

Classifications of Emission-Line Galaxies with SDSS & WISE

Final Project

Nicolás Guerra-Varas

Professor Dragana Ilić
Tutor Isidora Jankov
3 de junio de 2023

Introduction

Any study of active galactic nuclei (AGN) has to start by finding a proper sample of these objects. However, this is not always an easy task, as factors like redshift and extinction can complicate the analysis. It is becoming increasingly important to have effective classification schemes in place to accommodate for all the big data surveys that are coming soon such as the LSST and SKA.

Some existing classification schemes are the Baldwin, Phillips and Terlevich (BPT) diagram [1], the WHAN diagram [2], and the Trouille, Barger and Tremonti (TBT) diagrams, which all use spectral data publicly available from the Sloan Digital Sky Survey (SDSS). The Wide-field Infrared Survey Explorer (WISE) [3] also provides a classification scheme from its color-color diagram [4, 4].

The goal of this report is to test different classification schemes, specifically how and how many AGN they find. In the first two tasks, I used SDSS data to compare the BPT, TBT and WHAN diagrams. Then, in the third task, I used WISE data to the SDSS classification schemes with another survey.

Task 1: BPT and WHAN on a Complete Sample

In this first task, my goal is to evaluate and compare the classification scheme from the different BPT diagrams and the WHAN diagram, focusing on the number of AGN found.

I started by defining the sample and downloading the data using the sql query in Code 1. In Table 4, there is a statistical summary of the fluxes and equivalent widths of the sample. The result from the SQL query is in file `all_galaxies.csv`, and the same table including flux ratios and labels is in file `all_galaxies_after.csv`.

I checked the subclasses that SDSS assigned spectroscopically and found that there are 6820 starburst galaxies, 14626 starforming galaxies, 1881 AGN, 304 AGN broadline, 108 starforming broadline galaxies, 231 starburst broadline galaxies and 5485 unassigned. The larger class indicates there are 498 QSOs and 29502 inactive galaxies. Then, I saw that the objects in the QSO class are composed majoritarily by the starburst broadline, as well as other inactive galaxy classes. Similarly, the objects in the GALAXY (or inactive galaxy) class is also contaminated by AGN subclasses. This indicated that these labels are not fully reliable.

In addition, the SDSS query selects narrow line objects only, but there are several objects in the broadline subclasses. This also makes me think these labels are not the most reliable.

1. BPT Diagrams

Then, I plotted the BPT diagram of the full sample in Figure 1. I used `seaborn` [5] for this. The papers [6] and [7] propose a theoretical line to separate star-forming galaxies from AGN. The one from the former is stricter and poses a lower limit on the amount of AGN.

$$\log ([\text{OIII}]/\text{H}\beta) = \frac{0.61}{\log ([\text{NII}]/\text{H}\alpha) - 0.47} + 1.19 \quad (1)$$

$$\log ([\text{OIII}]/\text{H}\beta) = \frac{0.61}{\log ([\text{NII}]/\text{H}\alpha) - 0.05} + 1.30 \quad (2)$$

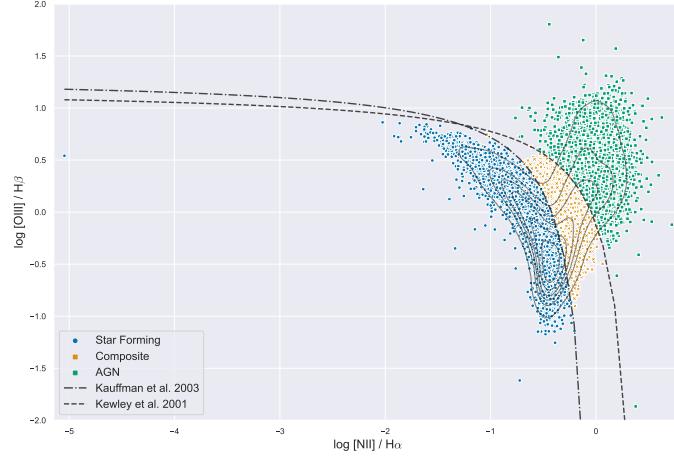


Figure 1: BPT Diagram: $[\text{NII}]/\text{H}\alpha$ vs $[\text{OIII}]/\text{H}\beta$. The lines correspond to the thresholds defined by [6] and [7].

The objects labeled as “Composite” in Figure 1 have a different classification by [6] and [7].

This classification resulted in 18988 star-forming galaxies (63 % of the sample), 5607 composites (19 %), and 5405 AGN (18 %) (see Table 1.b).

Then, I also plotted the BPT diagrams that use $[\text{SII}]/\text{H}\alpha$ and $[\text{OII}]/\text{H}\alpha$ instead of $[\text{NII}]/\text{H}\alpha$. The plot that uses the $[\text{SII}]$ lines is duplet with wavelengths 6717 and 6731 angstrom, so I used the summed them to get the BPT diagram. Only the paper by [6] defines thresholds in these diagrams:

$$\log ([\text{OIII}]/\text{H}\beta) = \frac{0.72}{\log ([\text{SII}]/\text{H}\alpha) - 0.32} + 1.30 \quad (3)$$

$$\log ([\text{OIII}]/\text{H}\beta) = \frac{0.73}{\log ([\text{OI}]/\text{H}\alpha) + 0.59} + 1.33 \quad (4)$$

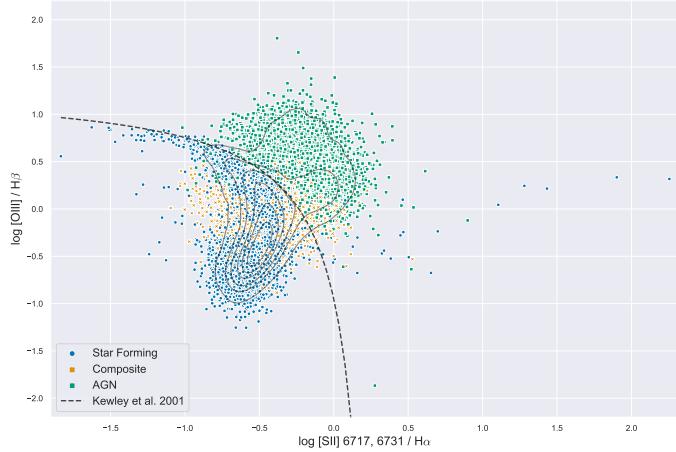


Figure 2: BPT Diagram: $[\text{SII}]/\text{H}\alpha$ vs $[\text{OIII}]/\text{H}\beta$. Objects above the threshold defined by [6]

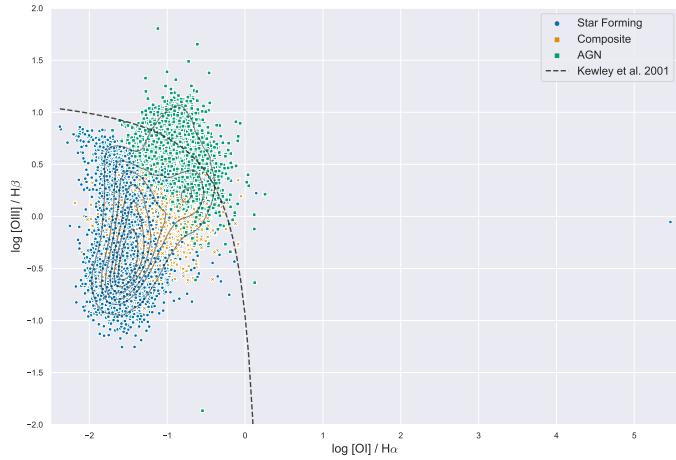


Figure 3: BPT Diagram: $[\text{OII}]/\text{H}\alpha$ vs $[\text{OIII}]/\text{H}\beta$.

I used the first diagram in Figure 1 to define the BPT labels of the objects, so the lines mark clear thresholds. In the BPT that uses $[\text{SII}]$ (see Figure 2), the threshold is clearly classifying the vast majority of the star-forming galaxies correctly, with only a few being outside of that region. Similarly, the vast majority of the AGN are above the threshold, and are classified correctly. So this diagram does a good job at preserving the same classification from the first BPT diagram. However, it seems to classify composite galaxies as star-forming galaxies by default. This is because this threshold is more restrictive than the one defined by [7]. On the other hand, the BPT diagram that uses $[\text{OI}]$ (see Figure 3) provides a worse classification. The threshold in this diagram passes directly through the population of AGN objects. So, even though it classifies the vast majority of star-forming galaxies correctly, this classification would provide a significantly underestimated number of AGN.

Thus, after the BPT with $[\text{NII}]$, the diagram that uses $[\text{SII}]$ is the most reliable one.

2. WHAN Diagram

Then, I plotted the WHAN diagram: H_{α} equivalent width $W_{H_{\alpha}}$ vs $[\text{NII}]/H_{\alpha}$ [2]. Since the equivalent widths provided by SDSS are majoritarily negative, I got their absolute value, and used that to plot the WHAN diagram in a logarithmic scale. In their paper, [2] define a criterion to classify objects:

- Pure star-forming galaxies: $\log [\text{NII}]/H_{\alpha} < -0.4$ and $W_{H_{\alpha}} > 3\text{\AA}$
- Strong AGN / Seyferts: $\log [\text{NII}]/H_{\alpha} > -0.4$ and $W_{H_{\alpha}} > 6\text{\AA}$
- Weak AGN: $\log [\text{NII}]/H_{\alpha} > -0.4$ and $3 < W_{H_{\alpha}} < 6\text{\AA}$
- Retired galaxies/Fake AGN (RGs): $W_{H_{\alpha}} < 3\text{\AA}$
- Passive/Lineless galaxies: $W_{H_{\alpha}} < 0.5\text{\AA}$ and $W_{[\text{NII}]} < 0.5\text{\AA}$

Figures 4 and 5 show the resulting WHAN diagrams with the WHAN and BPT labels respectively. In Figure 5, the divisions at $\log [\text{NII}]/H_{\alpha} = -0.4$ and at $W_{H_{\alpha}} = 3\text{\AA}$ seem to separate the star-forming galaxies from the rest fairly well.

Tabla 1: Classifications from the BPT and WHAN diagrams.

	# of Objects	%		# of Objects	%
Star-Forming	15995	53.3	Star-Forming	18988	63.3
Seyfert	8031	26.8	Composite	5607	18.7
Weak AGN	1777	5.9	AGN	5405	18.0
RG	4181	13.9			
Passive	16	0.05			

(a) Classification from the WHAN diagram.

(b) Classification from the BPT diagram.

The WHAN diagram classifies 8031 objects as Seyferts or strong AGN, whilst the BPT diagram only 5405 (see Table 1). I cross-checked the BPT labels of the WHAN Seyferts and the WHAN labels of the BPT AGN to see how they disagree. Pie charts presenting this are in Figure 6.

- The WHAN Seyferts have different BPT labels (see Figure 5). The majority of them are AGN or composites. However, about 36 % are BPT star-forming galaxies, which is not good.
- The BPT AGN are mostly RGs (fake AGN). This indicates that the BPT diagram cannot identify these as fake and defaults them to the AGN class. The rest are classified majoritarily as Seyferts and weak AGN, with only a small percentage as star-forming. So, the BPT diagram does well at identifying AGN but has the disadvantage that it cannot distinguish RGs from others.

Overall, the WHAN diagram is capable of providing a more detailed classification. However, it classifies a large portion of sources as RGs or fake AGN. With more time, I would like to see if the proportion of RGs is overestimated in this diagram since the threshold to define them doesn't seem restrictive enough.

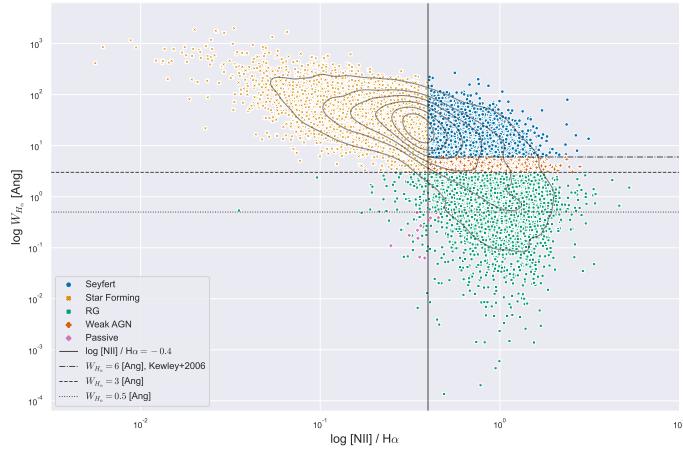


Figure 4: WHAN Diagram with its own labels: $[\text{NII}] / \text{H}\alpha$ vs W_{H_α} [Ang].

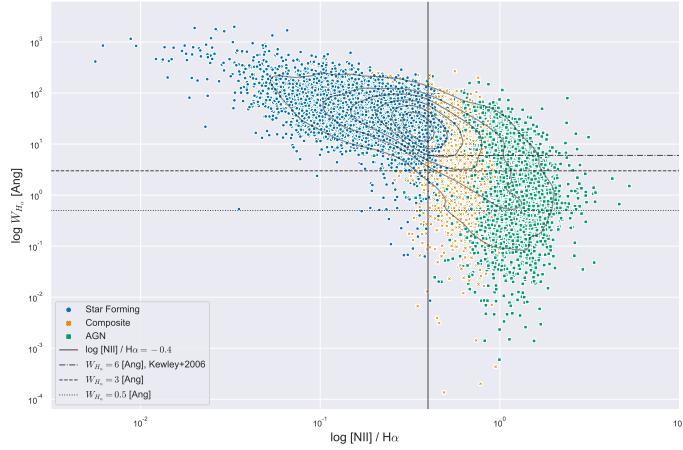
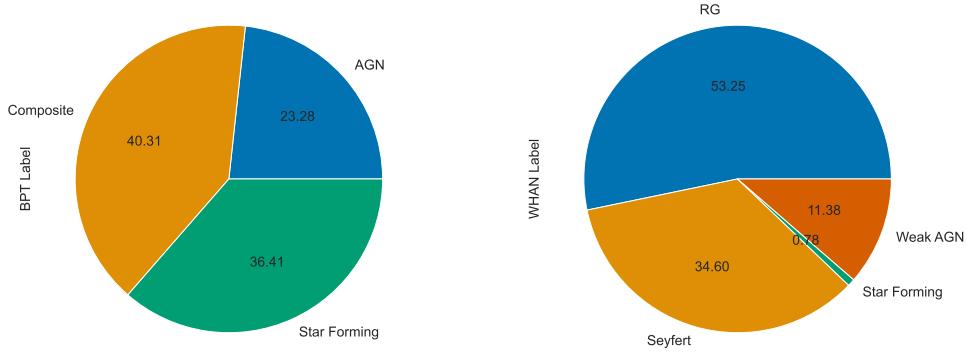


Figure 5: WHAN Diagram with BPT labels: $[\text{NII}] / \text{H}\alpha$ vs W_{H_α} [Ang].



(a) BPT labels of the objects classified as Seyferts by the WHAN diagram: 2924 star-forming galaxies, by the BPT diagram: 1870 Seyferts, 2878 RGs, 615 weak AGN and 42 star-forming.

Figure 6: Pie charts illustrating the labels of the objects marked as AGN and Seyferts by the BPT and WHAN diagrams.

Task 2: BPT and TBT on a Subsample

In this first task, my goal is to evaluate and compare the classification scheme from the different BPT diagrams and the WHAN diagram, focusing on the number of AGN found.

The next goal is to test the TBT diagram, compare it to the BPT diagram and see how the number of labeled AGN changes from one to another.

1. Defining the Sample

Within my python code, I defined a subsample using the same criteria as [8]. I kept sources that have H α , H β , [OIII], [NII], [OII] and [NeIII] fluxes with signal-to-noise (S/N) > 5. This results in 3688 objects (called TBT subsample, see file `TBT_subsample.csv`).

2. k-Correction

I corrected the color magnitudes ($g - z$) for redshift using the k-correction [9]. For this, I used the code available in [10]. I computed the correction with both the g and z filters to get their redshift-corrected values. Then I used this relation to obtain the k-corrected color:

$${}^0(g - z) = (g - z) - ({}^0g - {}^0z) \quad (5)$$

3. BPT and TBT Diagrams

Then, I plotted the BPT diagram as in the previous section in Figure 7. The TBT diagram was proposed by [8] and it uses the [NeIII] and [OII] absorption lines and the redshift-corrected color

$(g - z)$. I plotted this diagram in Figure 8. The objects that lie above the threshold:

$$^0(g - z) = -1.2 \times \log ([\text{NeIII}]/[\text{OII}]) - 0.4 \quad (6)$$

are classified as AGN, and the rest as star-forming galaxies. In Figure 8, the three BPT classes lie in fairly distinct places. The AGN class lies almost completely above the threshold, the composite class passes through it, and the star-forming galaxy class is mostly below it, although it does bleed into the other region of the plot. Table 2 shows the result of the BPT and TBT classifications. Like in the previous section, I cross-checked the BPT and TBT labels to see how they get confused (see Figure 9).

Tabla 2: Classifications from the BPT and TBT diagrams.

	BPT Diagram		TBT Diagram	
	# of Objects	% of the Sample	# of Objects	% of the Sample
Star-Forming	2263	61.36	2252	61.06
Composite	321	8.70	-	-
AGN	1104	29.93	1436	38.94

- The BPT diagram (see Figure 9.a) classifies 57.3 % of the TBT AGN as either AGN or composites, so it mostly agrees with the TBT classification scheme. However, it is more uncertain because it is not capable of necessarily distinguishing AGN from composites, and the TBT AGN subsample is heavily contaminated by star-forming galaxies (more than 40 %).
- The TBT diagram (see Figure 9.b) classifies 1436 objects as AGN, more than 99 % of which are also classified as AGN by the BPT diagram. The BPT AGN subsample is virtually not contaminated by star-forming galaxies. This indicates that most likely, if an object is an AGN according to the BPT diagram, it will also be an AGN according to the TBT diagram. So, these objects will certainly be AGN.

On the other hand, both classification schemes result in a very similar number of star-forming galaxies, so they will mostly agree when classifying this class of objects. Overall, the TBT diagram is very good at distinguishing between AGN and star-forming, but it has more uncertainty in the classification of the latter.

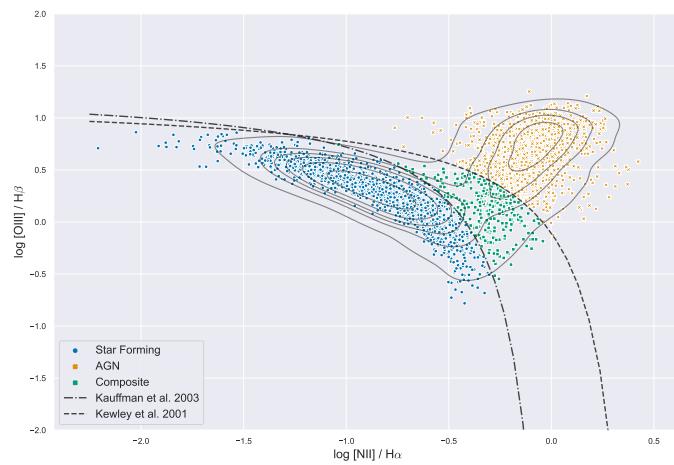


Figure 7: BPT Diagram: $[\text{NII}]/\text{H}\alpha$ vs $[\text{OIII}]/\text{H}\beta$ for the defined subsample of Task 2.

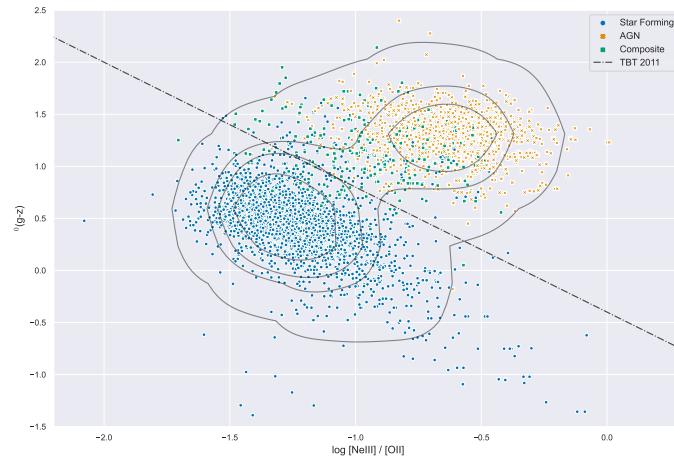
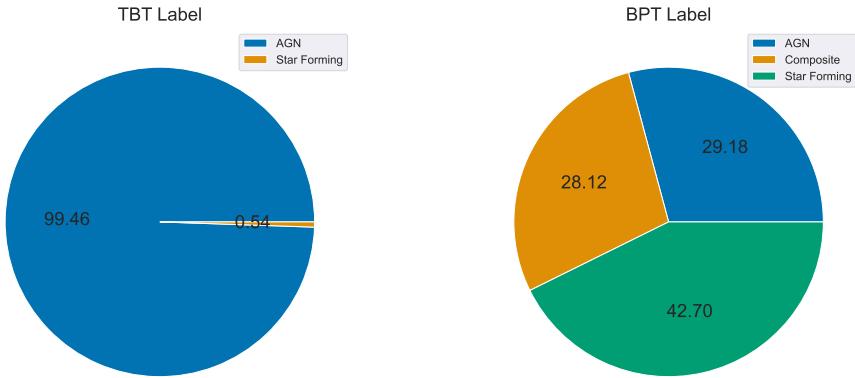


Figure 8: TBT Diagram: $[\text{NeIII}]/[\text{OII}]$ vs ${}^0(\text{g}-\text{z})$ for the defined subsample.



(a) TBT labels of the objects classified as AGN by the BPT diagram: 2 star-forming galaxies and 1102 by the TBT diagram: 87 star-forming galaxies, 247 composites and 1102 AGN.

Figure 9: Pie charts illustrating the labels of the objects marked as AGN by the BPT and TBT diagrams.

Task 3: WISE Color-Color Diagram on a Subsample

The goal of this last task is to classify a subsample with the WISE color-color diagram and compare it to the labels from SDSS data.

1. Defining the Subsample

I started by downloading the WISE colors and fluxes: W1, W2 and W3 which are at 3.4, 4.6 and $12 \mu\text{m}$ respectively. I used the SQL query in Code 2 for this. It has the same conditions as Code 1, but I did not select the top columns, I queried for the full table these conditions define. The result is in the `wise_full_query.csv` file and it consists of 194527 objects. Then, I used TOPCAT [11] to merge both tables based on their coordinates within a radius of $3''$. This is about half of the angular resolution at $3.4 \mu\text{m}$ ($6.1''$) [3], so this ensures the WISE counterparts will correspond to the correct object in the initial sample without confusion. The final subsample has 25499 objects (see the `wise_subsample.csv` file; the `wise_subsample_after.csv` file also includes the labels from the diagrams). In Table 6, a statistical summary of the colors magnitudes and fluxes is presented.

Out of 30000, 4501 objects were discarded to create the subsample, which is about 15 % of the original sample.

2. Color-Color Diagram

First, I started by defining labels from the color-color diagram as Figure 11 of [12] (see Figure 10). These criteria result in 13212 star-forming disks, 12119 intermediate disks, 2297 spheroids and 584 AGN & (u)LIRGs.

In the same plot, I included the criterion that [4] defines to select luminous AGN with X-ray detection using mid-IR WISE colors:

- $y = 0.315x (+0.297, -0.110)$
- $y = -3.172x + 0.436$

where $y = \log\left(\frac{f_{4.6}}{f_{3.4}}\right)$ and $x = \log\left(\frac{f_{12}}{f_{4.6}}\right)$. These boundaries define a wedge in both the flux ratios and color-color diagrams since:

$$W_i - W_j = 2.5 \times \log\left(\frac{f_j}{f_i}\right)$$

Before plotting, I converted the flux units from digital numbers [DN] to Janskys [Jy], which is the unit used by [4, 13]. I used this formula from [14]:

$$F_\nu[\text{Jy}] = F_\nu[\text{DN}] \times 10^{-m_{Vega}/2.5}, \quad (7)$$

where m_{Vega} is the corresponding color magnitude.

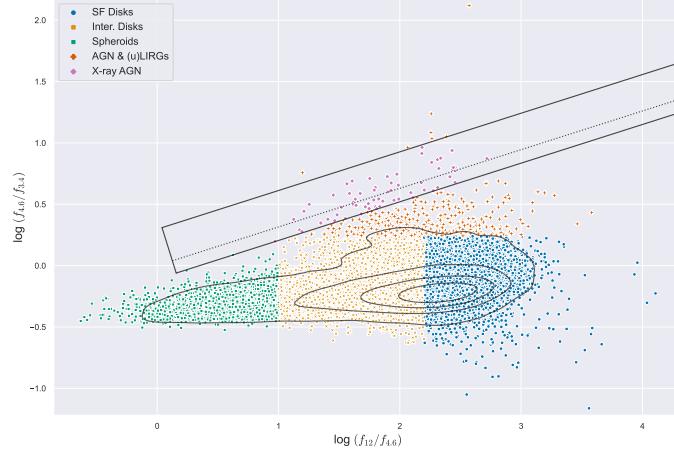


Figure 10: WISE color-color diagram as in [4, 13].

This classification scheme then resulted in 65 X-ray AGN (see Figure 10), 64 of which are classified as AGN & (u)LIRGs by [12], and the remaining one as a spheroid.

I also plotted the data on top of Figure 12 from Wright et al. 2010 [3], which illustrates the locations of different classes in the WISE color-color diagram (see Figure 11). In this plot, the data mostly falls in the Starburst/LINER/LIRG region, as well as the Seyfert, spirals and elliptical/star regions. The Figure 12 from [3] has the disadvantage that many regions of different classes overlap, which makes it difficult to decide which is the correct label for the objects. Perhaps these blobs could be projected in three dimensions, where they might be more separated. Either way, if WISE

is capable of providing a reliable classification with this diagram, it has a great advantage because these labels are detailed.

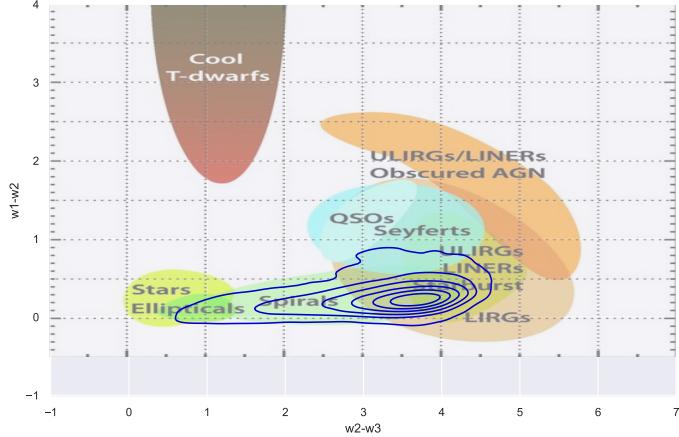


Figure 11: Density contours on top of Figure 12 from Wright et al. 2010 [3]. ^a

^a This plot was too busy with a scatterplot, so I prefer to present this and Figure 10 separately.

Tabla 3: SDSS labels, and classifications from the BPT diagram and WISE.

	#	%		#	%
Star-Forming	11942	46.83	Star-Forming	16060	62.98
Intermediate	11041	43.30	Composite	4970	19.49
Spheroids	2008	7.87	AGN	4469	17.53
AGN & (u)LIRGs	443	1.74			
X-Ray AGN	65	0.25			

(a) Classifications from WISE mid-IR color selections by [12] and [4, 13].

	#	%
Galaxy	25061	98.28
QSO	438	1.72

(c) SDSS classes.

The results from the WISE color-color classification scheme is presented in Table 3.a, as well as the SDSS classes (Table 3.c) and the BPT labels (Table 3.b).

The classes provided directly by SDSS only have 438 QSOs. By comparing with all the other classification schemes, I believe this number is heavily underestimated. Thus, the SDSS classes might not be very reliable.

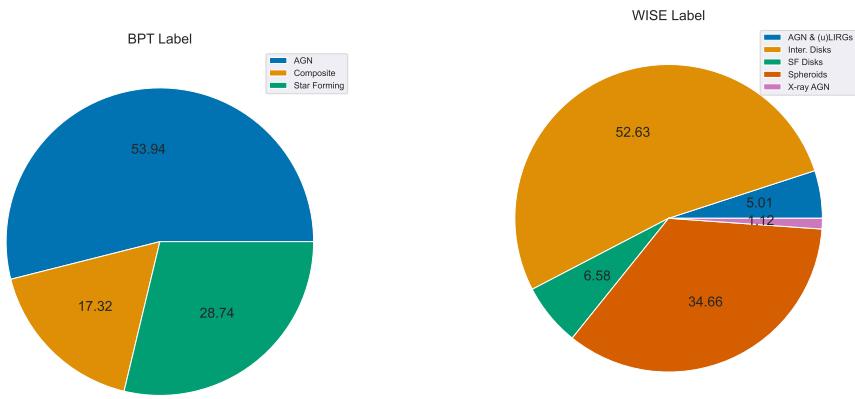
Just as in the other sections, I cross-checked the labels of the AGN objects between the WISE and BPT classification schemes with pie charts (see Figure 12).

The quantity of AGN identified by WISE is much less than that of the BPT diagram when considering the AGN & (u)LIRGs and X-ray AGN classes. However, there is a large quantity of intermediate disks (11041). The region in Figure 10 below the wedge that corresponds to intermediate disks is quite close to the AGN & (u)LIRGs and X-ray AGN regions, so most likely, a portion of

these are also AGN but perhaps with not such clear signatures. This may indicate the objects that are labeled as AGN by WISE are certainly part of this class, and probably have very clear AGN signatures. However, even though Figure ?? shows that the majority of WISE AGN objects are classified as AGN or composites by the BPT diagram, there are many star-forming galaxies in this population. Thus, it is not possible to say these classification schemes agree.

On the other hand, BPT AGN objects have a wider variety of WISE labels. The majority of these are labeled as intermediate disks, and only about 6 % as true AGN. This supports my idea that a portion of these intermediate disks are actually AGN. The rest ($\sim 40\%$) are classified as spheroids or star-forming disks. Once again, this indicates that WISE and BPT classifications are not consistent with one another. Furthermore, more information and a more detailed study would be necessary to see which of the two is more reliable, and it will depend on the given samples and data.

There are only 289 objects that are classified as AGN by both WISE and BPT. Given that these classification schemes do not agree as I discussed above, these 289 objects will certainly be AGN.



(a) BPT labels of the objects classified as either X- by the BPT diagram: 294 star-forming disks, 2352 ray AGN or AGN & (u)LIRGs by WISE: 274 AGN, intermediate disks, 1549 spheroids, 224 AGN & 146 star-forming galaxies and 88 composites. (b) WISE labels of the objects classified as AGN by the WISE and BPT diagrams.

Figure 12: Pie charts illustrating the labels of the objects marked as AGN by the BPT and TBT diagrams.

References

- [1] J. A. Baldwin, M. M. Phillips, and R. Terlevich, “Classification parameters for the emission-line spectra of extragalactic objects.,” *Publications of the Astronomical Society of the Pacific*, vol. 93, p. 5, feb 1981.
- [2] R. Cid Fernandes, G. Stasińska, A. Mateus, and N. Vale Asari, “A comprehensive classification of galaxies in the Sloan Digital Sky Survey: how to tell true from fake AGN?,” *MNRAS*, vol. 413, pp. 1687–1699, May 2011.
- [3] E. L. Wright, P. R. M. Eisenhardt, A. K. Mainzer, M. E. Ressler, R. M. Cutri, T. Jarrett, J. D. Kirkpatrick, D. Padgett, R. S. McMillan, M. Skrutskie, S. A. Stanford, M. Cohen, R. G. Walker, J. C. Mather, D. Leisawitz, I. Gautier, Thomas N., I. McLean, D. Benford, C. J. Lonsdale, A. Blain, B. Mendez, W. R. Irace, V. Duval, F. Liu, D. Royer, I. Heinrichsen, J. Howard,

- M. Shannon, M. Kendall, A. L. Walsh, M. Larsen, J. G. Cardon, S. Schick, M. Schwalm, M. Abid, B. Fabinsky, L. Naes, and C.-W. Tsai, “The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance,” *AJ*, vol. 140, pp. 1868–1881, Dec. 2010.
- [4] S. Mateos, A. Alonso-Herrero, F. J. Carrera, A. Blain, M. G. Watson, X. Barcons, V. Braito, P. Sevrgnini, J. L. Donley, and D. Stern, “Using the Bright Ultrahard XMM-Newton survey to define an IR selection of luminous AGN based on WISE colours,” *MNRAS*, vol. 426, pp. 3271–3281, Nov. 2012.
 - [5] M. L. Waskom, “seaborn: statistical data visualization,” *Journal of Open Source Software*, vol. 6, no. 60, p. 3021, 2021.
 - [6] L. J. Kewley, M. A. Dopita, R. S. Sutherland, C. A. Heisler, and J. Trevena, “Theoretical Modeling of Starburst Galaxies,” *ApJ*, vol. 556, pp. 121–140, July 2001.
 - [7] G. Kauffmann, T. M. Heckman, C. Tremonti, J. Brinchmann, S. Charlot, S. D. M. White, S. E. Ridgway, J. Brinkmann, M. Fukugita, P. B. Hall, Ž. Ivezić, G. T. Richards, and D. P. Schneider, “The host galaxies of active galactic nuclei,” *MNRAS*, vol. 346, pp. 1055–1077, Dec. 2003.
 - [8] L. Trouille, A. J. Barger, and C. Tremonti, “The OPTX Project. V. Identifying Distant Active Galactic Nuclei,” *ApJ*, vol. 742, p. 46, Nov. 2011.
 - [9] I. V. Chilingarian, A.-L. Melchior, and I. Y. Zolotukhin, “Analytical approximations of K-corrections in optical and near-infrared bands,” *MNRAS*, vol. 405, pp. 1409–1420, July 2010.
 - [10] “k-correction code.” <http://kcor.sai.msu.ru/getthecode/>, Date of access: June 2023.
 - [11] “Sdss sky server.” <https://www.star.bris.ac.uk/~mbt/topcat/>, Date of access: May 2023.
 - [12] T. H. Jarrett, M. E. Cluver, C. Magoulas, M. Bilicki, M. Alpaslan, J. Bland-Hawthorn, S. Brough, M. J. I. Brown, S. Croom, S. Driver, B. W. Holwerda, A. M. Hopkins, J. Loveday, P. Norberg, J. A. Peacock, C. C. Popescu, E. M. Sadler, E. N. Taylor, R. J. Tuffs, and L. Wang, “Galaxy and mass assembly (gama): Exploring the wise web in g12,” *The Astrophysical Journal*, vol. 836, p. 182, feb 2017.
 - [13] S. Mateos, A. Alonso-Herrero, F. J. Carrera, A. Blain, P. Sevrgnini, A. Caccianiga, and A. Ruiz, “Uncovering obscured luminous AGN with WISE,” *MNRAS*, vol. 434, pp. 941–955, Sept. 2013.
 - [14] “Wise data pre-processing.” https://wise2.ipac.caltech.edu/docs/release/allsky/expsup/sec4_4h.html, Date of access: June 2023.

Appendix: Computer Codes and Tables

0.1. SQL Queries

Code 1: SQL Query: Defines Sample of 30,000 Objects

```
1  SELECT TOP 30000 s.specobjid, s.plate, s.mjd, s.fiberID,
2    s.ra, s.dec, s.class, s.subclass, s.z,
3    g.oiii_5007_flux, g.oiii_5007_flux_err,
4    g.h_alpha_flux, g.h_alpha_flux_err,
5    g.h_beta_flux, g.h_beta_flux_err,
6    g.nii_6584_flux, g.nii_6584_flux_err,
7    g.nii_6584_reqw, g.nii_6584_reqw_err,
8    g.sii_6717_flux, g.sii_6717_flux_err,
9    g.sii_6731_flux, g.sii_6731_flux_err,
10   g.oi_6300_flux, g.oi_6300_flux_err,
11   g.h_alpha_reqw, g.h_alpha_reqw_err,
12   g.neiii_3869_flux, g.neiii_3869_flux_err,
13   g.oi_3726_flux, g.oi_3726_flux_err,
14   g.oi_3729_flux, g.oi_3729_flux_err,
15   p.psfMag_g, p.psfMagErr_g,
16   p.psfMag_z, p.psfMagErr_z
17
18  FROM GalSpecLine AS g
19  JOIN SpecObjAll AS s ON g.specobjid = s.specobjid
20  JOIN PhotoObjAll AS p ON s.bestobjid = p.objID
21
22 WHERE
23 (s.class = 'QSO' or s.class = 'GALAXY')
24 AND s.z <= 0.3
25 AND 2.355 * g.sigma_balmer <= 700
26 AND 2.355 * g.sigma_forbidden <= 700
27 AND g.oiii_5007_flux >> 0
28 AND (g.oiii_5007_flux / g.oiii_5007_flux_err) > 1
29 AND g.h_alpha_flux >> 0
30 AND (g.h_alpha_flux / g.h_alpha_flux_err) > 1
31 AND g.h_beta_flux >> 0
32 AND (g.h_beta_flux / g.h_beta_flux_err) > 1
33 AND g.nii_6584_flux >> 0
34 AND (g.nii_6584_flux / g.nii_6584_flux_err) > 1
35 AND g.sii_6717_flux >> 0
36 AND (g.sii_6717_flux / g.sii_6717_flux_err) > 1
37 AND g.sii_6731_flux >> 0
38 AND (g.sii_6731_flux / g.sii_6731_flux_err) > 1
39 AND g.oi_6300_flux >> 0
40 AND (g.oi_6300_flux / g.oi_6300_flux_err) > 1
41 AND g.neiii_3869_flux >> 0
42 AND (g.neiii_3869_flux / g.neiii_3869_flux_err) > 1
43 AND g.oi_3726_flux >> 0
44 AND (g.oi_3726_flux / g.oi_3726_flux_err) > 1
45 AND g.oi_3729_flux >> 0
46 AND (g.oi_3729_flux / g.oi_3729_flux_err) > 1
```

```

47     AND (p.psfMag_g / p.psfMagErr_g) > 1
48     AND (p.psfMag_z / p.psfMagErr_z) > 1
49

```

Code 2: SQL Query: Cross-ID to find WISE Colors

```

1   SELECT --count(*)
2     s.specobjid, s.plate, s.mjd, s.fiberID,
3     s.ra, s.dec, s.class, s.subclass, s.z,
4     W.w1mpro AS w1, W.w2mpro AS w2, W.w3mpro AS w3,
5     W.w1sigmpro AS w1, W.w2sigmpro AS w2, W.w3sigmpro AS w3,
6     W.w1flux, W.w2flux, W.w3flux,
7     W.w1sigflux, W.w2sigflux, W.w3sigflux
8
9
10    FROM GalSpecLine AS g
11    JOIN SpecObjAll AS s ON g.specobjid = s.specobjid
12    JOIN PhotoObjAll AS p ON s.bestobjid = p.objID
13    JOIN wise_xmatch AS x ON s.bestobjid = x.sdss_objid JOIN wise_allsky AS w ON w.cntr =
14      ↪ x.wise_cntr
15
16    WHERE
17      (s.class = 'QSO' or s.class = 'GALAXY')
18      AND (W.w1flux / W.w1sigflux) > 1
19      AND (W.w2flux / W.w2sigflux) > 1
20      AND (W.w3flux / W.w3sigflux) > 1
21      AND (W.w1mpro / W.w1sigmpro) > 1
22      AND (W.w2mpro / W.w2sigmpro) > 1
23      AND (W.w3mpro / W.w3sigmpro) > 1
24      AND s.z <= 0.3
25      AND 2.355 * g.sigma_balmer <= 700
26      AND 2.355 * g.sigma_forbidden <= 700
27      AND g.oiii_5007_flux <> 0
28      AND (g.oiii_5007_flux / g.oiii_5007_flux_err) > 1
29      AND g.h_alpha_flux <> 0
30      AND (g.h_alpha_flux / g.h_alpha_flux_err) > 1
31      AND g.h_beta_flux <> 0
32      AND (g.h_beta_flux / g.h_beta_flux_err) > 1
33      AND g.nii_6584_flux <> 0
34      AND (g.nii_6584_flux / g.nii_6584_flux_err) > 1
35      AND g.sii_6717_flux <> 0
36      AND (g.sii_6717_flux / g.sii_6717_flux_err) > 1
37      AND g.sii_6731_flux <> 0
38      AND (g.sii_6731_flux / g.sii_6731_flux_err) > 1
39      AND g.oi_6300_flux <> 0
40      AND (g.oi_6300_flux / g.oi_6300_flux_err) > 1
41      AND g.neiii_3869_flux <> 0
42      AND (g.neiii_3869_flux / g.neiii_3869_flux_err) > 1
43      AND g.oii_3726_flux <> 0
44      AND (g.oii_3726_flux / g.oii_3726_flux_err) > 1
45      AND g.oii_3729_flux <> 0
46      AND (g.oii_3729_flux / g.oii_3729_flux_err) > 1
47      AND (p.psfMag_g / p.psfMagErr_g) > 1
48      AND (p.psfMag_z / p.psfMagErr_z) > 1

```

2. Plots

Code 3: $[\text{H}\alpha]$ vs $[\text{OIII}]/\text{H}\beta$). The other BPT diagrams are done similarly.]Function to plot the BPT diagram ($[\text{NII}]/\text{H}\alpha$ vs $[\text{OIII}]/\text{H}\beta$). The other BPT diagrams are done similarly.

```

1 def plot_BPT(data, x, y, hue, xlabel, ylabel,
2 Kauffman=True, Kewley=True,
3 ylim=(-2.0, 2.0),
4 density=False,
5 save=False, save_name=None):
6
7 plt.figure(figsize=(12, 8))
8
9 plt.ylim(ylim)
10
11 sns.scatterplot(data=data, x=x, y=y, palette='colorblind',
12 hue=hue, style='BPT final', size=hue, sizes=(30, 15))
13 if density:
14     sns.kdeplot(data=data, x=x, y=y, levels=7, alpha=0.5, cut=2, color='k')
15
16 # -----
17 # Kauffman
18 if Kauffman:
19     x_to_plot = np.linspace(galaxies[x].min(), 0)
20     plt.plot(x_to_plot, 0.61 / (x_to_plot - 0.05) + 1.3, label='Kauffman et al. 2003',
21     linewidth=1.5, alpha=0.8, c='k', linestyle='dashdot')
22
23 # -----
24 # Kewley
25 if Kewley:
26     x_to_plot_1 = np.linspace(galaxies[x].min(), 0.4)
27     plt.plot(x_to_plot_1, 0.61 / (x_to_plot_1 - 0.47) + 1.19, label='Kewley et al. 2001',
28     linewidth=1.5, alpha=0.8, c='k', linestyle='dashed')
29
30 # -----
31
32 plt.xlabel(xlabel, fontsize=13)
33 plt.ylabel(ylabel, fontsize=13)
34
35 plt.legend(loc='lower left', fontsize=12)
36
37 if save:
38     plt.savefig('./BPT_Diagrams/' + save_name + '.pdf')
39
40 plt.show()
41

```

Code 4: Code to Plot the WHAN diagram in Figure 4. Figure 5 is done similarly.

```

1 plt.figure(figsize=(12, 8))
2
3 sns.scatterplot(data=galaxies, x='nii_h_alpha', y='h_alpha_reqw_abs', palette='colorblind',
4 hue='WHAN label', style='WHAN label', size='WHAN label', sizes=(30, 15))
5
6 sns.kdeplot(data=galaxies, x='nii_h_alpha', y='h_alpha_reqw_abs', levels=7,
7 alpha=0.5, cut=2, color='k', log_scale=True)
8
9 plt.ylim(10**(-4.2), 10**(3.8))
10 plt.xlim(10**(-2.5), 10)
11
12 plt.vlines(10**(-0.4), 10**(-4.2), 10**(3.8), linewidth=1.0, alpha=0.8, colors='k', linestyle='solid',
13 label=r'log [NII] / H$\alpha$ = -0.4$')
14 plt.hlines(6, 10**(-0.4), 10, linewidth=1.0, alpha=0.8, colors='k', linestyle='dashdot',
15 label=r'$W_{H\alpha}$ = 6$ [Ang], Kewley+2006')
16 plt.hlines(3, 10**(-2.5), 10, linewidth=1.0, alpha=0.8, colors='k', linestyle='dashed',
17 label=r'$W_{H\alpha}$ = 3$ [Ang]')
18 plt.hlines(0.5, 10**(-2.5), 10, linewidth=1.0, alpha=0.8, colors='k', linestyle='dotted',
19 label=r'$W_{H\alpha}$ = 0.5$ [Ang]')
20
21 plt.xlabel(r'log [NII] / H$\alpha$', fontsize=13)
22 plt.ylabel(r'log $W_{H\alpha}$ [Ang]', fontsize=13)
23
24 plt.xscale('log')
25 plt.yscale('log')
26
27 plt.legend(loc='lower left', fontsize=10)
28
29 plt.savefig('./WHAN_diagrams/WHAN_diagram_whan_label.pdf')
30
31 plt.show()
32
```

Code 5: Code to Plot the TBT diagram in Figure 8

```

1 plt.figure(figsize=(12, 8))
2
3 sns.scatterplot(data=sub_gals, x='neiii_oii_log', y='g_z_color_kcorr', palette='colorblind',
4 hue='BPT final', style='BPT final', size='BPT final', sizes=(30, 15))
5
6 sns.kdeplot(data=sub_gals, x='neiii_oii_log', y='g_z_color_kcorr', levels=5,
7 alpha=0.5, cut=2, color='k', log_scale=False)
8
9 x_to_plot = np.linspace(10**(-2.7), 10**(0.5), 50)
10 plt.plot(np.log10(x_to_plot), - 1.2 * np.log10(x_to_plot) - 0.4,
11 linewidth=1.2, alpha=0.8, c='k', linestyle='dashdot', label='TBT 2011')
12
13 plt.ylim(-1.5, 2.5)
14 plt.xlim(-2.2, 0.3)
15
16 plt.xlabel(r'log [NeIII] / [OII]', fontsize=11)
17 plt.ylabel(r'$^0(g-z)$', fontsize=11)
18
```

```

19 plt.legend(loc='upper right', fontsize=10)
20
21 # plt.savefig('./TBT_diagrams/TBT_full_subsample.pdf')
22
23 plt.show()
24

```

Code 6: Code to Plot the WISE color-color diagram in Figure 10

```

1 fig, ax = plt.subplots(figsize=(10, 8))
2
3 sns.scatterplot(data=wise_sub, x='flux3/flux2_Jy_log', y='flux2/flux1_Jy_log', palette='
4     &gt; colorblind',
5 hue='final_wise_label', style='final_wise_label', size='final_wise_label', sizes=(30, 15))
6
7 sns.kdeplot(data=wise_sub, x='flux3/flux2_Jy_log', y='flux2/flux1_Jy_log',
8 levels=5, alpha=0.7, cut=2, color='k', log_scale=False)
9
10 x_to_plot = np.linspace((0.436 - 0.297) / (3.172 + 0.315), 4.3, 50)
11 plt.plot(x_to_plot, 0.315 * x_to_plot + 0.297, linewidth=1.2, alpha=0.8, c='k', linestyle='solid')
12
13 x_to_plot_1 = np.linspace((0.436 + 0.110) / (3.172 + 0.315), 4.3, 50)
14 plt.plot(x_to_plot_1, 0.315 * x_to_plot_1 - 0.110, linewidth=1.2, alpha=0.8, c='k', linestyle='
15     &gt; solid')
16
17 x_to_plot_2 = np.linspace(0.436 / (3.172 + 0.315), 4.3, 50)
18 plt.plot(x_to_plot_2, 0.315 * x_to_plot_2, linewidth=1.2, alpha=0.8, c='k', linestyle='dotted')
19
20 x_to_plot_3 = np.linspace((0.436 - 0.297) / (3.172 + 0.315), (0.436 + 0.110) / (3.172 + 0.315),
21     &gt; 50)
22 plt.plot(x_to_plot_3, - 3.172 * x_to_plot_3 + 0.436, linewidth=1.2, alpha=0.8, c='k', linestyle='
23     &gt; solid')
24
25 plt.xlabel(r'log $ \left( f_{12} / f_{4.6} \right) $', fontsize=13)
26 plt.ylabel(r'log $ \left( f_{4.6} / f_{3.4} \right) $', fontsize=13)
27
28 plt.legend(loc='upper left', fontsize=11)
29
# plt.savefig('./WISE_plots/wise_final_label.pdf')
plt.show()

```

Code 7: Code to Plot the WISE color-color diagram on top of Figure 12 of Write et al. 2010 [3] in Figure 11

```

1 fig = plt.figure(figsize=(522*2/120, 573*2/120), dpi=120)
2
3 img = plt.imread('./WISE_plots/Fig12_Wright_2010.jpg')
4
5 plt.imshow(img, zorder=1, extent=[-1.01, 7, -0.5, 4], aspect='equal', alpha=0.5)
6
7 sns.kdeplot(data=wise_sub, x='w2-w3', y='w1-w2', levels=7, alpha=1.0, cut=2, color='

```

```

    ↪ mediumblue', zorder=5)

8
9 # plt.savefig('./WISE_plots/color_color_top_Wright.pdf')
10
11 plt.show()
12

```

Code 8: Code to Plot the Pie Charts in Figure 6. Figures 9 and Figures 12 are done similarly.

```

1 plt.figure(figsize=(7, 7))
2
3 galaxies.loc[galaxies['BPT final']=='AGN'].groupby('TBT label').size().plot(kind='pie', autopct='
4     ↪ %.2f', labels=None,
5     textprops={'fontsize': 14},
6     colors=['#0173b2', '#de8f05', '#029e73'],
7     legend=True)
8
9 plt.title('TBT Label', fontsize=14)
10 plt.ylabel(None)
11
12 # plt.savefig('./TBT_diagrams/tbt_label_pie.pdf')
13
14 plt.show()
15
16 # -----
17
18 plt.figure(figsize=(7, 7))
19
20 galaxies.loc[galaxies['TBT label']=='AGN'].groupby('BPT final').size().plot(kind='pie', autopct='
21     ↪ %.2f', labels=None,
22     textprops={'fontsize': 14},
23     colors=['#0173b2', '#de8f05', '#029e73'],
24     legend=True)
25
26 plt.title('BPT Label', fontsize=14)
27 plt.ylabel(None)
28
29 # plt.savefig('./TBT_diagrams/bpt_label_pie.pdf')
30
31 plt.show()

```

Tabella 4: Description of the fluxes [10^{-17} erg s $^{-1}$ cm $^{-2}$] and equivalent widths [Å] of the full sample.

	[OIII] $\lambda 5007$	H_{α}	H_{β}	[NII] $\lambda 6584$	$W_{[NII]}$	[SII] $\lambda 6717$	[SII] $\lambda 6731$	[OI] $\lambda 6300$	$W_{H_{\alpha}}$	[NeIII] $\lambda 3869$	[OI] $\lambda 3726$	[OI] $\lambda 3729$	g	z
Mean	4.9×10^6	1.2×10^7	3.2×10^6	2.7×10^6	-9.6	2.0×10^6	1.6×10^6	4.7×10^5	-30.2	5.1×10^5	3.6×10^6	4.3×10^6	19.4	18.2
std	2.3×10^8	4.1×10^8	1.1×10^8	9.9×10^7	8.1	6.7×10^7	5.2×10^7	1.7×10^7	52.9	2.0×10^7	1.3×10^8	1.7×10^8	0.7	0.9
Minimum	1.4	1.6	1.2	1.4	-170.1	1.4	1.1	0.9	-1986.6	1.3	1.5	2.2	15.9	14.5
25%	14.9	102.4	24.5	43.0	-12.4	23.7	16.8	5.3	-36.2	4.6	21.5	24.0	19.0	17.6
Median	31.8	206.4	49.1	78.2	-7.5	41.2	29.8	8.9	-19.0	6.8	41.8	46.8	19.5	18.1
75%	87.4	398.8	97.9	145.4	-4.4	73.5	54.3	15.5	-8.0	11.5	92.3	101.3	19.9	18.7
Maximum	2.3×10^{10}	2.7×10^{10}	8.5×10^9	7.9×10^9	7.8	3.4×10^9	2.9×10^9	1.1×10^9	2.2	1.9×10^9	8.6×10^9	1.1×10^{10}	25.1	27.0

Table 5: Description of the fluxes and magnitudes of the selected TBT subsample. The colors and their k-corrected values are included

	H_{α}	H_{β}	[NeIII] $\lambda 3869$	[OII] $\lambda 3726$	[OII] $\lambda 3729$	$(g - z)$	$^0(g - z)$
Mean	1.6×10^7	4.1×10^{61}	5.5×10^5	4660122.03	5526051.95	1.13	-1.01
STD	4.4×10^8	1.2×10^8	1.7×10^7	1.4×10^8	1.7×10^8	0.41	1.80
Minimum	21.85	6.47	1.27	8.53	8.46	-7.22	-23.03
25%	167.58	42.76	4.58	37.73	42.22	0.84	-1.25
Median	276.18	68.58	6.64	60.45	67.78	1.12	-1.02
75%	451.44	109.58	9.80	101.69	113.01	1.42	-0.80
Maximum	1.9×10^{10}	5.6×10^9	8.7×10^8	7.3×10^9	8.5×10^9	5.75	213.09

Tabla 6: Description of WISE colors [magnitudes] and fluxes [10^{-3} Jy] of the selected subsample.

	W1	W2	W3	$f_{3.4}$	$f_{4.6}$	f_{12}
Mean	14.21	13.95	10.66	3.4	2.5	476.3
STD	0.86	0.86	1.02	21.2	47.1	6811.1
Minimum	10.00	8.73	5.38	0.0	0.0	0.1
25 %	13.72	13.45	10.02	0.3	0.2	13.3
Median	14.24	13.96	10.67	0.6	0.4	46.0
75 %	14.74	14.49	11.35	1.7	1.1	151.9
Maximum	17.57	17.16	13.82	1.5961	6584.6	789652.5