Active Galactic Nuclei and the Correlated Properties of Neighboring Galaxies

Research Thesis

Presented in partial fulfillment of the requirements for graduation with research distinction in Astronomy and Astrophysics in the undergraduate colleges of The Ohio State University.

by

Robert Von Holle

The Ohio State University
April 2021

Project Advisor: Professor Barbara Ryden, Department of Astronomy

Abstract

When a galaxy's central supermassive black hole is actively accreting matter, it becomes an active galactic nucleus (AGN). AGN are most commonly found in the "green valley", a region of relatively low density on a color/magnitude diagram of galaxies located between the denser red and blue sequences. I performed an analysis on a sample of galaxies collected from the Sloan Digital Sky Survey (SDSS) database in search of the properties possessed by a target galaxy's neighboring galaxies that correlate with the probability of finding an AGN in that target. I found that, among galaxies with AGN, galaxies with more neighbors – galaxies within a projected distance of 0.5 megaparsecs (Mpc) – tend to be redder than galaxies with few neighbors for the majority of the color range in my sample; however, among galaxies without AGN, this trend only exists outside of the green valley. Additionally, I found that galaxies with fewer neighbors are more likely to contain AGN than galaxies with many neighbors.

1. Introduction

Some galaxies feature compact central regions that are much more luminous at certain wavelengths than normal, with this increased luminosity not attributable to an increased density of stars. The central regions of these galaxies are known as Active Galactic Nuclei (AGN). Figure 1 shows the spectra of several galaxies belonging to a variety of classes of AGN compared to the spectrum of a "normal" galaxy with an inactive nucleus. The several active galaxies show clear emission lines that are not present in the spectra of the inactive galaxy. An AGN is understood to be the result of the central supermassive black hole of its host galaxy actively accreting matter. This accreted matter heats up, resulting in the characteristic emission lines of the AGN[1].

[1] Peterson (1997)

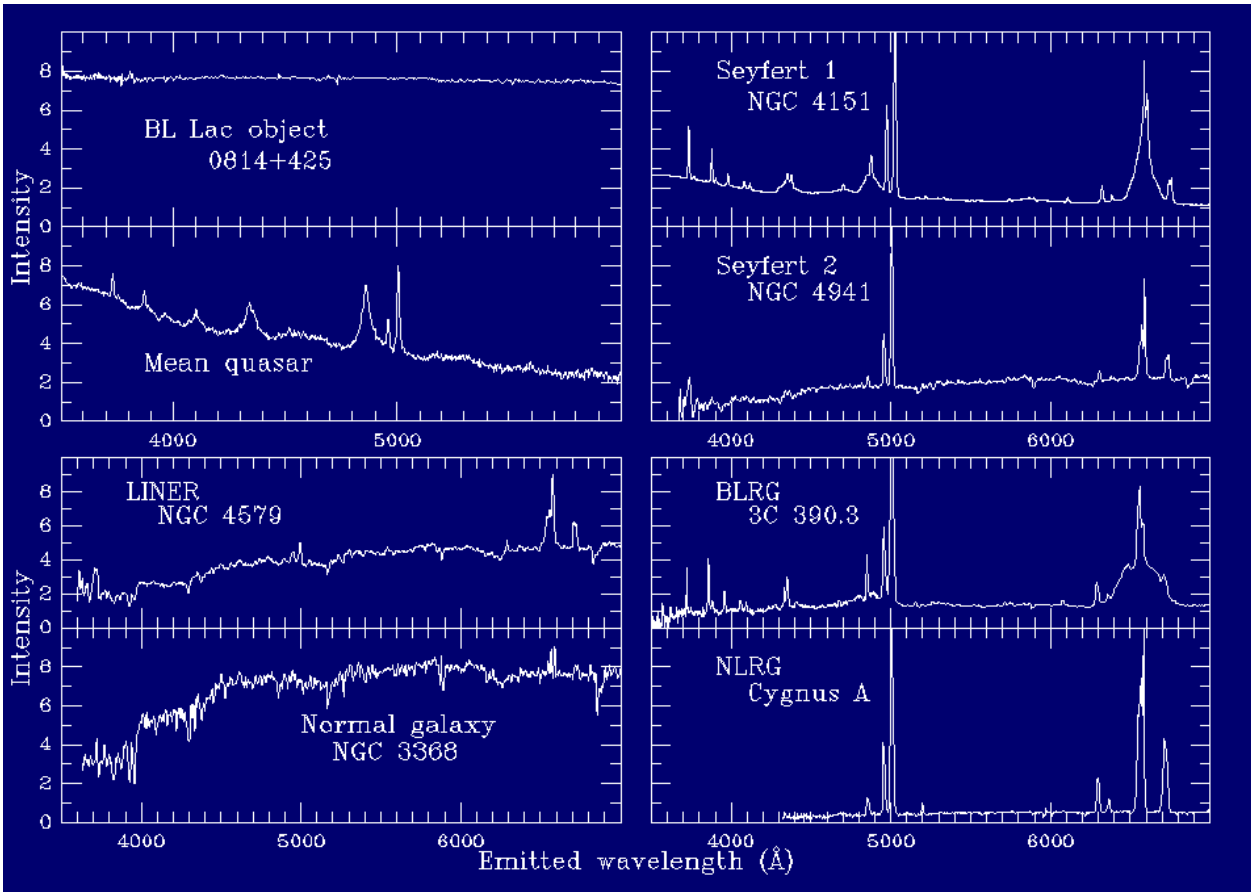

Figure 1: The spectra of a variety of galaxies. The lowermost left plot displays the spectra of an inactive galaxy, while each other spectra displays a different type of active galaxy[2].

When galaxies are plotted according to the entire galaxy's color (the difference between the ultraviolet u-band magnitude and the red r-band magnitude), they tend to lie in two dense regions, a "blue" region representing a smaller difference between these magnitudes, and a "red" region representing a larger difference between these magnitudes. Blue light from galaxies is predominantly produced by young, hot stars. As such, blue galaxies tend to be younger and actively forming stars, while red galaxies tend to be older, with most star formation having ceased. Between the blue and red sequences lies the "green valley", which is not representative of a literal green color but of a range of colors between those of the red and blue sequences and is less densely populated than these sequences. These red and blue sequences, as well as the green valley, can be seen in Figure 2. Galaxies that lie in the green valley have a higher probability of hosting AGN than galaxies that lie in the red or blue sequences[3]. This suggests that AGN feedback may have a role in quenching the star formation of their host galaxies, which makes intuitive sense because stars tend to form from cold gas, but the accretion of matter into

[2] Keel (2002)
[3] Nadra et al. (2007)

an AGN will heat the gas of the host galaxy[4]. AGN may also be a result of the star formation having been quenched through another mechanism, such as stellar feedback or mass loss[5]. As such, the study of AGN is beneficial to the understanding of galaxy evolution.

Goulding et al. suggest that interactions between galaxies may also have a hand in causing AGN formation. Specifically, they identify that AGN are far more common in one or both galaxies undergoing a merger than in isolated galaxies, most likely as a result of gas inflows being directed to one or both galactic nuclei[6]. However, little else is known about how interactions between neighboring galaxies impact the formation of AGN. In order to better understand this, I investigated a number of groups of galaxies to determine which properties correlate with an increased probability of finding AGN among the galaxies in these groups. Specifically, I investigated the dependence of the galaxies' color distribution on both the presence or lack of AGN and the number of neighboring galaxies within a specified distance of each galaxy, as well as the correlation between the number of neighbors a galaxy has and that galaxy's probability of hosting an AGN.

2. Data Sample

I collected the data used in this study from Data Release (DR) 12 of the Sloan Digital Sky Survey (SDSS). The SDSS used a dedicated 2.5-meter telescope at Apache Point Observatory, New Mexico to develop a detailed map of large portions of the sky, having mapped about one fourth of the sky at the time of DR 12[7]. Data was collected through repeated photometric exposures, followed by repeated spectroscopic exposures which employed optical fibers positioned by holes drilled in thin metal plates that fed into the SDSS spectrographs. Automated image processing was then used to correct various image defects and deblend overlapping objects[8]. The data collected by SDSS is supplemented by a number of Value Added Catalogues (VAC) contributed by various collaborators[9].

The online SDSS workbench CasJobs was used to collect the data for my sample. The sample was limited to include only galaxies with spectroscopic redshift of no less than 0.004 in order to exclude foreground objects and no greater than 0.25 to reduce the effects of galactic evolution on my results. An upper apparent u-band magnitude limit of 24.25 and an upper apparent r-band magnitude limit of 17.77 were used to conform to the limits of SDSS at these redshifts because, although the spectroscopic survey may contain some galaxies fainter than this, these are the highest apparent magnitudes for which completeness is guaranteed. After collecting data for each galaxy in the DR 12 database, a CasJobs query function that returned all objects

[4] Krumholz (2012)
[5] Schwaninski et al. (2014)
[6] Goulding et al. (2018)
[7] "About the SDSS"
[8] "Glossary of SDSS Terminology"
[9] "Galaxy Properties from the Portsmouth Group"

within a specified angular distance of a specific point in the sky was used to find which galaxies neighbored each other. Finally, the Portsmouth Stellar Kinematics and Emission Line Fluxes VAC was used to determine whether each galaxy hosts AGN based on that galaxy's emission line spectra[10]. For the purposes of this investigation, I considered the presence of or lack of AGN in each host galaxy a binary and made no distinction between types of AGN.

The sample consists of $6.86 * 10^5$ galaxies, approximately 53.9% of which host AGN. Using the data collected from SDSS, the color of each galaxy was defined as the corresponding galaxy's u-band magnitude minus r-band magnitude. As galaxies tend to lie within two distinct areas of high density on a plot comparing color and absolute magnitude, a color divider line was employed, defined by the equation:

$$(u - r) = -0.018(M_r + 21)^2 - 0.137(M_r + 21) + 2.20$$

where u represents the u-band magnitude, r represents the r-band magnitude, and M_r represents the absolute magnitude[11]. Galaxies lying to the left of this line are considered "blue", making up 28.4% of the sample, and galaxies lying to the right of this line are considered "red", making up 71.6% of the sample.

[10] Thomas et al. (2013)
[11] James (2017)

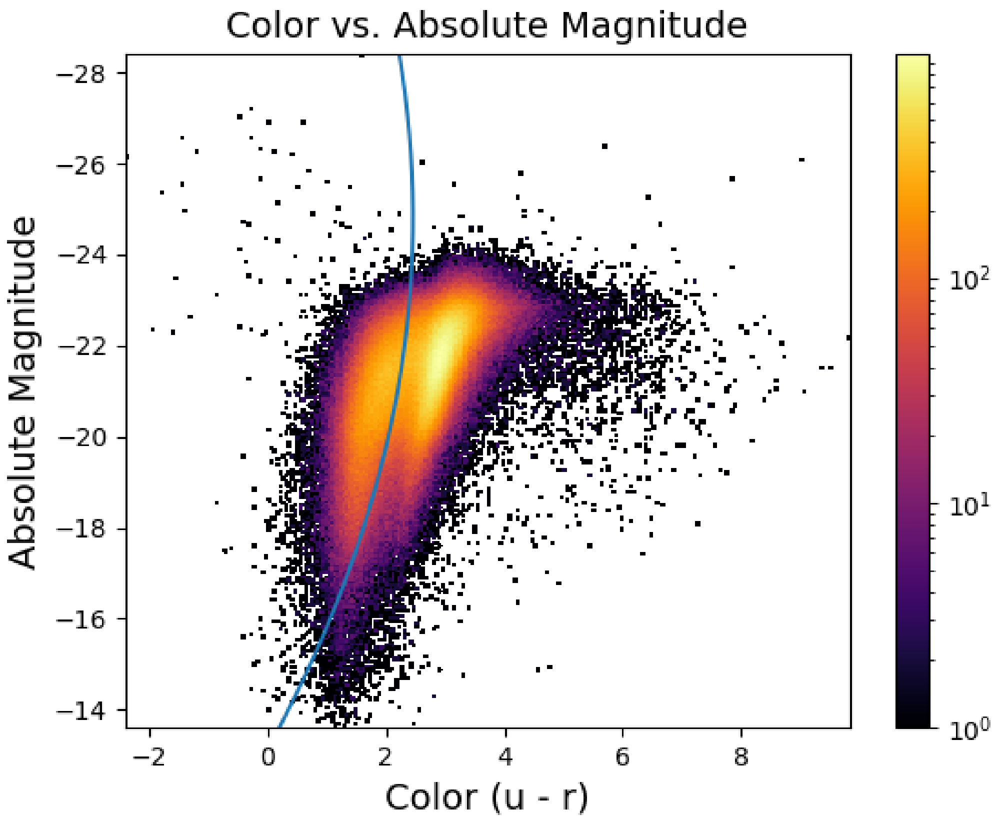

Figure 2: Color/Magnitude diagram displaying the galaxies in our sample, including a blue line representing the dividing line between the red and blue sequences.

3. Data Subsamples

Following the collection of the data, the galaxies in my sample were organized into subsamples based on the presence or lack of AGN in each, and were from there further subdivided based on the number of galaxies within a projected distance of 0.5 Megaparsecs (Mpc) of the target. For this subsample investigation, only galaxies with 4 or fewer nearby neighbors were considered because the more nearby neighbors a galaxy has, the more likely it is that these results are subject to projection error. Galaxies with 4 or fewer nearby neighbors make up 94% of the sample.

Color-magnitude diagrams similar to Figure 2 were created for each of these subsamples. Figure 3 shows these plots for galaxies with no nearby neighbors and with 4 nearby neighbors for both the subsample with AGN and the subsample without AGN. The division between blue and red galaxies visible in the color/magnitude diagram of the entire sample is much clearer in galaxies without AGN than in galaxies with AGN in both displayed cases, which is unsurprising due to the abundance of AGN in the green valley. As a result, galaxies without AGN exist in wider range of colors than galaxies with AGN, especially in the case of 4 nearby neighbors. In this case, the standard deviation (STD) in color of galaxies without AGN is 0.61, while the STD in color of galaxies with AGN is little more than half this at 0.32. In the case of galaxies with no nearby neighbors, this remains true, although to a less extreme degree; galaxies without AGN have a STD of 0.65, while galaxies with AGN have a STD of 0.47. As shown by Table 1, the difference between the STDs of galaxies with AGN and galaxies without AGN steadily increases with luminosity; however, the fractional difference between STDs remains relatively consistent.

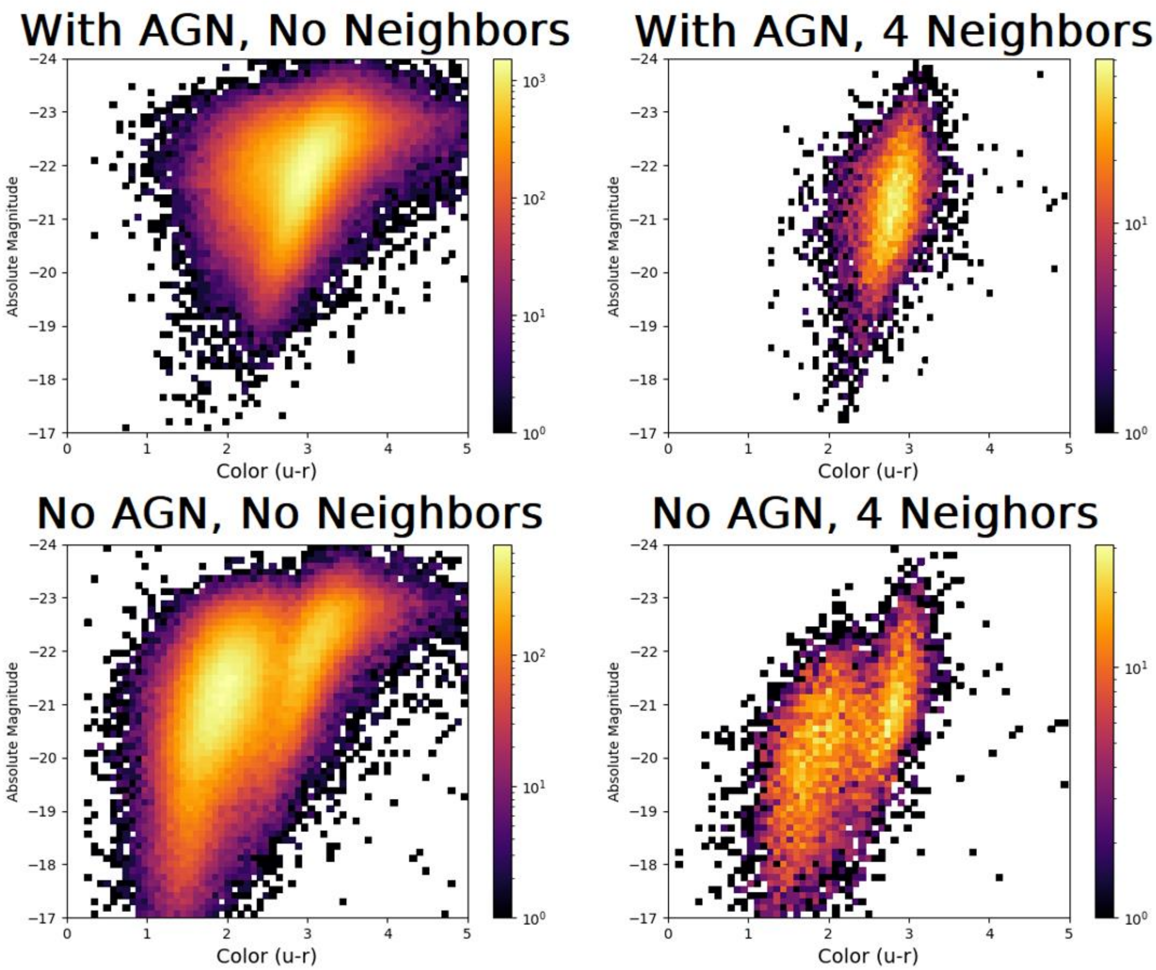

Figure 3: Color/Magnitude diagrams for several data subsamples. The upper left shows galaxies with AGN and no neighbors within 0.5 Mpc. The upper right shows galaxies with AGN and 4 neighbors within 0.5 Mpc. The lower left shows galaxies without AGN and with no neighbors within 0.5 Mpc. The lower right shows galaxies without AGN and with 4 neighbors within 0.5 Mpc.

Luminosity Range	Mean Color (AGN)	Mean Color (no AGN)	STD (AGN)	STD (no AGN)	STD difference	STD fractional difference
-17.0 to -18.75	2.172	1.547	0.449	0.354	0.095	0.268
-18.75 to -20.5	2.597	1.840	0.323	0.431	0.108	0.251
-20.5 to -22.25	2.835	2.235	0.396	0.540	0.144	0.267
-22.25 to -24.0	3.160	2.942	0.510	0.690	0.180	0.261

Table 1: A comparison of color for galaxies with no nearby neighbors. The "Luminosity Range" column displays the luminosity range of the corresponding row. The "Mean (AGN)" column shows the mean color of galaxies with AGN and no nearby neighbors. The "Mean (no AGN)" column shows the mean color of galaxies with no AGN and no nearby neighbors. The "STD (AGN)" column shows the STD of galaxies with AGN and no nearby neighbors. The "STD (no AGN)" column shows the STD of galaxies with no AGN and no nearby neighbors. The "STD difference" column shows the difference between these two STDs. The "STD fractional difference" column shows the difference between these two STDs as a fraction of the STD of galaxies with no AGN and no neighbors.

Additionally, a two sample Kolmogorov-Smirnov (KS) test was used to compare the sample with AGN to the sample without AGN for each number of nearby neighbors from 0 to 4. For example, galaxies with AGN and with 2 nearby neighbors were compared to galaxies without AGN and with 2 nearby neighbors, etc. The KS test compares the Cumulative Distribution Function (CDF) of two distributions and returns two values: the KS statistic, which indicates the maximum separation between the CDFs of the samples, and the P value, which represents the probability of obtaining this KS statistic assuming both samples were drawn from the same parent distribution. Figure 4 displays the CDFs for galaxies with no more than 4 nearby neighbors, with each number of neighbors receiving its own CDF. Additionally, the functions for galaxies are divided based on the presence or lack of AGN. The P value for each of these tests was effectively 0, indicating once again that the color distribution among galaxies with AGN is notably different than the color distribution among those without. There was no clear correlation between number of nearby neighbors and KS statistic. When the test was performed again to compare galaxies within the same subsample but with differing numbers of neighbors – for example, galaxies with AGN and 0 nearby neighbors were compared with galaxies with AGN and 1 nearby neighbor, etc. – the P values were again effectively zero, except in the case of galaxies without AGN and with 3 neighbors compared to galaxies without AGN and with 4 neighbors. In this case, the P value was approximately 0.4.

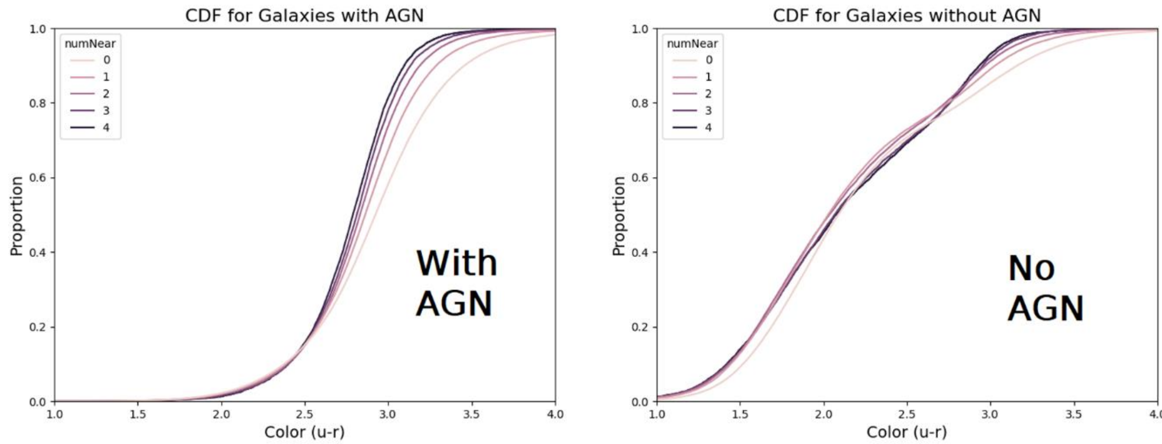

Figure 4: CDFs for several data subsamples. The left shows color CDFs for galaxies with AGN. The right shows color CDFs for galaxies without AGN.

4. AGN Probability and Number of Neighbors

The most notable result obtained through my investigation of these data subsamples was that AGN are most commonly found in galaxies with the fewest nearby neighbors. As seen in Figure 5, there is a strong trend towards fewer AGN in galaxies with more nearby neighbors, both when "nearby" is defined as within a projected separation of 0.5 Mpc and as within a projected separation of 0.1 Mpc. The number of galaxies in each bin also drastically decreases as the number of nearby neighbors increases. As such, the bin representing galaxies with no nearby neighbors contains about 62% of the sample when "nearby" is defined as within a projected separation of 0.5 Mpc, and about 87% of the sample when "nearby" is defined as within a projected separation of 0.1 Mpc. I assigned uncertainties, represented by the error bars in these histograms, by dividing the square root of the number of galaxies with AGN in each bin by the total number of galaxies in each bin. These errors tend to be so small as to barely noticeable for most bins representing few neighbors and become much larger for bins which represent larger numbers of neighbors and as such contain much fewer galaxies.

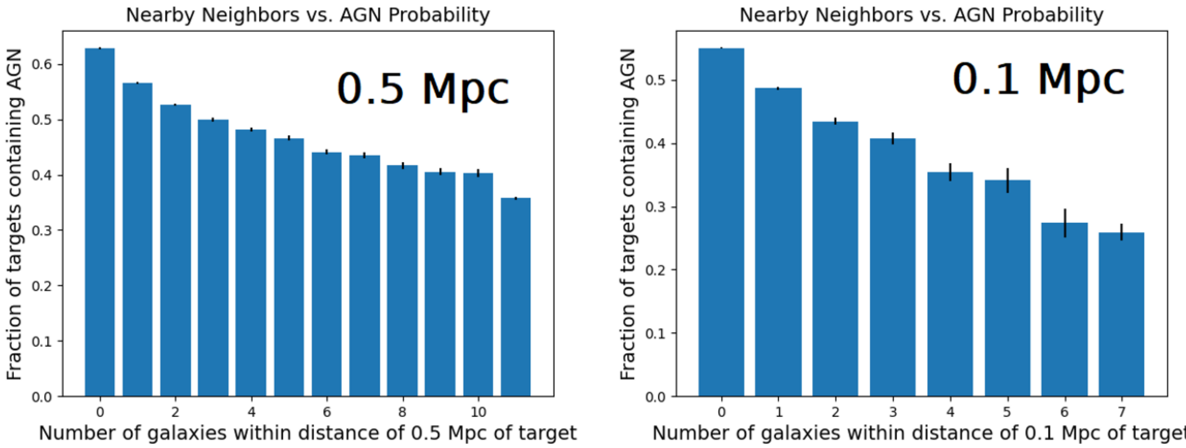

Figure 5: The fraction of each bin of galaxies containing AGN when galaxies are binned based on number of nearby neighbors. In the left plot, "nearby" is defined as "within 0.5 Mpc", while in the right plot, "nearby" is defined as "within 0.1 Mpc". The rightmost bin in each histogram represents galaxies with at least this number of neighbors.

Bin Number	Number of Galaxies (0.5 Mpc)	Number of Galaxies (0.1 Mpc)
0	$4.24 * 10^5$	$5.99 * 10^5$
1	$1.25 * 10^5$	$6.47 * 10^4$
2	$5.22 * 10^4$	$1.30 * 10^4$
3	$2.72 * 10^4$	$4.10 * 10^3$
4	$1.62 * 10^4$	$1.72 * 10^3$
5	$1.03 * 10^4$	903
6	$6.97 * 10^3$	519
7	$4.94 * 10^3$	$1.45 * 10^3$
8	$3.66 * 10^3$	N/A
9	$2.80 * 10^3$	N/A
10	$2.13 * 10^3$	N/A
11	$1.03 * 10^4$	N/A

Table 2: The number of galaxies in each bin of the histograms in Figure 5. The "Number of Galaxies (0.5 Mpc)" column displays the number of galaxies in each respective bin of the plot on the left of Figure 5, while the "Number of Galaxies (0.1 Mpc)" column displays the number of galaxies in each respective bin of the plot on the right of Figure 5.

Certain selection biases and sources of uncertainty do, of course, apply here. For example, this study only includes galaxies found in the SDSS spectroscopic survey, meaning that many faint galaxies have been excluded. The use of holes drilled in metal plates to position the optical fibers collecting data for this survey also impacts the survey's ability to detect nearby neighbors. Because these holes have a finite, non-zero size, there must be a finite, non-zero distance between them, which imposes a lower bound of about 55 arcseconds on the angular separation between galaxies measured by the survey. In other words, if two galaxies are too near each other in the sky, they cannot both be measured by the survey. This is further complicated by the fact that some areas of the sky are captured on multiple plates due both to a desire to collect data on objects nearer to each other than this lower bound and to the fact that two-dimensional disks cannot completely cover the three-dimensional dome of the sky without some overlap. Additionally, due to projection error, two galaxies with a great physical distance between them can appear to be neighbors if they are in the same area of the sky. I imposed some redshift restrictions while collecting the data to mitigate this effect – if the difference between two galaxies' redshifts is at least 0.08 (corresponding to a distance of about 350 Mpc), they are not considered to be neighbors – but because the difference in redshift corresponding to a difference in distance of 0.5 Mpc is smaller than the redshift uncertainty recorded in the SDSS database, I was unable to eliminate this effect entirely. Finally, the SDSS deblending algorithm will occasionally "shred" galaxies; bright, star forming regions outside a galaxy's central bulge are sometimes interpreted as separate galaxies. An investigation of a number of images collected during the survey indicates that, as of DR12, shredding affects only 5% of galaxies[12]. Because these star forming regions will predominantly appear in galaxies in the blue sequence while AGN predominantly appear in galaxies in the green valley, it is unknown what proportion of these shredded galaxies host AGN.

5. Conclusion

In this paper, I investigated how the probability of finding an AGN in a given target galaxy correlates with the galaxies surrounding that target. I accomplished this through the use of data on a number of galaxies collected through the spectroscopic survey of SDSS DR 12. I found that, among galaxies with AGN, galaxies with few neighbors tend to be redder than galaxies with many neighbors consistently for all values of color greater than 2.5, a range containing about 80% of the sample. While this trend appears for some color ranges among galaxies without AGN, it is only consistent outside of the green valley. Additionally, I found that the likelihood of a galaxy to host an AGN steadily declines as the number of neighbors that the galaxy has increases, and that this trend is consistent across multiple distances. Over 60% of galaxies with no nearby neighbors within a projected distance of 0.5 Mpc host AGN, but less than half of galaxies with 3 or more neighbors within this projected distance host AGN. When "nearby" is defined as "within a projected distance of 0.1 Mpc", galaxies hosting AGN make up the majority of only the bin representing no nearby neighbors, and they make up less than half of every other bin, steadily trending downwards as the number of neighbors increases.

This latter result appears to contradict the conclusions of Goulding et al. who assert that a merger between two galaxies is far more likely to produce AGN than a solitary galaxy due to gas from the merging galaxies flowing towards one or both galactic nuclei[13]; however, there are several caveats which may provide solutions to this contradiction. For example, many of the galaxies neighboring each other may not have begun, or may be early in, the process of merging. In order to merge, two galaxies must obviously be extremely close to each other, and the

[12] "Deblending Overlapping Objects"
[13] Goulding et al. (2018)

projected distances used in this investigation allow for a much greater physical separation and, thus, much less significant tidal interactions. This is supported by the fact that galaxy mergers lead to new star formation, meaning that galaxies will usually become bluer while merging. Because galaxies with few nearby neighbors tend to be redder than galaxies with many neighbors, it is likely that most galaxies with few nearby neighbors are not undergoing mergers. On the other hand, galaxies with few or no neighbors according to this investigation may, in fact, be merging with smaller galaxies that are below the spectroscopic survey's lower luminosity bound, or they may have merged recently enough that their nuclei are still consuming gas inflows. Further investigation is required to determine which of these potential phenomena, if any, affects a significant enough portion of the sample to explain this discrepancy.

References:

"About the SDSS". Sloan Digital Sky Survey Skyserver. URL: http://skyserver.sdss.org/dr12/en/sdss/sdsshome.aspx

"Deblending Overlapping Objects". Sloan Digital Sky Survey Skyserver. URL: https://www.sdss.org/dr12/algorithms/deblend/

"Galaxy Properties from the Portsmouth Group". Sloan Digital Sky Survey Skyserver. URL: https://www.sdss.org/dr14/spectro/galaxy_portsmouth/

"Glossary of SDSS Terminology". Sloan Digital Sky Survey Skyserver. URL: https://www.sdss.org/dr12/help/glossary/

Goulding, A. D. et al. (2018), "Galaxy interactions trigger rapid black hole growth: An unprecedented view from the Hyper Suprime-Cam survey". Publications of the Astronomical Society of Japan, vol. 70, Issue SP1. January 2018. DOI: https://doi.org/10.1093/pasj/psx135

James, Derrick and Barbara Ryden (2017). "The Green Valley: Separating Galaxy Populations in Color-Magnitude Space".

Keel, B. (2002), "Optical spectra of various kinds of active galactic nuclei". URL: https://pages.astronomy.ua.edu/keel/agn/spectra.html

Krumholz, M. R. (2012), "STAR FORMATION IN ATOMIC GAS". The Astrophysical Journal, vol. 759. pp. 9-18, October 2012. DOI: https://doi.org/10.1088/0004-637X/759/1/9

Nadra, K. et al. (2007), "AEGIS: The Color-Magnitude Relation for X-Ray-selected Active Galactic Nuclei". The Astrophysical Journal, vol. 660. L11-L14, April 2007. DOI: https://doi.org/10.1086/517918

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

Schwaninski, K. et al. (2014), "The green valley is a red herring: Galaxy Zoo reveals two evolutionary pathways towards quenching of star formation in early- and late-type galaxies". Monthly Notices of the Royal Astronomical Society, vol. 440. pp. 889-907, March 2014. DOI: https://doi.org/10.1093/mnras/stu327

Thomas, D. et al. (2013), "Stellar velocity dispersions and emission line properties of SDSS-III/BOSS galaxies". Monthly Notices of the Royal Astronomical Society, vol. 431. pp 1383-1397, March 2013. DOI: https://doi.org/10.1093/mnras/stt261