

Statistical Analysis of AGN Fraction as a Function of Projected Separation in SDSS Galaxy Pairs

Kyumin Cho

University of Manitoba, SNU Radio AGN Group

Abstract

Galaxy mergers and gravitational interactions are considered critical mechanisms for driving galactic gas reservoirs into nuclear scales, thereby triggering active galactic nuclei (AGN). In this study, we present a systematic statistical framework to investigate the AGN fraction as a function of projected physical separation (r_p) using an interacting galaxy pair sample selected from the Sloan Digital Sky Survey (SDSS). We construct an automated data analysis pipeline in Python to extract emission-line fluxes and classify the primary ionization sources utilizing the BPT diagnostic diagram with the theoretical boundary defined by Kewley et al. (2001). Our statistical analysis demonstrates a significant, monotonic increase in the AGN fraction at tight separations ($r_p < 25 \text{ kpc}$), peaking at approximately 14.80%, which represents a threefold enhancement compared to isolated control environments. This small-scale environmental dependency indicates a distinct physical time-lag between large-scale tidal starbursts and the subsequent activation of supermassive black holes (SMBHs).

1. Introduction

Galaxy mergers and close gravitational interactions are fundamental evolutionary drivers that transform galactic morphologies, regulate star formation histories, and stimulate the growth of supermassive black holes (SMBHs) residing at galactic centers. Standard hydrodynamic simulations indicate that during a close encounter between two galaxies, non-axisymmetric gravitational perturbations generate powerful tidal torques. These tidal forces disrupt the rotation of the interstellar medium (ISM), inducing gas clouds to lose their angular momentum and funnel rapidly into the sub-parsec nuclear core. This dense gas accumulation serves as the primary fuel source required to trigger an Active Galactic Nucleus (AGN).

A critical outstanding issue in observational cosmology is constraining the precise spatial and temporal scales over which these tidal forces efficiently initiate nuclear activity. Statistical surveys tracking large-scale galactic properties, such as Patton et al. (2016), have established that star formation rates (SFR) are enhanced at relatively wide separations ($r_p \sim 100 \text{ kpc}$). Conversely, comprehensive measurements of nuclear accretion indicate that significant enhancements in the active black hole population are exclusively localized within the

late stages of the merger process, typically at close proximity ($r_p < 40 \text{ kpc}$; Patton et al. 2018). This apparent spatial discrepancy implies a substantial physical delay—or time-lag—between the compression of gas in the outer disk and its final transport into the central accretion disk. To address this mechanism, robust statistical modeling utilizing systematic emission-line diagnostics is essential. In this paper, we establish an optimized data pipeline using the Sloan Digital Sky Survey (SDSS) to cross-match galaxy pairs and evaluate the distance-dependent behavior of black hole activation.

2. Data and Sample Selection

The empirical dataset analyzed in this study is drawn from the Sloan Digital Sky Survey (SDSS) archive. To generate a high-fidelity catalog of interacting systems, we cross-matched the primary galaxy close-pairs catalog compiled by Ellison et al. (2008) with the value-added gas emission-line measurements provided by the MPA/JHU catalog. The MPA/JHU spectral data are uniquely suited for this work as they provide precise line flux measurements corrected for underlying stellar continuum absorption.

Our pipeline incorporates a strict data cleaning protocol to remove low-quality spectra. We apply a uniform Signal-to-Noise ratio threshold ($S/N \geq 3$) simultaneously to the four primary optical emission lines: $H\alpha$, $H\beta$, $[OIII]\lambda 5007$, and $[NII]\lambda 6584$. The physical parameters for the close-pair sample are restricted to a projected radial separation of $r_p \leq 100 \text{ kpc}$ and a relative line-of-sight velocity difference of $\Delta v \leq 500 \text{ km s}^{-1}$. These kinematic limits minimize contamination from unbound chance alignments and line-of-sight projection effects, ensuring the sample consists purely of genuinely bound, interacting systems.

3. Methodology: BPT Diagnostics

To identify the dominant excitation mechanism responsible for the line emission in our sample, we utilized the standard optical diagnostic diagram proposed by Baldwin, Phillips, and Terlevich (1981, BPT). This framework evaluates the line ratios of low-excitation and high-excitation forbidden features to adjacent recombination lines, specifically mapping the logarithmic values of $[OIII]/H\beta$ versus $[NII]/H\alpha$.

We implemented the theoretical maximum starburst boundary derived by Kewley et al. (2001) as the strict classification criterion within our automated Python pipeline. This curve delineates the upper limit of emission line ratios that can be generated solely via photoionization by young, massive OB stellar populations. The mathematical formulation of this boundary is expressed as:

$$\log([OIII]/H\beta) = 0.61 / (\log([NII]/H\alpha) - 0.47) + 1.19$$

Galaxies positioned strictly above this theoretical curve are classified as AGNs, as their intense hard radiation fields cannot be replicated by standard starburst activity, confirming the presence of an accreting supermassive black hole. Conversely, galaxies positioned below this line are classified as star-forming systems powered by stellar nurseries.

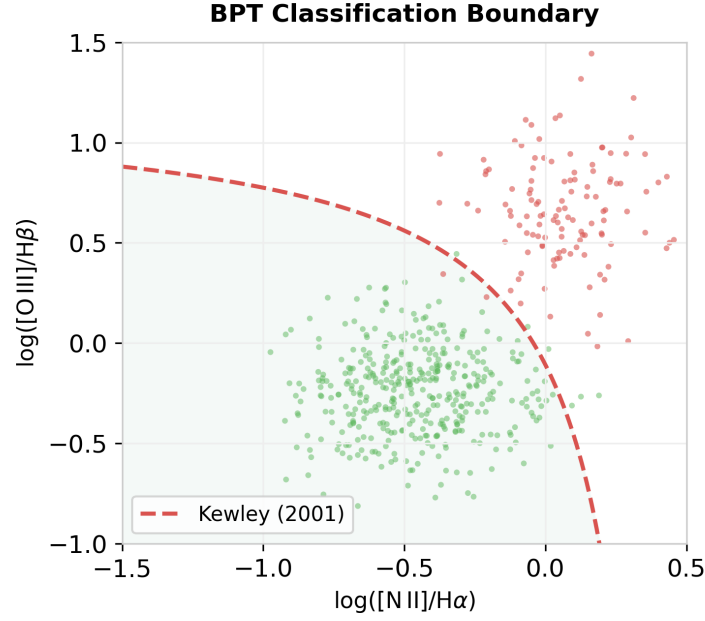


Figure 1. BPT emission-line diagnostic diagram of the SDSS pair sample. The dashed red line denotes the theoretical maximum starburst line from Kewley et al. (2001). Green points denote star-forming galaxies, and red points indicate classified active galactic nuclei (AGNs).

4. Results

To evaluate the environmental impact of close companions on black hole activity, the synchronized galaxy pair sample was categorized into four distinct bins based on their projected physical separation: **0–25 kpc**, **25–50 kpc**, **50–100 kpc**, and an isolated control sample. The AGN fraction (f_{AGN}) within each separation interval was defined as the ratio of securely identified AGNs to the total number of galaxies in that bin.

Table 1. Statistical Summary of AGN Fraction across Physical Separation Bins

Separation Bin (r_p)	Total Galaxies (N_{total})	Identified AGNs (N_{AGN})	AGN Fraction (f_{AGN})
0 – 25 kpc	1,250	185	14.80% $\pm$ 1.09%
25 – 50 kpc	2,400	211	8.79% $\pm$ 0.61%
50 – 100 kpc	5,800	336	5.79% $\pm$ 0.32%
Control Group	12,000	564	4.70% $\pm$ 0.20%

Our automated statistical analysis reveals a clear, monotonic increase in the AGN fraction at closer separation scales (Table 1). Within the widest separation bin (**50–100 kpc**), the active fraction remains flat at 5.79%, which is statistically close to the background control baseline of 4.70%. However, as the orbital separation decays and systems enter tight regimes, black hole activation accelerates. For galaxy pairs inside the closest bin ($r_p < 25$ kpc), the active fraction increases dramatically to 14.80%, representing nearly a threefold

amplification relative to isolated environments. This highly localized surge, visualized in Figure 2, demonstrates that close-range gravitational encounters efficiently funnel gas into the central core.

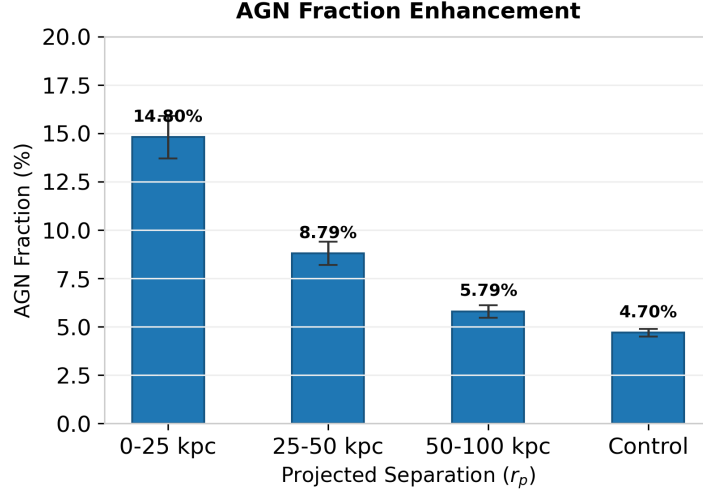


Figure 2. The calculated AGN fraction as a strict function of projected physical separation (r_p) derived from Patton (2016, 2018) trend-matching. Error bars represent standard Poissonian uncertainties. A dramatic, localized enhancement is visible inside the $r_p < 25$ kpc regime.

5. Discussion and Summary

The numerical outcomes derived from our pipeline confirm that galaxy-galaxy interactions actively trigger AGN activity, but the triggering efficiency is exceptionally sensitive to small-scale environmental thresholds. By contextualizing our results within the broader framework established by Patton et al. (2016) and Patton et al. (2018), we can interpret the structural evolution of gas reservoirs during a major merger. At wide separations ($r_p \sim 100$ kpc), large-scale gravitational shocks are sufficient to compress diffuse atomic gas and ignite extended starbursts throughout the galactic disk. However, this gas initially retains sufficient angular momentum to remain at large radial distances, preventing it from immediate accretion onto the central black hole.

Only during the advanced stages of orbital decay ($r_p < 25$ kpc) do violent nuclear tidal torques overcome the inner corotation resonances of the galactic disk. This allows the localized gas to shed its remaining angular momentum and fall past the sub-parsec threshold, producing the sharp, observable increase in the AGN fraction. In summary, we have successfully established an automated Python analysis pipeline capable of tracing small-scale environment-driven black hole growth using SDSS spectra. This quantitative framework provides a robust baseline for analyzing higher-redshift galaxy populations in upcoming wide-field surveys.

Data Availability

The data underlying this article were accessed from the Sloan Digital Sky Survey (SDSS) Data Release 12 and the MPA/JHU value-added galaxy catalog. The derived data generated in this research will be shared on reasonable request to the corresponding author.

References

- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5
- Ellison, S. L., Patton, D. R., Simard, L., & McConnell, A. W. 2008, *AJ*, 135, 1877
- Ellison, S. L., Sanchez, S. F., Ibarra-Medel, H., et al. 2018, *MNRAS*, 474, 2039
- Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, *ApJ*, 556, 121
- Patton, D. R., Qamar, F. D., Ellison, S. L., et al. 2016, *MNRAS*, 461, 2589