 5643

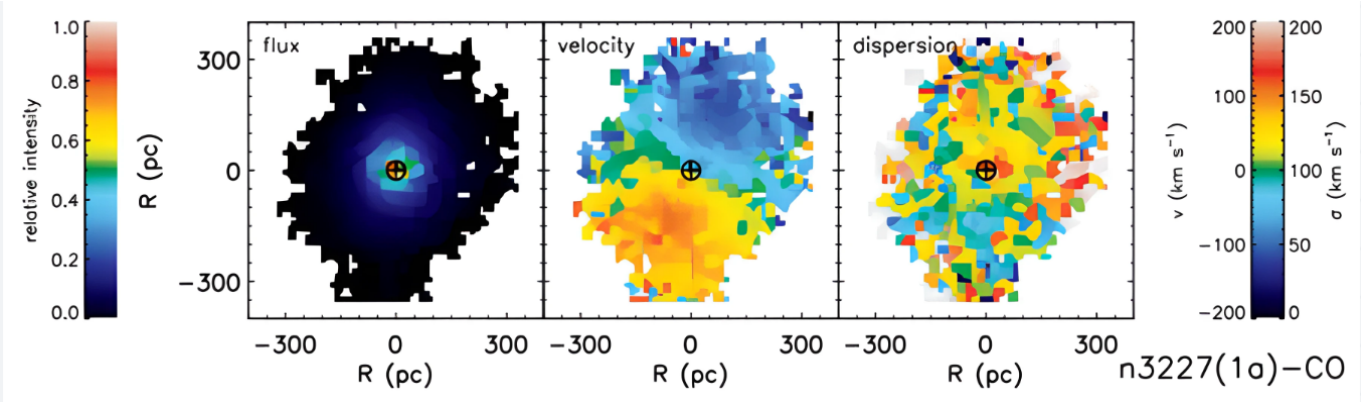

Fig. 11 Hicks et al.[4] - Seyfert 1.5 NGC 3227

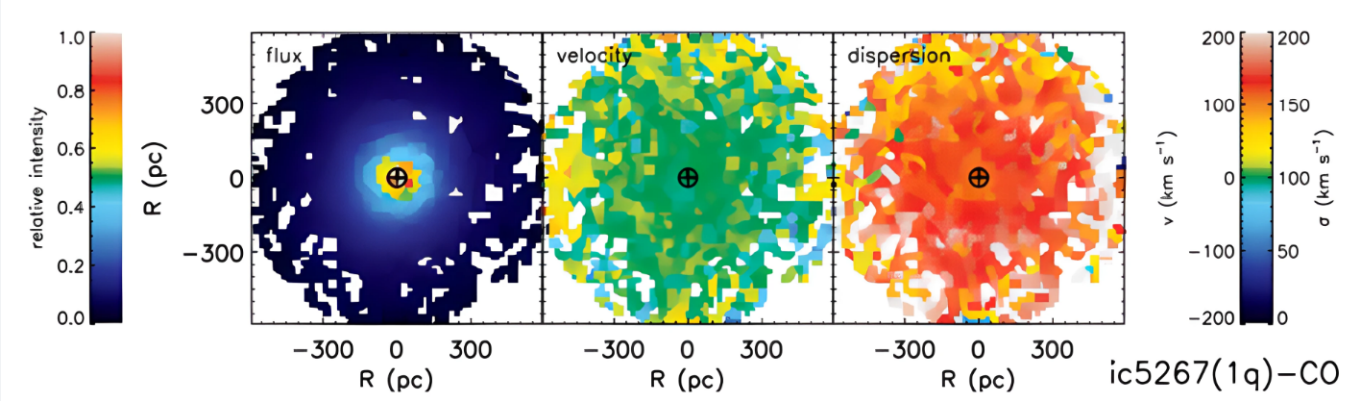

Fig. 12 Hicks et al.[4] - Quiescent galaxy IC 5267

Discussion

The graph at the end of Reines et al.[2] suggests that the more massive the AGN, typically, the larger the surrounding galaxy is in terms of stellar mass. This is with the exception of dwarf galaxies which do not seem to follow a line of best fit. Elliptical galaxies tend to have higher mass AGN as well as higher mass host galaxies. This backs up the notion that a more spherical shape is caused by higher mass. Finally, as expected, the paper states that quasar galaxies which have more luminous AGN will have a higher surrounding galactic mass.

McKernan et al.[3] suggests that except for some outliers, AGN have a thermal emission limit regardless of mass. It also states that an SMBHs accretion disk will lead to the creation of IMBH.

By extrapolating on this notion, this may be indirect evidence that SMBHs, which all AGN are centered around, are primordial or the result of a massive IMBH merging event. This is also supported by Celotti et al.[5] which states that massive star death creates stellar mass black holes as opposed to SMBHs.
Hicks et. al[4] indicates that galactic nucleus activity feeds central SMBH growth, meaning that in the context of the ideas established in the first paper, active galaxies cause overall galaxy growth. Next, they suggest that Seyfert or other AGN typically reside in younger galaxies where there is enough fuel for them to become active. This means that a large amount of primordial AGN have lost their fuel and become quiescent (possibly including our Milky Way). The paper makes the following assertions: Seyfert AGN can feed on stellar fuel within a radius of several kiloparsecs, and quasars are likely caused by galactic mergers. This contradicts the preceding notion of primordial AGN, as Seyfert AGN are known to be less massive than a quasar yet have a gravitational pull strong enough to feed off several kiloparsecs of surrounding material, indicating that a quasar should also have enough gravity to accumulate its mass on its own or with the help of a merger. The figures at the end of Hicks et. al[4] demonstrates that AGNs effects on host galaxies are strongly evident in the galaxies overall composition.
Specifically, more active or luminous AGN contribute to an increase in central intensity and a decrease in overall galactic radius.
The relationship between AGN and their host galaxies investigated in these three papers assist in lending context to the subsequent data analysis.

Research Paper

Design & Selection Criteria

In order to determine AGNs effects on their parent galaxies, the SDSS catalog was used to download the fits files of AGN. Using this catalog, over 35 fits files of quasar AGN will be graphed and fitted using Jupyter Notebook. Forty files were chosen specifically for a pragmatic, yet statistically robust analysis. This will help to classify these AGN by plotting their emissions on a BPT diagram. Raw spectra from the Sloan Data Sky Survey (SDSS) Data Release 16 (DR16) were downloaded for analysis. A reference BPT Diagram is plotted below.
Next, quasar-type AGN in a redshift </= 0.03 z, specifically imported from SDSS, were chosen to focus on AGN that are relatively local in the universe (0.03z $\sim$ 414 million light years) in order to get spectra that are more precise. The specific wavelengths for the hydrogen alpha, hydrogen beta, oxygen III, and nitrogen II were utilized. Below are the given wavelengths of these chemical compounds.
To determine their abundance, flux will be graphed and measured. This can be done with a Jupyter Notebook tutorial on plotting galactic emission lines and producing a fitted curve. A

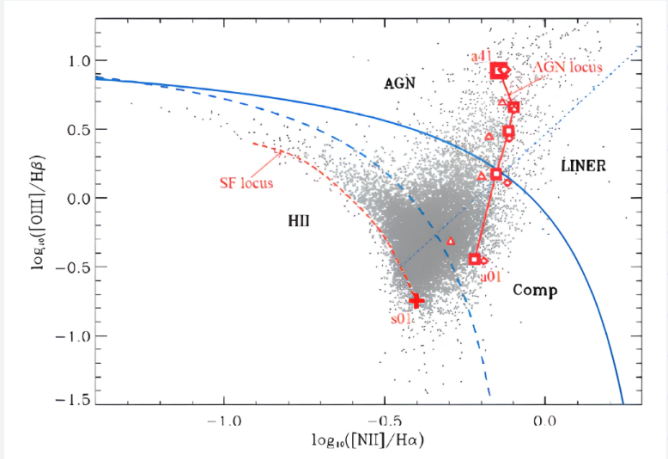

Fig. 13 Richardson et al.[6] - BPT Diagram

tutorial by Luke Holden, *Measuring Emission Line Fluxes with Python*[7], can be utilized to fit the curves and find the measurements. The commonly used threshold for wavelength significance is 3:1 as stated by Iowa States Center for Nondestructive Evaluation. So, the amplitude of a wavelength for the above chemical compounds must have a peak of at least 3 times the flux of the surrounding average or noise. This would correspond to a signal-to-noise ratio of about 3:1. These measurements are determined by a Gaussian function and will be represented in a normal distribution. In the Gaussian Function below, Equation 3, a is the amplitude of the
emission peak, xa is the wavelength of the amplitude, σ is the standard deviation of the distribution, x is the wavelength value in lambda, and e represents Eulers number.

$$f(x) = ae^{-\frac{(x-x_a)^2}{2\sigma^2}} \tag{3}$$

With this ratio measured, each individual ratio will be compared to the emissions of other chemical compounds and then plotted on the BPT diagram to fit onto the axis shown by Figure 13. By plotting these quasar spectra on the BPT diagrams and measuring their respective chemical abundance, the classification of galaxies from data release 16 of the optical SDSS will be attempted on at least 35 galaxy spectra. This equation is utilized in the code of the above python tutorial. Based on deductions from the literature review about AGN and their characteristics combined with data on the chemical composition of quasar AGN, a general correlation should be able to be made between the qualities of an AGN and the evolution and composition of the host galaxy.

Revisions

AstroimageJ was used to extract flux measurements from the fitz files when the above tutorial proved incompatible with the data. It was also used to extract log-lambda and flux values from the

© The National High School Journal of Science 2025

Table 1 Chemical compounds necessary for BPT plotting and their wavelengths, including oxygen III, 2 lines of nitrogen II, hydrogen alpha, and hydrogen beta.

	Hydrogen β	**Oxygen III**	**Nitrogen II**	**Hydrogen α**	**Nitrogen II**
Wavelength (Å)	4861.333	5006.843	6548.050	6562.819	6583.460

35 downloaded fitz files off of the SDSS database. AstroimageJ is an extension of ImageJ which provides an astronomy-specific image display environment and tools for astronomy specific image calibration and data reduction. Because of this, it was easily able to interpret the fits files from SDSS. The log-lambda values extracted are the solution to the logarithmic equation below where $\log\lambda$ is given by the AstroimageJ extraction.

$$f(x) = \log_{10}\lambda = \log\lambda \tag{4}$$

This extraction gave the necessary log lambda and flux measurements for the data analysis. A single example of the measurements is below in Figure 15.

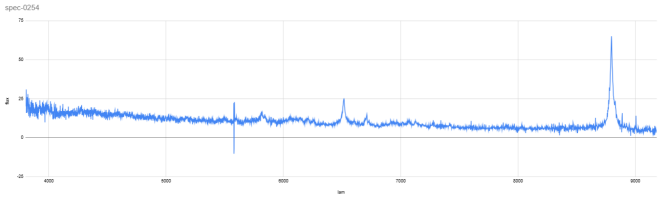

Fig. 14 An example of a measurement and flux given by AstroImageJ (spec-0003) see Table 1 in the analysis results. AstroImageJ typically produced anywhere between 3835 and 3850 unique λ values, each with specific flux variables. Chemical emissions were measured with an error bound of ± 50 from the given wavelengths in Figure 14. This error bound was chosen to fully encompass the emissions of the specific compound. It was selected after examining the graphed spectra and determining the average difference between the two ends of each curve.

Upon obtaining the log lambda values, the actual lambda value could be obtained using exponentiation of the set with a base of 10. It is this lambda vs. flux that is graphed above.
Shown below is an example of a hydrogen alpha/nitrogen II emission which is significant enough to be measured. (Noticeable signal:noise ratios where the signal is the curve amplitude and the noise is the spectral background.)
The emissions for this wavelength can be calculated using an equation with the area between two curves and the Fundamental Theorem of Calculus. Equation 5 below was derived during this data analysis to explicitly measure the emissions of a specific chemical and the signal-to-noise ratio necessary to plot on the BPT Diagram.

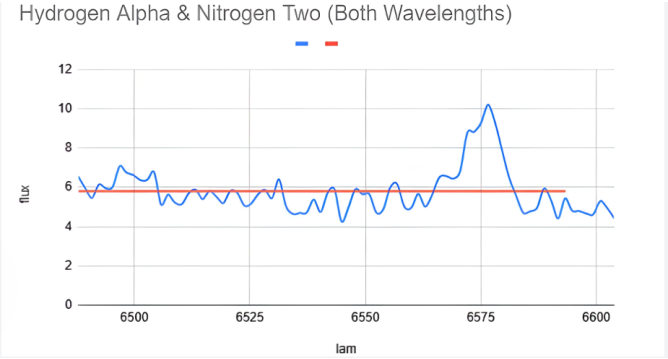

Fig. 15 Nitrogen II and hydrogen alpha emission curve for spec-0165. See Table 1 in Results. The red line is the flux average in the above data, and the blue line is the actual flux of this spectra.

$$\frac{\int_{\lambda(a)}^{\lambda(b)} ke^{-\frac{(x-n)^2}{2\sigma^2}}\,dx}{\int_{\lambda(a)}^{\lambda(b)} \mu\,dx} = \frac{S(\text{signal})}{N(\text{Noise})} \tag{5}$$

This equation was utilized to find the signal-to-noise ratio where S(total) represents the total emission of the wavelength, k the amplitude of the Gaussian function (which represents the peak of the emission), h the wavelength at which this peak takes place, and σ the standard deviation of the $+/-$ 50 lambdas corresponding flux values. The lambda bounds are given when the peak curve exceeds the μ (average) flux value of the specific emission. The μ is determined by taking the average flux of the surrounding 50 wavelength fluxes on either side in order to remove the possibility of an inaccurate equation due to gradual background noise over a massive amount of lambda values. An example of this is shown below.

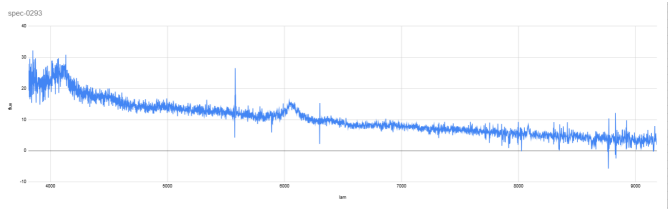

Fig. 16 In this example spec 0293 a total μ value of the entire spectra wouldnt make sense as the μ of the entire spectra is likely to be higher than even the peaks at λ = 9000

Anticipated Error

Most errors occurring from the usage of these equations are from the approximations used in the input variables. Integrating the total over the base gives a precise measurement for the signal-to-noise ratio. For instance, selecting +/- 50 wavelength for the integration of the emission is an approximate average $(+/-)$ which was taken based on the start and end of the emission across the thirty-five spectra that were graphed. It is probable that curves could start or end at very slightly smaller or higher values than the bounds, however, the bounds cannot be customized to the curve in order to preserve the integrity of the normal distribution equation and its essential symmetry. Every other step in this procedure such as calculating the standard deviation of the values is in the +/- 50 range which may lead to the same type of approximation error.

Finally, the noise ratios vary based on the average flux of the emission. For instance, a low average emission can cause an extremely high ratio as integration is based on the x-axis. In the event of a negative flux average (should not realistically be possible), the equation would fail entirely as there cannot be a negative ratio. (*When referring to average flux of emission, it is the red line in Figure 15.) There is a probability that a negative ratio could occur if the signal is so weak that an offset in the spectra, potentially accounting for background noise from Earth, causes a negative flux.

Method Test

Below is a graph and equation bar of sample spectra, spec-0165, to test the derived equation. (See Table 1 in Results.)

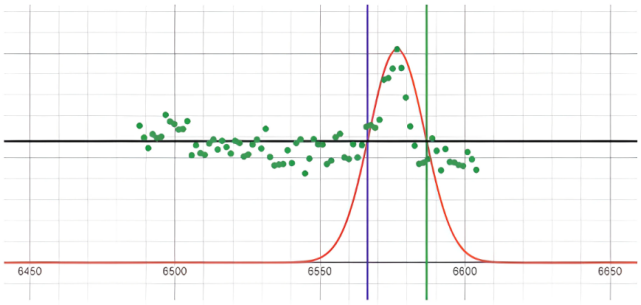

Fig. 17 Visualization of the equation above on the Desmos Graphing Calculator. The scattered points represent the NII and hydrogen alpha emissions of spec-0165. (See Table 1 in Results.) The green and blue lines represent a and b coordinates for integration. These a and b coordinates give an accurate depiction of the emission of the curve, focusing on when the curve exceeds the average/noise of the sample represented by the black line. This figure confirms that the red line given in Equation 3 correctly fits/represents the data.

This is very much the expected result from a visual standpoint on the graph. This method repeatedly proved successful when measuring abundant emissions. The remaining data was analyzed within these new parameters.

Fig. 18 Equation tab on Desmos Graphing Calculator for emission measurements. The Gaussian function of the distribution is line 1, the total emission is line 2, and the additional signal minus the background noise is line 3. In this case, line 3 also represents the area between two curves given in Figure 17. These emissions are compared to the flux over the interval in which the emission curve exceeds the sample average within a wavelength $+/-$ 50 lambda. Line 4 is the S/N ratio needed for BPT plotting given by Equation 4. The solution given in line four is the correct signal-to-noise ratio for spec-0165, representing the peak-to-average shown in Figure **??**.

Results

The analyzed sample of AGN spectra consisted of 35 quasar spectra measured with AstroimageJ to analyze their flux and wavelengths. AstroimageJ measured around ~3850 wavelength plottable values per quasar spectra. Table 2 below is the documentation of all 35
spectra. All calculation, and emission plots are within $+/-$ 50 lambda of their corresponding wavelengths lambda values given in Figure 14. Hydrogen alpha and Nitrogen II emission measurements were grouped because their approximate wavelengths were within 50 lambda of each other. Below is a table which shows the measured ratios for all 35 samples downloaded from SDSS.

© The National High School Journal of Science 2025

	Spectra #	Hydrogen B	Oxygen III	Nitrogen II / Hydrogen α
		S/N Ratio - (3 for significance)		
1	spec-0003	1.27109818	1.32284337	1.425998781
2	spec-0005	1.420187389	1.263752442	1.261522089
3	spec-0009	1.429790928	1.613624203	1.20711661
4	spec-0042	249.7970059	Underfined	2.091338594
5	spec-0050	4.958226103	8.765979067	1.569591013
6	spec-0052	1.405429416	1.226879318	1.152225055
7	spec-0074	1.266519452	1.17450689	1.212239852
8	spec-0124	1.135891876	1.179049962	1.216363488
9	spec-0133	1.279075477	1.193172716	1.204746823
10	spec-0165	1.202317816	1.186390589	1.866873277
11	spec-0184	1.258676072	1.317365227	1.221471678
12	spec-0190	1.145324793	1.169938761	1.214501388
13	spec-0216	1.141316882	1.14415617	1.190115449
14	spec-0222	1.219416925	1.167195997	1.11997005
15	spec-0232	1.156653172	1.19803447	1.220289033
16	spec -0254	1.142826366	1.132813346	1.623636363
17	spec - 0267	1.126406995	1.145040706	1.140609665
18	spec-0281	1.147072944	1.097206816	1.133515584
19	spec-0293	1.134118829	1.125547755	1.176186798
20	spec-0304	1.180876154	1.171283668	1.176186798
21	spec-0325	1.293076253	1.22709742	1.262135917
22	spec-0330	1.343450297	1.410599153	1.305034045
23	spec-0352	1.321256783	1.237051796	1.336707278
24	spec-0360	1.258027403	1.259255	1.284435415
25	spec-0372	1.803711873	1.848968301	2.371660536
26	spec-0385	1.170778371	1.181572502	1.138024208
27	spec-0388	1.26906944	1.396352807	1.261394083
28	spec-0495	1.13537252	1.137176182	1.073955882
29	spec-0502	1.289963786	1.198470924	1.168382287
30	spec-0518	1.133860616	1.13935793	1.116383508
31	spec-0565	1.521180769	1.408323283	1.48006623
32	spec-0579	1.2782217	1.262641153	1.369586479
33	spec-0608	2.858008757	3.684269521	2.453035409
34	spec-0620	1.217722796	1.26024905	1.208901963
35	spec-0635	1.096844724	1.107604816	1.112558596

Fig. 19 Each number represents the emission ratio for all 35 spectra. There needs to be a ratio of 3 or more to hold significance.

Conclusion

Although several spectra had emissions that were measured, not a single source had the abundance in all four compounds necessary to be plotted on the BPT diagram. However, it is important to note that some emissions did meet the standard for significance. It is also important to acknowledge the errors anticipated in the previous section. For the latter, spec-0042 is an excellent example of a very weak spectra. The average flux for the Oxygen III emission was negative, causing an undefined answer. Similarly, for its Hydrogen Beta emission, there were so many negative values that the average flux neared zero, and the signal ratio ended up being 249 times larger than the average despite the lowest (negative) fluxes having an absolute value larger than the amplitude. The measurements from spec-0050 are a similar scenario with the sample $+/-$ 50 lambda containing negative flux values. The Oxygen III emission in spec-0608 was the only emission which met the standard for AGN significance without negative values in its emission. Spec-0372 and spec-0565 also have noticeably higher ratios compared to the others. It is possible that a ratio of three is too high a standard for emission significance or that enough spectra were analyzed. Finally, spectra had emission abundance at other lambda values. The below spectra is an example of an abundance of chemical compounds other than the ones needed for BPT plotting:

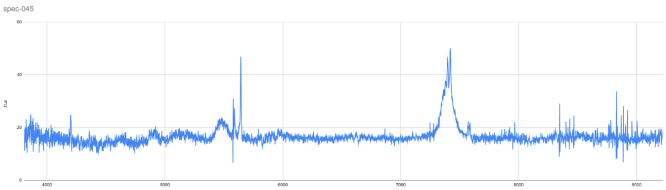

Fig. 20 A spectra from the SDSS sample (spec-0495). The massive fluxes between 7000 and 8000 lambda may be grounds for further research as these chemicals may further explain optical characteristics of the specific AGN.

Nevertheless, since none of the sources exhibited all four emission lines - hydrogen alpha, hydrogen beta, nitrogen II, and oxygen III - they could not be plotted on the BPT diagram. Consequently, this limitation prevented drawing conclusive evidence from the data analysis to support the claims discussed. However, examination of these spectra reveals numerous avenues for future research.

Next Steps

Further steps could involve processing more AGN spectral data in order to make the BPT diagram or using an already processed database which holds candidate galaxies with the correct chemical abundance for analysis. For example, rather than using raw data from SDSS, other catalogs such as the MUSE Wide Survey could be utilized which hold AGN spectra that have emission lines with the correct abundance. SDSS was simply used for the opportunity to learn
with raw data. The MUSE Wide Survey could be used as select sources for the BPT diagram. It is a promising database for further analysis. As well, although the intent of the data analysis was unfulfilled by the lack of specific chemical abundance, the 35 AGN spectra plotted in Table 1 seemed to have an abundance in other wavelengths. This could serve as a breadcrumb for a future analysis.

Acknowledgements

First, I would like to thank my father and mother who supported this research opportunity. Secondly, I would like to thank my mentor, Dr. Marziye Jafariyazani, who did an excellent job in helping me understand material from various textbooks to prepare for the writing of this paper. Dr. Jafariyazani never failed

to make difficult material easy, and she guided me through the conceptualization, formatting, and referencing for this paper. I would also like to thank her for the opportunity I had to analyze data with Python. I am grateful to AstroImageJ, Jupyter Notebooks, Google Suite, and YouTube programming tutorials for their incredible open source platforms.

I would also like to extend my appreciation to my teachers at Nashua High School South, including Mr. Jeffrey Durnford from AP Statistics, Ms. Kellie Gabriel from AP Calculus, and Mr. Anthony Spero from AP Computer Science, who helped me troubleshoot the mathematics and technical errors with the software. I would like to thank Lumiere Education who provided me with the opportunity to work with a researcher in my desired sector of astronomy. The Lumiere program assisted in editing the scientific writing, preparing for publication, and adhering to a timeline. As a senior in high school, I am grateful to have worked on AGN at my stage of learning, and I look forward to further research in the fields of astronomy, astrophysics, and mathematics during undergraduate studies, graduate school, and in my future career.

As a high school student, it was a privilege to undertake this project. Coming into this, I hoped to learn more about working with black holes and the genesis of galaxies. I had previously worked with exoplanets and their lightcurves and transits, but this paper delved more into the physics of astronomical phenomena in which I am keenly interested. Outside of astronomy, I learned more about the scientific method and the expertise needed to write a scholarly paper. Through this experience, I learned that hypotheses do not always end in hoped for results, that failure is a necessary part of scientific inquiry, and that resilience is key to finding alternative solutions. Regardless of what comes next, the insight I gained from this paper will be invaluable to my future research in astrophysics.

References

1 The Open University, *An Introduction to Active Galaxies*, 2016.

2 A. E. Reines and M. Volonteri, *Astrophysical Journal*, 2015, **813**, 82.

3 B. McKernan, K. E. S. Ford, B. Kocsis, W. Lyra and L. M. Winter, *Monthly Notices of the Royal Astronomical Society*, 2014, **441**, 900.

4 E. K. Hicks, R. I. Davies, W. Maciejewski, E. Emsellem, M. A. Malkan, G. Dumas, F. Müller-Sánchez and A. Rivers, *The Astrophysical Journal*, 2013, **768**, 107.

5 A. Celotti, J. C. Miller and D. W. Sciama, *Classical and Quantum Gravity*, 1999, **16**, A3.

6 C. T. Richardson, J. T. Allen, J. A. Baldwin, P. C. Hewett and G. J. Ferland, *Monthly Notices of the Royal Astronomical Society*, 2014, **437**, 2376–2403.

7 L. Holden, *Tutorial – Measuring Emission Line Fluxes with Python and Deriving Gas Properties*, 2023, `https://lukeholden.com/blog/measuring_emission_line_fluxes_python.html`, Accessed: 30 Jan 2023.

 © The National High School Journal of Science 2025