The Astrophysical Journal, 999:35 (14pp), 2026 March 1 https://doi.org/10.3847/1538-4357/ae3d9b

Anomalous Narrow Line Seyfert I Galaxies from SDSS DR17

Arihant Tiwari[1], Rachana[1,2], M. Vivek[1], and Suvendu Rakshit[3]
[1] Indian Institute of Astrophysics, 100 Feet Road, Santhosapuram, 2nd Block, Koramangala, Bengaluru, Karnataka, India; tivariarihant@gmail.com
[2] Joint Astronomy Programme, Department of Physics, Indian Institute of Science, Bengaluru 560012, India
[3] Aryabhatta Research Institute of Observational Sciences, Nainital 263001, Uttarakhand, India
Received 2025 September 16; revised 2026 January 10; accepted 2026 January 22; published 2026 February 23

Abstract

We present an analysis of 22,656 narrow-line Seyfert 1 galaxies (NLSy1s) from the Sloan Digital Sky Survey (SDSS) DR17 ($0.1 \leqslant z \leqslant 0.9$), identifying a sample of spectroscopically anomalous sources. These anomalies were detected via the Spectroscopic Quasar Anomaly Detection (SQuAD) algorithm, which employed principal component analysis and hierarchical *k*-means clustering. Various physical diagnostic analyses were performed such as color excess ($E_{(B-V)}$) calculations, Wide-field Infrared Survey Explorer color analysis, probing [O III] equivalent width as an inclination indicator, the Baldwin–Phillips–Terlevich (BPT) diagram, and eigenvector 1 diagram. We detected 620 anomalous NLSy1s classified into two groups. The first is 246 red NLSy1s, exhibiting host-galaxy-dominated spectra with a low-luminosity active galactic nucleus (AGN) core revealed by the emission line widths. The other set consists of 374 blue NLSy1s, which are highly luminous galaxies with enhanced AGN activity, bluer continua as compared to a typical NLSy1, and have stronger Fe II emission. Finally, we identify a third group of 257 outliers, identified as intermediate Seyferts, which are a class of Seyfert galaxies identified by composite emission profiles and extremely strong emission lines paired with virtually no continuum. These sources also exhibit rare and high ionization emission lines unseen in any other NLSy1 spectra (e.g., [Ne v] λ3345, Ne v λ3426, Ne III λ3869, etc.). We conclude that the differentiating factor between red and blue NLSy1s is not dust obscuration or orientation effects, but intrinsic distinction in AGN activity. The resulting sample is presented as a value-added catalog.

Unified Astronomy Thesaurus concepts: AGN host galaxies (2017); Active galaxies (17); Seyfert galaxies (1447); Galaxy spectroscopy (2171); Catalogs (205); Quasars (1319); Active galactic nuclei (16); Supermassive black holes (1663)

Materials only available in the online version of record: machine-readable tables

1. Introduction

Active galactic nuclei (AGNs) are among the most energetic and dynamic objects in the Universe, powered by the accretion of matter onto supermassive black holes (SMBHs) located at the centers of galaxies (I. Shlosman et al. 1990). These luminous sources display a wide range of observational characteristics, leading to their classification into various subtypes based on their spectral features, luminosity, and variability. Quasars, the most luminous class of AGNs, typically display both broad and narrow emission lines, along with a continuum that follows a power law, $F(\nu) \propto \nu^{-\alpha}$, with a spectral index of approximately 0.44 between 1300 and 5000 Å. This index steepens to about 2.45 at wavelengths beyond ~5000 Å, a spectral break that is primarily attributed to increasing host galaxy contamination at lower redshifts (D. E. Vanden Berk et al. 2001). Moreover, the evolution of AGNs follows the quasar luminosity function (QLF), which peaks at redshifts $z \sim 2$–3, tracing the peak era of black hole growth in the Universe. This evolution provides crucial insights into the role of quasars in cosmic feedback processes and the coevolution of galaxies and their central black holes (Q. Yu & S. Tremaine 2002).

Among these various AGN subclasses, narrow-line Seyfert 1 galaxies (NLSy1s) represent a particularly distinct and intriguing population. Spectroscopically, they are characterized by a relatively narrow "broad" component of the Hβ emission line (FWHM < 2000 km s^{-1}), weak [O III] emission with a flux ratio of [O III]/H$\beta \leqslant 3$, and strong Fe II emission, typically with Fe II/H$\beta_{\rm tot} > 0.5$ (D. E. Osterbrock & R. W. Pogge 1985; R. W. Goodrich 1989; M. P. Véron-Cetty et al. 2001). The prominent Fe II emission in AGNs is often found to be anticorrelated with both the strength of [O III] and the width of the broad Balmer lines, suggesting a complex interplay between the ionized gas and the central engine (T. A. Boroson & R. F. Green 1992). The distinct observational characteristics of NLSy1s are believed to stem from the unique physical conditions in their central engines. Black hole masses in NLSy1s are typically estimated using the FWHM of the broad Hβ line and the continuum luminosity via single-epoch methods calibrated against reverberation mapping (S. Kaspi et al. 2000, 2005; B. M. Peterson 2015). In these sources, the broad Hβ profiles are generally better represented by Lorentzian rather than Gaussian shapes. The second-order moment, often used for Gaussian profiles (E. Dalla Bontà et al. 2020), is therefore not suitable in this case. These estimates suggest that NLSy1s host relatively low-mass SMBHs, generally in the range of 10^6–$10^8\,M_\odot$ (B. M. Peterson 2011), and accrete at high rates, often close to the Eddington limit (K. A. Pounds et al. 1995).

Due to their low black hole masses and high accretion rates, NLSy1s exhibit some of the most extreme multiwavelength signatures among Type 1 AGNs, particularly in the X-ray and UV regimes. In the X-ray regime, NLSy1s frequently exhibit steep soft X-ray spectra ($E < 2$ keV), a pronounced soft excess, and rapid, large amplitude variability, often exceeding 40%, on timescales ranging from hours to years (T. Boller et al. 1996; K. M. Leighly 1999a, 1999b; D. Grupe 2004). These properties

are typically more extreme than those observed in their broad-line Seyfert 1 galaxy (BLSy1) counterparts. Spectral complexity is further enhanced by signatures of ionized winds and outflows (T. A. Boroson 2002; A. J. Blustin et al. 2005). Multi-wavelength monitoring has revealed correlated variability across X-ray, UV, and optical bands, supporting a scenario where fluctuations in the inner accretion flow modulate reprocessed emission at longer wavelengths (e.g., D. Grupe et al. 2010). While some NLSy1s are X-ray weak, emitting less X-ray flux relative to their optical–UV output, others, especially radio-loud or jet-dominated sources, exhibit strong nonthermal X-ray emission from synchrotron or inverse Compton processes (M. Orienti et al. 2015; S. Yao & S. Komossa 2023; S. C. Chaudhary & R. Prince 2025).

The presence of strong nonthermal X-ray emission in some NLSy1s is closely tied to their radio and high-energy behavior. A subset of NLSy1s, classified as radio loud, exhibit compact radio cores and flat or inverted spectra, with some sources showing signs of relativistic jet activity such as high brightness temperatures or apparent superluminal motion. In several cases, they have also been detected at γ-ray energies by the Fermi γ-ray space telescope (A. A. Abdo et al. 2009; L. Foschini et al. 2011; H. Yang et al. 2018), confirming the presence of powerful jets. Intriguingly, some of these jet-dominated NLSy1s host relatively massive black holes, with estimates reaching $\sim 10^7$–$10^8\,M_\odot$, i.e., at the upper end of the typical NLSy1 range (10^6–$10^8\,M_\odot$; L. Foschini et al. 2011), and comparable to those of blazars and BLSy1s (G. Calderone et al. 2013; R. D. Baldi et al. 2016). Their broadband spectral energy distributions (SEDs) often resemble those of blazars, with synchrotron and inverse Compton peaks spanning the radio to γ-ray regime. However, L. Foschini (2020) pointed out that these analyses have several limitations and suggested that some jetted NLSy1s might still host relatively low-mass black holes. In the UV regime, NLSy1s display distinctive spectral features that further underscore their complexity. Many NLSy1s exhibit redder UV continua compared to typical quasars, often attributed to intrinsic dust extinction or line-of-sight absorbers (A. Constantin & J. C. Shields 2003). Broad, blueshifted absorption lines in high-ionization species such as C IV and N V have been observed in several cases, signaling the presence of radiation-driven disk winds or warm absorbers (WAs; e.g., K. M. Leighly 2004; K. M. Leighly & J. R. Moore 2004; F. Tombesi et al. 2013). Highly ionized X-ray ultrafast outflows have also been detected in some NLSy1s (J. N. Reeves et al. 2023). These features provide valuable diagnostics of the coupling between the accretion disk and the outflowing gas, and, when combined with X-ray WA signatures, offer a multiphase picture of AGN feedback in NLSy1s.

NLSy1s have also been extensively studied in the IR band. AGN emission in this regime is primarily driven by thermal radiation from the dusty circumnuclear torus, which absorbs high-energy UV and optical photons from the accretion disk and reemits them at longer wavelengths. This reprocessed mid-IR emission serves as a powerful tracer of nuclear activity, particularly in obscured systems where optical and UV light are significantly attenuated by dust. Using data from the Wide-field Infrared Survey Explorer (WISE), S. Rakshit et al. (2019) analyzed 492 NLSy1s and reported a mean $W1-W2$ color of 0.99 ± 0.18 mag. For general AGN selection, D. Stern et al. (2012) and R. J. Assef et al. (2013) proposed $W1-W2$ thresholds of $\geqslant 0.8$ and $\geqslant 0.662$, respectively, as robust criteria to distinguish AGNs from inactive galaxies, since redder $W1-W2$ colors trace stronger hot-dust emission from the torus. B. Mingo et al. (2016) further refined these diagnostics by combining $W1-W2$ and $W2-W3$ color cuts, identifying star-forming Seyfert galaxies with $W1-W2>0.5$ and $0.6<W2-W3<3.4$. These red mid-IR colors reflect strong thermal emission from the AGN-heated torus, while lower values typically indicate increased host galaxy contamination. As a result, mid-IR color selection remains effective for uncovering both unobscured (Type 1) and heavily obscured (Type 2) AGNs, including those often missed in optical and X-ray surveys due to dust extinction or weak nuclear signatures.

Due to the diversity of emission mechanisms and complex physical conditions in their nuclear regions, some NLSy1s exhibit unusual or anomalous spectroscopic behavior. For example, W. Zhang et al. (2022) discovered a late-time X-ray flare and anomalous emission-line enhancement after the nuclear optical outburst. C. L. Steinhardt & J. D. Silverman (2013) present a sample of quasars with anomalous Hβ profiles, featuring several NLSy1s. N. M. Nagar et al. (2002) presented strong evidence that NGC 5506 is an NLSy1, despite previously being classified as an intermediate or Type 2 Seyfert. This galaxy displays several anomalous properties, including being optically obscured, hosting both a Type 1 AGN and a nuclear water vapor megamaser, and having a relatively high black hole mass for an NLSy1. These examples underscore the need for systematic identification of outliers within the NLSy1 population that can provide in situ conditions to study certain properties of NLSy1s in isolation for a deeper understanding.

This project leverages the Spectroscopic Quasar Anomaly Detection (SQuAD) algorithm (A. Tiwari & M. Vivek 2025) to identify anomalous NLSy1s based on their spectroscopic features and to perform detailed follow-up analyses of the most intriguing cases. This work represents the second installment in a series of studies aimed at systematically uncovering and characterizing spectroscopic outliers in AGN populations. In this paper, we focus specifically on anomalies within the rest-frame optical spectra of a curated sample of NLSy1s drawn from Sloan Digital Sky Survey (SDSS) Data Release 17 (DR17; Abdurro'uf et al. 2022), as compiled by V. S. Paliya et al. (2024). V. S. Paliya et al. (2024) classified the NLSy1s solely on the basis of the definition markers, i.e., Hβ FWHM and the [O III] to Hβ ratio, making it possible for contaminants such as intermediate Seyferts and quasars to be included in the catalog. We expect to utilize the efficiency of the SQuAD algorithm in objectively picking out abnormal spectra to extract such sources too along with the any anomalous NLSy1s. The paper is organized as follows: In Section 2, we describe the sample selection, and the application of the SQuAD algorithm to extract the anomalous NLSy1s. Section 3 presents our main findings, with visual representation of the detected anomalies and their observed spectral features. In Section 4 we discuss various physical diagnostic procedures utilized to reveal the physical properties of the anomalies and their implications. Finally, Section 5 summarizes the findings and traces out the follow-up study plans.

2. Methodology

2.1. Data

The galaxies analyzed in this project are obtained from the V. S. Paliya et al. (2024) catalog, which contains 22,656 sources identified as NLSy1s by carrying out a detailed decomposition of optical spectra of quasars and galaxies from SDSS DR17. The sample contains sources in the redshift range $z \in [0.13, 0.90]$, the spectra for which were obtained using the SDSS and Baryon Acoustic Oscillation Spectroscopic Survey (BOSS) spectrographs from the SDSS survey, covering wavelengths from 3600 to 10400 Å (with good throughput between 3650 and 9500 Å) and a resolution of 1560–2270 in the blue channel and 1850–2650 in the red channel (S. A. Smee et al. 2013). The redshift range translates to a common wavelength window of 3600 to 5400 Å, capturing six prominent emission lines in a typical NLSy1 spectrum: [O III] $\lambda\lambda$5007, 4959, Hβ λ4861, H$\gamma\lambda$4340, Hδ λ4102, [Ne III] λ3869, and [O II] λ3727 (e.g., E. O. Schmidt et al. 2016; A. Constantin & J. C. Shields 2003). The spectra are analyzed as per the SQuAD algorithm to detect, group, and characterize the anomalous NLSy1s. The following steps were performed chronologically to obtain the anomalous objects.

2.2. SQuAD

During this work, we followed the SQuAD algorithm exactly, with no additional modifications. Details of the algorithm can be found in A. Tiwari & M. Vivek (2025). We briefly summarize the algorithm below as applied here.

1. *Preprocessing*. The raw spectra obtained from SDSS are resampled with a 2 Å binning to bring all the spectra to the same wavelength grid and simultaneously reduce the size of the spectral array by a factor of 2. These resampled spectra are smoothed using a Savitzky–Golay filter (A. Savitzky & M. J. E. Golay 1964), which helps reduce the noise and unwanted artifacts. Once smoothed, all spectra are max normalized to bring the flux range to $[-1, 1]$ to ensure uniformity, needed for the anomaly detection algorithm. These processed spectra are then fed into the anomaly detection pipeline.
2. *Principal component analysis (PCA)*. We apply PCA (H. Hotelling 1933) with 20 components to reduce the size of each spectrum from a 900-dimensional wavelength–flux vector to 20-dimensional vector in the PCA hyperspace. The number of components was chosen as to achieve more than 95% cumulative explained variance, which was achieved with 20 components accounting for 95.2% total explained variance.
3. *Clustering*. We then employed k-means clustering (J. MacQueen 1967) to categorize the NLSy1s based on the Euclidean separation between the PCA eigenvalues. k-means clustering requires a predefined number of clusters (k) into which the dataset is supposed to be divided. We calculated the optimum number of clusters using elbow method on the sum of squared errors and silhouette coefficients. Based on this $k = 3$ was obtained as the optimum number of clusters for our dataset.
4. *Anomaly detection*. Once clustered, we computed the Euclidean distance of each data point from its respective cluster centroid and analyzed the resulting distribution. Statistical tests confirmed a nonnormal, long-tailed spread, best described by a Gumbel distribution (S. Nadarajah & S. Kotz 2004). Given this, we used the interquartile range (IQR; J. W. Tukey1977) method to identify outliers, marking data points beyond $Q3 + 1.5 \times IQR$ as anomalies. This resulted in 944 outliers, which we refer to as "anomalous galaxies" here onward.
5. *Anomaly grouping*. On visual inspection, it was observed that the detected anomalies exhibited repeating trends. Hence, to further categorize these anomalies, we reapplied k-means clustering solely on the outliers or anomalies detected in the previous step. Using a 20-component PCA, which captured 97% of the variance, and the elbow method, we determined an optimal cluster number of three ($k = 3$). This step ensured that similar spectral anomalies were grouped together, independent of their original cluster assignment.

A detailed flowchart of the SQuAD algorithm can be found in Figure A.1 in the appendix of A. Tiwari & M. Vivek (2025).

3. Results

Hereafter, the three classifications done by the initial k-means clustering applied to the complete dataset will be referred to as clusters, while the classifications of the detected anomalous spectra from the subsequent k-means clustering will be called groups.

3.1. Clusters

The initial k-means clustering was performed on the PCA-transformed dataset, and the resulting clusters were visualized using a two-dimensional projection of the PCA hyperspace. Specifically, we plotted the coefficients of the first two principal components, which capture the maximum variance in the data. The left panel of Figure 1 displays the full quasar sample, segmented into three distinct clusters by the k-means algorithm. The identified spectral anomalies (see Section 4) are overlaid as yellow star markers in the same figure for visual comparison.

To investigate the distinguishing features among the clusters, we computed and normalized the mean spectrum for each cluster (see right panel of Figure 1). The resulting mean spectra reveal that the primary differentiating factor is the strength of the emission lines, rather than the line widths or continuum shape. Cluster 1 exhibits the most prominent emission lines (e.g., average [O III] equivalent width (EW) $\approx$36.26, Hβ EW $\approx$ 91.15), while cluster 3 shows the weakest (e.g., average [O III] EW $\approx$ 2.35, Hβ EW $\approx$ 71.92), with cluster 2 falling in between (e.g., average [O III] EW $\approx$ 10.55, Hβ EW $\approx$ 79.28). This trend is particularly pronounced in the [O III] and [O II] emission lines. In cluster 1, the [O III] line is notably narrow and prominent, indicative of strong and possibly low-velocity narrow-line region (NLR) emission. A similar but less pronounced pattern is observed in the Balmer lines. Interestingly, the trend is reversed for the Fe II emission, with cluster 3 displaying the strongest Fe II emission features and cluster 1 the weakest. This inverse relationship between Fe II strength and high-ionization narrow-line emission (e.g., [O III]) is consistent with known AGN spectral sequences (e.g., eigenvector I of T. A. Boroson & R. F. Green 1992), and may reflect differences in the accretion rate, ionizing continuum shape, or orientation effects across

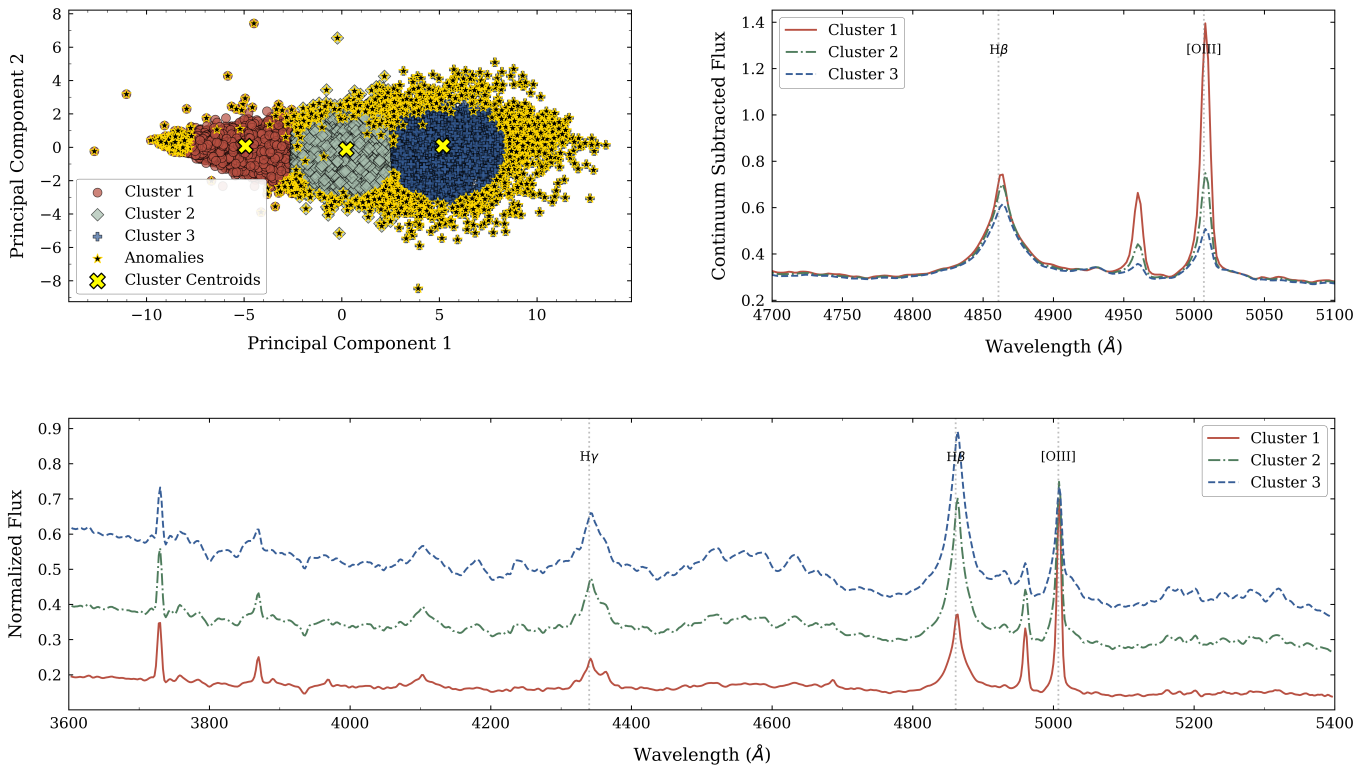

Figure 1. Top left: a 2D projection (PC1 vs. PC2 coefficients) of the dataset. The dataset is divided into three clusters (cluster 1: brown, cluster 2: green, and cluster 3: blue) within the 20-dimensional PCA hyperspace, using k-means clustering. Sources classified as anomalous after applying a $Q3 + 1.5 \times IQR$ threshold are shown as golden scatter points overlaid on top of the cluster members. The cluster centroids are marked by yellow crosses. Top right: continuum-subtracted mean composite spectrum for each cluster to highlight the differences in emission-line strength. Bottom: mean composite spectrum for each cluster of the dataset. The color of each spectrum corresponds to the color of the cluster as shown in the left panel.

the clusters (e.g., P. Marziani et al. 2001; Y. Shen & L. C. Ho 2014).

The number of members and anomalies in each cluster is given in Table 1.

3.2. Groups

As outlined in Section 2.2, we performed a second k-means clustering exclusively on the subset of anomalous galaxies. This clustering yielded three distinct groups, visualized in the left panel of Figure 2 as a two-dimensional scatter plot based on the coefficients of the first two principal components of PCA. The right panel of the same figure displays the mean composite spectra corresponding to each of the three anomaly groups, with colors matching their cluster assignments. In contrast to the relatively similar continua of the mean spectra of the clusters (see Figure 1), the composite spectra of the anomaly groups exhibit substantial diversity in both continuum slope and emission-line characteristics.

3.2.1. Group 1

The first group is characterized by galaxies exhibiting exceptionally strong emission lines superimposed on an almost flat continuum. These objects occupy the leftmost region in the PCA component space, shown as orange scatter points in the left panel of Figure 2. Their spectra are visually striking—dominated entirely by sharp and strong atomic emission lines with little to no underlying continuum emission (see orange spectra in the right panel of Figure 2). These spectral features are characteristic of intermediate Seyfert galaxies, a detailed analysis of which is presented in Section 4. The absence of a prominent continuum enhances the visibility of numerous weak atomic lines that are typically obscured in conventional AGN and NLSy1 spectra, making this group particularly unique. The spectral shape and continuum emission-line contrast features are remarkably similar to the C IV peakers identified by A. Tiwari & M. Vivek (2025).

Notably, this group is mildly affected by instrumental artifacts, most commonly in the form of sharp, narrow spikes that resemble strong, unresolved emission lines. These features —likely due to cosmic-ray hits or processing glitches—often mimic the narrow [O III] profile that is characteristic of this group, leading to coincident grouping. However, these contaminant spectra were few in number and were manually identified and removed during visual inspection. In total, 257 such galaxies were identified, constituting approximately 29.3% of the full set of detected outliers.

Table 1
Number of Members and Detected Anomalies in Each Cluster

Dataset	Cluster 1	Cluster 2	Cluster 3
Members	7244	6518	8746
Anomalies	312	384	298

3.2.2. Group 2

The second group comprises galaxies characterized by a noticeably bluer optical continuum compared to the average (mean composite) NLSy1 spectrum, along with a significantly

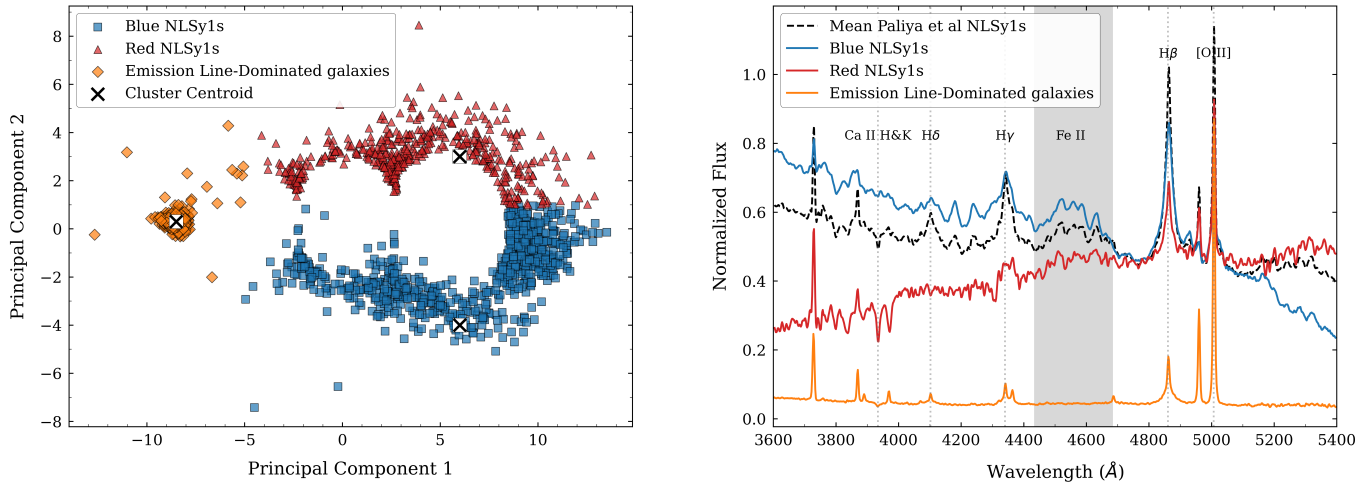

Figure 2. Left: a 2D projection (principal component (PC1) vs. second principle component (PC2) coefficients) of the anomalous galaxies as divided into three groups (group 1: orange, group 2: blue, and group 3: red), in the 20-dimensional PCA hyperspace, with the second k-means clustering applied only on the anomalous galaxies. The groups are characterized by further analysis (see Section 4). Right: mean composite spectrum of each anomaly group. The colors of the spectra correspond to the color of their group in the left panel. The black dashed spectrum is the mean composite of all the galaxies in the V. S. Paliya et al. (2024) catalog. The red, blue, and black spectra are arbitrarily normalized for better comparison.

enhanced optical Fe II emission bump. These objects populate the lowermost region of the PCA component space, as indicated by the blue scatter points in the left panel of Figure 2. The PC2 coefficients primarily characterize spectral reddening effects, as demonstrated by A. Tiwari & M. Vivek (2025). Hence, their presence at the low end of the PC2 axis reflects a lower degree of continuum reddening or intrinsically bluer spectral slopes. Owing to these characteristics, we refer to this population as blue NLSy1s.

Spectroscopically, the blue NLSy1s exhibit the strongest optical Fe II emission, with an average R4570 ≈ 0.647 as compared to $\sim$0.532 for red NLSy1s and $\sim$0.103 for intermediate Seyferts. The blue NLSy1s also have the weakest [O III] λ5007 emission among the three anomaly groups (see blue spectrum in the right panel of Figure 2) with average [O III] EW ≈ 5.12 Å as compared to $\sim$8.28 Å for red NLSy1s and $\sim$165.35 Å for intermediate Seyferts. This inverse correlation between Fe II and [O III] strength is a hallmark of the so-called eigenvector 1 (EV1) parameter space, which captures one of the principal axes of quasar spectral diversity and is commonly linked to variations in Eddington ratio, orientation, and broad-line region (BLR) structure (T. A. Boroson & R. F. Green 1992). In total, 374 such anomalies were identified, accounting for 42.64% of the detected anomalies and constituting the largest group in the anomaly sample.

3.2.3. Group 3

The third group consists of NLSy1s that exhibit a markedly redder optical continuum compared to the average NLSy1 spectrum. These objects are concentrated in the upper region of the PCA projection (see red scatter points in the left panel of Figure 2), corresponding to elevated PC2 values. High values along this principal component reflect a steep positive continuum slope, consistent with significant reddening. We refer to this class hereafter as red NLSy1s.

The mean spectrum of the red NLSy1s reveals moderate [O III] emission and relatively weak Fe II features (see red spectrum in the right panel of Figure 2), positioning them between the blue NLSy1s and the intermediate Seyferts in terms of narrow-line strength and iron emission. However, their most distinctive attribute is a reddened continuum, which gives rise to their classification. In total, we identified 246 such red NLSy1s, accounting for 28.05% of the detected anomalies.

The number of anomalies in each group is given in Table 2.

Table 2
Number of Anomalies in Each Group after Manual Removal of Bad Spectra Caused Due to Instrumental Artifacts

Group	Group 1	Group 2	Group 3
Count	257	374	246

4. Discussion

The V. S. Paliya et al. (2024) catalog was curated by selecting sources solely on the basis of quantitative definition markers of an NLSy1. This makes it possible for contaminants to seep into the sample that are not actually NLSy1s but fall into the numerical definition. Since we are dealing with outliers from the V. S. Paliya et al. (2024) catalog, it becomes especially important to confirm the true classification of these sources before proceeding with an in-depth spectral and physical analysis. By careful spectral decomposition of the detected outliers using PyQSOFit (H. Guo et al. 2018) we found that two of the three anomaly groups, i.e., red NLSy1s and blue NLSy1s are truly NLSy1s. We define a true NLSy1 to be a Seyfert 1 galaxy with an Hβ FWHM less than 2000 km s^{-1}, an [O III] to Hβ ratio < 3, strong Fe II emission, and (in our case) a broad Hβ emission profile with negligible to no narrow component. Group 1 members, on the other hand were found to be intermediate Seyfert galaxies. Their nature was confirmed by (i) the presence of a composite (narrow + broad) profile of the Hβ and Hα emission lines (R. D. Cohen 1983) and (ii) an Hβ narrow to Hβ broad flux ratio of $\approx$0.2 as compared to average ratio $\sim$0.04 for the total sample. We also followed the classification scheme provided by M. Whittle (1992) and found that out of 247 sources, 47 fall in the Seyfert 1.2 category while the rest can be classified as

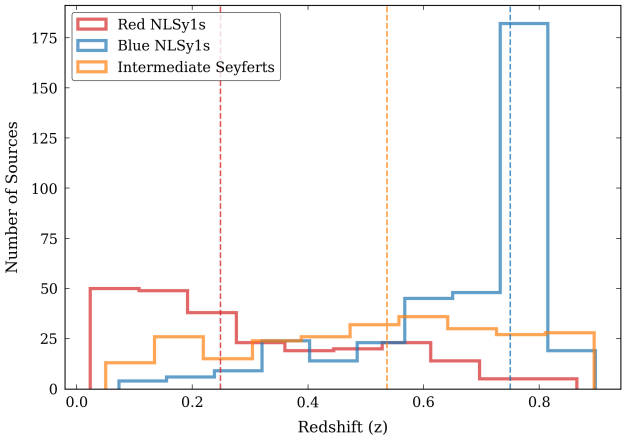

Figure 3. The redshift distributions of red (red) and blue (blue) NLSy1s, and intermediate (orange) Seyferts. The median redshift for each class is marked by a dashed line.

Seyfert 1.5. The Seyfert type calculation as given by H. Netzer (1990) also provided a similar distribution, confirming the intermediate Seyfert identity of Group 1 members. Sources with a low signal-to-noise ratio, for which we were unable to obtain a reliable classification, were left unmarked but are still given in the catalog for future observations.

To gain a deeper insight into the nature of the three groups, we carried out a comprehensive investigation of their physical properties. This includes an examination of their redshift distributions to probe potential selection effects and cosmological trends, estimation of color excess $E(B-V)$ as a proxy for internal reddening, IR color diagnostics to assess dust attenuation and thermal emission properties, and orientation-dependent effects inferred through the EW of the [O III] λ5007 line. All the parameters used in the following analysis, such as $M_{\rm BH}$, bolometric luminosity, line properties, etc., are taken from the V. S. Paliya et al. (2024) catalog. These diagnostics collectively allow us to characterize the underlying physical drivers responsible for the observed spectroscopic diversity among the anomaly classes and to place them in the broader context of AGN unification and evolution.

4.1. Redshift and Magnitude Distributions

Figure 3 illustrates the redshift distribution of the three identified anomaly types. The red NLSy1s predominantly occupy the lowest redshift range, with their numbers declining steadily as redshift increases. In contrast, the blue NLSy1s are more prevalent (heavily concentrated at $z\sim0.8$) at higher redshifts and are virtually absent at $z\sim0$. The intermediate Seyferts span an intermediate redshift range and exhibit a relatively uniform distribution across the redshift space.

The decline in the number of red NLSy1s with increasing redshift is likely not intrinsic to their astrophysical properties but rather a consequence of observational bias, primarily driven by the sensitivity limitations of the SDSS eBOSS spectrograph (R. H. Lupton et al. 2002). The red NLSy1s are observed to exhibit low luminosity (see lower left panel of Figure 4), and as the redshift increases, their apparent brightness diminishes significantly due to cosmological dimming (see V. Calvi et al. 2014) and increasing luminosity distance. Consequently, beyond a certain redshift threshold, many red NLSy1s fall below the SDSS detection limit and are thus underrepresented in the sample. This selection effect creates the appearance of a strong redshift dependence, when in fact simply reflects the survey's inability to detect faint sources at greater distances. Therefore, any inferred evolutionary trends or redshift-dependent population changes among the red NLSy1s must be interpreted with caution, as they may be dominated by this selection bias rather than by any underlying astrophysical evolution.

On the other hand, the scarcity of blue NLSy1s at low redshifts can be understood in the context of the evolution of the QLF (P. F. Hopkins et al. 2007; X. Shen et al. 2020). Quasar activity is known to peak at redshifts around $z\sim2$, often referred to as the "quasar epoch," and declines both toward lower and higher redshifts (e.g., B. J. Boyle et al. 2000; G. T. Richards et al. 2006). As a result, the probability of detecting luminous quasar-like objects such as the blue NLSy1s is significantly higher at intermediate to high redshifts. This naturally leads to a larger observed population of blue NLSy1s at $z\gtrsim0.7-0.8$, with their numbers falling off as redshift decreases. The drop-off toward lower redshifts is therefore consistent with the global decline in quasar activity and abundance, as traced by the QLF.

Visual inspection of the SDSS images for the red NLSy1s and blue NLSy1s revealed that the majority of the red NLSy1s have a visible host galaxy. This is likely facilitated by two factors: (i) the red NLSy1s have a low AGN luminosity making the center of the image not as bright as compared to the host hence making it possible to see the host, and (ii) the red NLSy1s are present primarily at low redshifts, due to which the host galaxy is easily identifiable in the images. Notably, most of these host galaxies have a spiral morphology and exhibited a bar (D. M. Crenshaw et al. 2003; I. Varglund et al. 2023). In contrast, blue NLSy1s have a "blue dot" appearance, consistent with their higher AGN luminosities and larger redshifts, which render host galaxy detection challenging.

Interestingly, the redshift distribution of the intermediate Seyferts is nearly uniform throughout the observed redshift range with a median $z\sim0.6$. Their relative absence at lower redshifts (as compared to red NLSy1s) can be reasonably attributed to the evolution of the QLF, which predicts a decline in AGN activity as one moves away from the peak epoch near $z\sim2$ (G. T. Richards et al. 2006). However, the missing intermediate Seyfert population at higher redshifts ($z>0.6$; like blue NLSy1s) appears to arise from selection effects inherent to the SDSS target selection algorithm. Specifically, the SDSS AGN selection is partly based on the observed $g-i$ color, where a minimum threshold is required for a source to be classified as a quasar candidate (A. D. Myers et al. 2015). As redshift increases beyond $z\sim0.6$, the Hβ and [O III] $\lambda\lambda$4959, 5007 emission-line complex—one of the most prominent features in the intermediate Seyferts spectra—shifts into the i-band filter. This leads to an artificial brightening of the i-band magnitude, which in turn suppresses the $g-i$ color. As a result, some of these objects fall below the quasar classification threshold and are excluded from the SDSS spectroscopic quasar sample. This selection bias offers a compelling explanation for the subtle decline in the intermediate Seyferts population at higher redshifts, despite the intrinsic strength of their emission lines.

The top-left panel of Figure 4 shows the bolometric luminosity distributions of all the classes. The blue NLSy1s

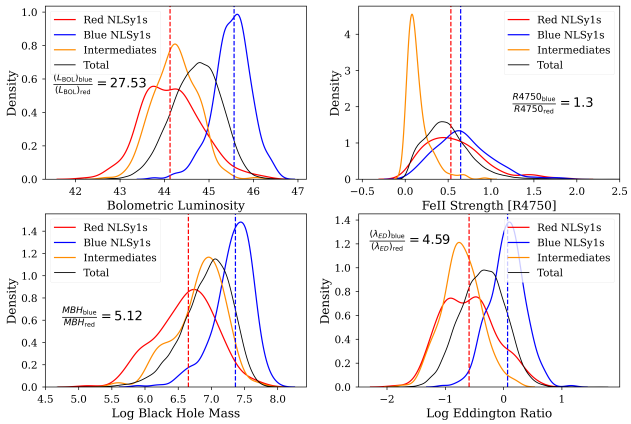

Figure 4. The density-normalized distributions of parameters for the total V. S. Paliya et al. (2024) dataset and the three anomalous groups. Top left: the distribution of Eddington ratios on log scale. Top right: distribution of the Hβ-derived black hole masses on log scale. Bottom left: log-scale distribution of bolometric luminosities. Bottom right: distribution of Fe II emission strengths (R4750). In all the cases, the red NLSy1s lie at the lowest end and blue NLSy1s at the highest end. Intermediate Seyferts are the galaxies with the lowest Fe II emission strength in the complete dataset. The number annotation in each plot indicates the ratio of the respective value for blue to red NLSy1s.

exhibit the highest values, lying mostly at the high end tail of the overall distribution. In contrast, the red NLSy1s and intermediate Seyferts occupy the lower region of the overall distribution and have smaller median bolometric luminosities as compared to the entire sample. This bias in the luminosities of blue NLSy1s is also due to the fact that only high-luminosity sources would be visible at higher redshifts, leading to this distribution. Additionally, due to this difference in the luminosity distribution, the connected properties, i.e., $M_{\rm BH}$ and the Eddington ratio in the lower panels of Figure 4, also show similar trends, where blue NLSy1s occupy the highest end. The Fe II emission strength in the upper-right panel of Figure 4 shows that the red NLSy1s and blue NLSy1s have nearly similar Fe II emission strengths with the blue NLSy1s having a subtle excess. The intermediate Seyferts on the other hand have exceptionally low Fe II emission, occupying the lower-most region of the overall distribution. This is consistent with R. D. Cohen (1983), who used weak Fe II emission as a criterion for selecting intermediate Seyfert samples.

Additionally, as shown in Figure 5, most of the blue NLSy1s exhibit very high values for the absolute B-band magnitude (more than 23) as compared to the other two classes. This indicates that most of the blue NLSy1 sources are more like a quasar and less of a Seyfert galaxy pertaining to their strong luminosity. Red NLSy1s on the other hand, have the lowest B-band magnitudes, which is consistent with their low AGN activity and host galaxy dominated emission. The intermediate Seyferts also lie toward the lower end of the distribution with a median B-band magnitude slightly higher than that of the red NLSy1s.

Given the reddened continuum and lower luminosity of the red NLSy1s, it is tempting to interpret them as dust-obscured AGNs, potentially viewed through the edge of the torus (e.g., Z. Li et al. 2015; X. Pan et al. 2017; S. Zhang et al. 2017). This interpretation was debunked as reddening was found to be caused by host galaxy dominance instead of obscuration (see Section 4.2). On the other hand, the blue NLSy1s —with their bright, blue continua—appear to be face-on sources with minimal extinction or attenuation. Such an interpretation would imply that the observed differences in bolometric luminosity between these classes could be driven primarily by orientation effects or dust obscuration, rather than intrinsic differences in AGN power. To critically assess this possibility, we conduct a series of diagnostic analyses aimed at disentangling the contributions of orientation, dust attenuation, and intrinsic AGN luminosity. These studies serve to clarify whether the observed spectral and luminosity differences among the three anomaly types—red NLSy1s, blue NLSy1s, and intermediate Seyferts—are the result of physical distinctions or observational biases.

4.2. *Color Excess:* E(B − V)

Dust obscuration in AGNs can be quantified using the color excess, $E(B-V)$, which measures the reddening of the continuum due to dust absorption (D. Calzetti et al. 2000; R. Maiolino et al. 2001). In order to calculate the $E(B-V)$ values, all spectra were dereddened to account for Milky Way dust extinction (E. L. Fitzpatrick 1999) using the dust maps by D. J. Schlegel et al. (1998). Then we fit these dereddened spectra $f(\lambda)$ with a template spectrum (the D. E. Vanden Berk et al. 2001 composite) $f_t(\lambda)$ using the parameterization following M. Vivek et al. (2012):

$$f(\lambda) = \left[a f_t(\lambda) + b\left(\frac{\lambda}{\lambda_0}\right)^{\alpha} \right] e^{-\tau_\lambda}. \quad (1)$$

We utilize χ^2 minimization (masking out the emission-line regions, marked as gray in Figure 6) to obtain the optimum values for the parameters a, b, α, λ_0 and τ_λ. The first term in Equation (1) gives the scaling difference, while the second term denotes the difference in spectral index between the observed and template spectra. The optical depth for dust τ_λ is obtained with the Small Magellanic Cloud–like extinction curve (K. D. Gordon et al. 2003). The above described analysis was done using `dust_extinction` Python package (K. D. Gordon 2024).

As illustrated in Figure 7, the distribution of $E(B-V)$ values for the intermediate Seyferts is majorly concentrated around zero and spans up to 0.5. On the other hands, the distribution for the red NLSy1s primarily spans from 0.3 to 0.8, with a median of 0.54. The $E(B-V)$ values for all the blue NLSy1s were calculated to be zero consistently and hence are not shown in Figure 7 for visual clarity. Several previous studies provide context for interpreting these $E(B-V)$ values in terms of host galaxy effects, reddening, and AGN activity. For example, analyses of SDSS galaxies in terms of extinction and star formation histories by G. Kauffmann et al. (2003a) and J. Brinchmann et al. (2004), as well as investigations of dust attenuation in X-ray-selected NLSy1s by A. Caccianiga et al. (2004), show that the $E(B-V)$ range for red NLSy1s is well within that of typical star-forming galaxies. On the other hand, the values for intermediate Seyferts and blue NLSy1s suggest little to no dust obscuration. Broader studies of dust extinction in galaxies across various redshifts further support this interpretation (T. Garn & P. N. Best 2010; V. Buat et al. 2011).

Crucially, the moderate $E(B-V)$ values observed in the red NLSy1s are insufficient to classify them as dust-obscured AGNs. For instance, the well-studied obscured AGN NGC 6860 exhibits a significantly higher $E(B-V) \sim 0.72$, as reported by S. Lipari et al. (1993). More recent work by C. Andonie et al. (2025), based on SED fitting of

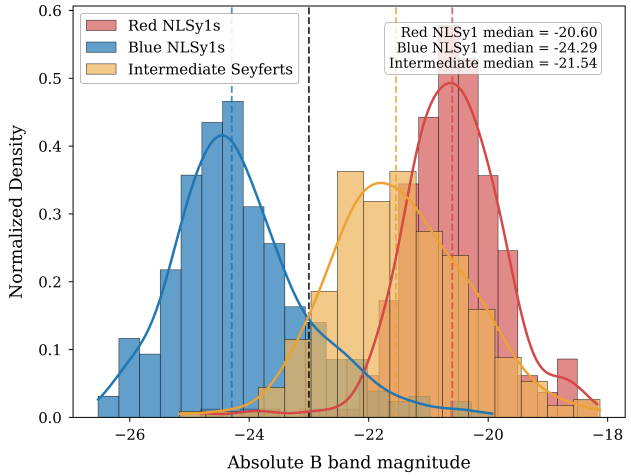

Figure 5. Absolute B-band magnitude distribution for the red and blue NLSy1s, and intermediate Seyferts. The respective median is marked by a dashed line for all classes. The black dashed line marks $M_B = -23$, which is considered to be a distinction limit between quasars and Seyfert galaxies. (V. S. Paliya et al. 2024)

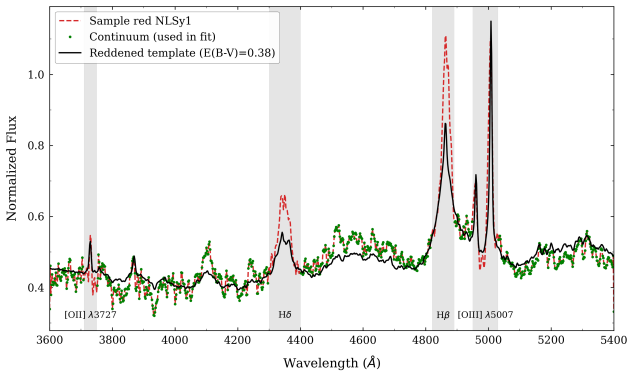

Figure 6. Sample $E(B-V)$ calculation for an NLSy1 spectrum using χ^2 minimization. The red line shows the reddened D. E. Vanden Berk et al. (2001) composite spectrum matched to the Milky Way extinction-corrected NLSy1 spectrum (blue). The gray shaded regions mark the emission-line regions masked during the χ^2 calculation.

obscured AGNs, and theoretical extinction models from A. Schnorr-Müller et al. (2016), suggest that $E(B-V)$ values in the range of 0.7–1 are indicative of significant dust obscuration, while values above one are typical of fully extinguished AGNs. The fact that the red NLSy1s fall just below this threshold supports the interpretation that their reddened appearance is not primarily due to heavy dust attenuation. Hence, the redder appearance is more plausibly attributed to intrinsic properties of the host galaxy, such as a weak AGN continuum being overwhelmed by starlight or modest dust distributed on galactic rather than nuclear scales.

As an additional measure to confirm the nature of the continuum reddening in the red NLSy1s, we compared the mean red NLSy1s spectrum with an SDSS galaxy template spectrum (see Figure 8). The comparison revealed a remarkably good match across the full wavelength range, indicating that the optical continuum of these sources is dominated by host galaxy starlight rather than AGN emission. The most notable deviations from the template arise in the broad emission lines are particularly Hβ and [O III], which are significantly wider in the red NLSy1s spectrum compared to

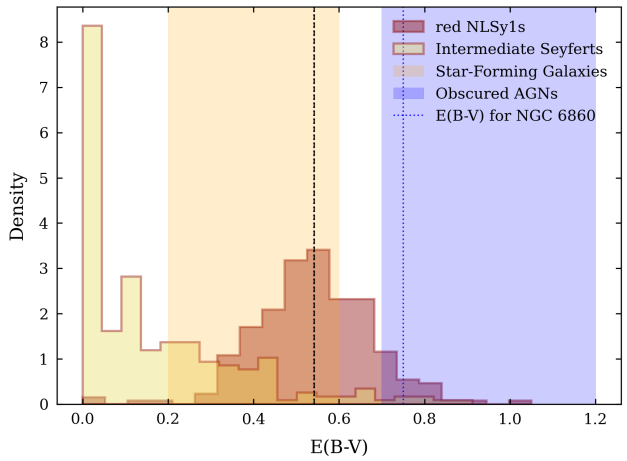

Figure 7. $E(B-V)$ distribution for the red NLSy1s and intermediate Seyfert galaxies. The $E(B-V)$ values for all the blue NLSy1s were calculated to be zero consistently and hence are not shown in the figure for visual clarity. The yellow shaded region indicates the typical $E(B-V)$ range for star-forming galaxies while the blue region indicates dust-obscured AGNs. The black dashed line indicates the median $E(B-V)$ for the red NLSy1s of 0.54, falling inside the star-forming range. The blue dotted line indicates the $E(B-V) = 0.72$ value for NGC 6860 (a known dust-obscured AGN; S. Lipari et al. 1993).

those in the galaxy template. This confirms the presence of a BLR and an NLR and validates the classification of these sources as Type 1 AGNs despite the apparent host-dominated continuum. Such instances where the presence of an AGN is only made evident by the emission lines have been studied in several literature studies, such as B. Grossan et al. (1996) and G. Kauffmann et al. (2003b), and the references therein.

A subtle excess in flux is also observed at the blue end of the spectrum, suggesting a weak but nonnegligible AGN contribution that becomes more apparent at shorter wavelengths where AGN activity overpowers the galactic emission. In this scenario, the AGN is intrinsically weak in the optical band, and the host galaxy effectively overshines the nuclear continuum across most of the spectrum (e.g., L. C. Ho et al. 1997; L. C. Ho 2008; J. Fernández-Ontiveros et al. 2023). This interpretation is further supported by the clear presence of the 4000 Å break (D4000) and the Ca II H and K absorption doublet—prominent features characteristic of a galactic stellar population (see B. M. Poggianti & G. Barbaro 1997), and typically absent or diluted in AGN-dominated spectra. Additionally, the $E(B-V)$ values for the host-galaxy-subtracted spectra for the red NLSy1s were found to be zero consistently, indicating that nearly all the reddening was being contributed by galactic emission.

Therefore, the overall agreement between the galaxy template and the observed continuum, along with the stellar absorption features and moderate $E(B-V)$ values, strongly suggest that the red colors of these NLSy1s are likely not driven solely by dust-induced extinction, but by a combination of weak AGN emission and strong stellar light from the host galaxy. This suggests that red NLSy1s occupy a regime where AGN activity is either intrinsically faint or temporarily diminished, allowing the galaxy's stellar population to dominate the SED. Such systems may represent the low-luminosity end of the NLSy1 population or possibly an evolutionary stage where AGN activity is starting (young

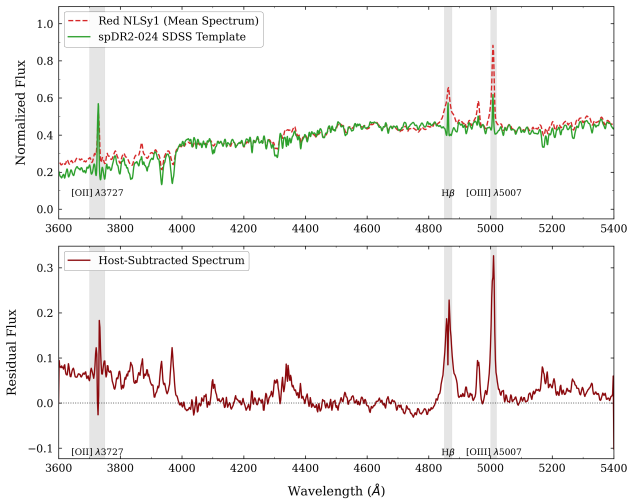

Figure 8. Top: a comparison of the mean red NLSy1s spectrum (red) with a galaxy spectral template (green) from the SDSS. The template is arbitrarily normalized to best match the mean red NLSy1s spectrum. AGN activity is apparent in the subtle continuum excess at the blue end and the broad Hβ and [O III] emission lines. Bottom: host-galaxy-subtracted spectrum for the red NLSy1s.

AGNs), but is not strong enough to wash out the host galaxy spectral signature.

Additionally, it is interesting to note that for the red NLSy1s and blue NLSy1s, the Hβ emission line was found to be better fitted with a Lorentzian profile (two Gaussian components used to fit the broad component while using PyQSOFit; H. Guo et al. 2018) while for the intermediate Seyferts a Gaussian profile was better suited. M. Berton et al. (2020) suggest that a Lorentzian profile is indicative of more pronounced turbulent motion often associated with young sources. This is consistent with the red NLSy1s being galaxies where AGN activity has just began. On the other hand, the Gaussian profile for the intermediate Seyferts is supposed to indicate a class of more evolved objects.

4.3. [O III] as an Inclination Indicator

To evaluate the role of orientation in shaping the observed properties of our anomaly types, we reproduced the analysis of G. Risaliti et al. (2011), where they modeled the distribution of [O III] EWs as a power law shaped by orientation-dependent projection effects. A prerequisite for this analysis is that the [O III] luminosity serves as a reliable proxy for bolometric luminosity. As shown in Figure 9, we find a strong correlation between these quantities (r-value = 0.731, p-value = 1.89×10^{-42}), with a best-fit slope of 0.93, supporting this assumption. Interestingly, the [O III] luminosities of the red NLSy1s and blue NLSy1s slightly underestimate the bolometric luminosity, while for the intermediate Seyferts, the [O III] luminosity tends to overestimate it—suggesting intrinsic differences in their NLR properties or covering factors.

The convolution framework of G. Risaliti et al. (2011) involved fitting the log distribution of the [O III] EWs with a double Gaussian from 2 to 30 Å and a power law from 30 to 500 Å. We performed the same analysis for our NLSy1 sample, and found that it yields a best-fit power-law slope of $\Gamma = -2.55 \pm 0.1$. This was significantly shallower than the $\Gamma = -3.55$ reported for an orientation-unbiased quasar sample. This lower Γ value likely indicates that our sample is restricted to lower inclination angles—either due to selection biases or intrinsic obscuration limits—and thus omits high-inclination sources where the torus obscures the central engine, naturally aligning with the Type 1 AGN classification. This supports the conclusion drawn from the color excess analysis in Section 4.2, which suggests a lack or negligible presence of dust obscuration. Moreover, the [O III] EW distributions of the red NLSy1s and blue NLSy1s are statistically indistinguishable (see Figure 10), implying no significant inclination difference between them. The distinct optical colors ($[u-g]$, $[u-r]$, $[g-r]$, and $[g-i]$) between these two classes must therefore arise from intrinsic differences rather than geometric effects as discussed earlier. In contrast, the intermediate Seyferts show the highest [O III] EWs among the three types, suggesting that they are being observed at relatively higher inclinations consistent to their inherent nature.

4.4. WISE Color Analysis

While several papers use WISE colors to distinguish AGN types or study dust obscuration (e.g., X.-B. Wu et al. 2012; R. Nikutta et al. 2014), the direct use of WISE color as a pure orientation indicator is debated. L. Klindt et al. (2019) specifically argue that differences in WISE color distributions between red and blue quasars are not solely due to orientation, but also to evolutionary effects. V. Fawcett et al. (2020) have used WISE color differences as a signature of radio and intrinsic property differences in red and blue quasars.

Based on the WISE color analysis shown in Figure 11, we investigate the mid-IR properties of three classes: red NLSy1s, blue NLSy1s, and intermediate Seyferts. The top panel shows the kernel density distribution of W1 − W2 colors for these subclasses. The blue NLSy1s exhibit a systematically higher W1 − W2 color peak (centered near ∼1.1 mag) than both red NLSy1s and intermediate Seyferts, with the latter two showing broader distributions. This is particularly intriguing, as the W1 − W2 ⩾ 0.8 color cut (black dashed line) from D. Stern et al. (2012) and W1 − W2 ⩾ 0.662 cur from R. J. Assef et al. (2013) are commonly used AGN selection criteria. A majority of blue NLSy1s clearly lie above both thresholds, implying robust AGN dominance in their mid-IR emission. This is consistent with the "quasar-like" properties of blue NLSy1s, indicated by their absolute B-band magnitude being greater than 23, as discussed in Section 4.1. Conversely, a significant fraction of red NLSy1s fall below these cuts, suggesting increased host galaxy contamination or intrinsically weaker AGN emission in the mid-IR, consistent with dilution effects. The bottom panel of Figure 11 (a color–color diagram, with thresholds adapted from B. Mingo et al. 2016) reinforces this interpretation: blue NLSy1s and many intermediate Seyferts occupy the AGN–quasar region (purple shaded region), whereas red NLSy1s are more scattered into the starburst and spiral galaxy regions. These patterns collectively suggest that blue NLSy1s exhibit stronger nuclear activity with more prominent hot-dust emission, while red NLSy1s may either be more host dominated or have dust obscuration from the galaxy that affects their mid-IR colors differently. Intermediate Seyferts span a wide color space, reflecting potential diversity in their dust geometry or evolutionary phase.

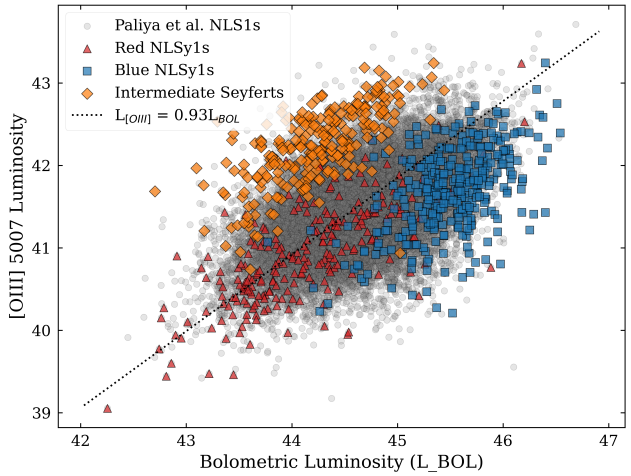

Figure 9. [O III] luminosity vs. bolometric luminosity for the entire V. S. Paliya et al. (2024) NLSy1 sample and the detected anomaly groups, i.e., red NLSy1s (red), blue NLSy1s (blue), and intermediate Seyferts (orange). The black dotted line indicates the 1:1 relation.

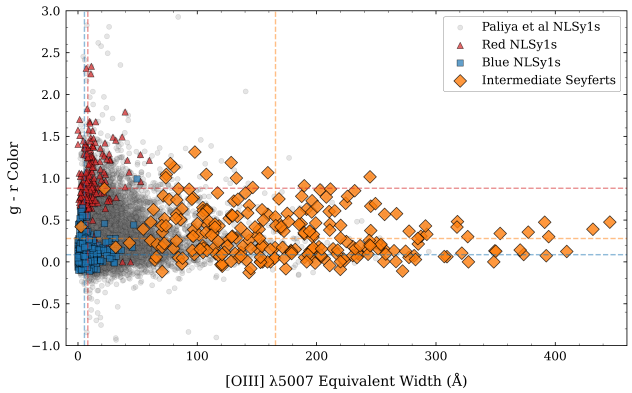

Figure 10. [O III] EW vs. $(g - r)$ optical color for the entire V. S. Paliya et al. (2024) catalog and the three anomaly groups. The horizontal and vertical dotted lines indicate the median $(g - r)$ and [O III] EW values, respectively, for each type.

4.5. Baldwin–Phillips–Terlevich Diagram

The Baldwin–Phillips–Terlevich (BPT) diagram (J. A. Baldwin et al. 1981) is a widely used diagnostic tool in extragalactic astrophysics to classify the dominant ionization mechanism in galaxies based on strong optical emission-line ratios. Typically plotted using [O III] $\lambda5007/H\beta$ versus [N II] $\lambda6584/H\alpha$, it provides a clear separation between star-forming galaxies, AGN (Seyfert and LINERs), and composite systems. Theoretical and empirical boundaries, such as the maximum starburst line from L. J. Kewley et al. (2001), the empirical demarcation of composite systems from G. Kauffmann et al. (2003b), and the LINER–Seyfert division from G. Stasińska et al. (2006) help interpret the nature of ionizing sources.

In Figure 12, we show the BPT distribution of our two anomaly classes of NLSy1s—red and blue NLSy1s and intermediate Seyferts—overlaid on the classical AGN–star formation diagnostic lines. The overall distribution lies well above the star-forming region and nearly coincident on the G. Stasińska et al. (2006) line, confirming their AGN nature. This supports their classification as Seyfert and/or NLSy1s, where high-ionization emission originates from the AGNs' NLRs.

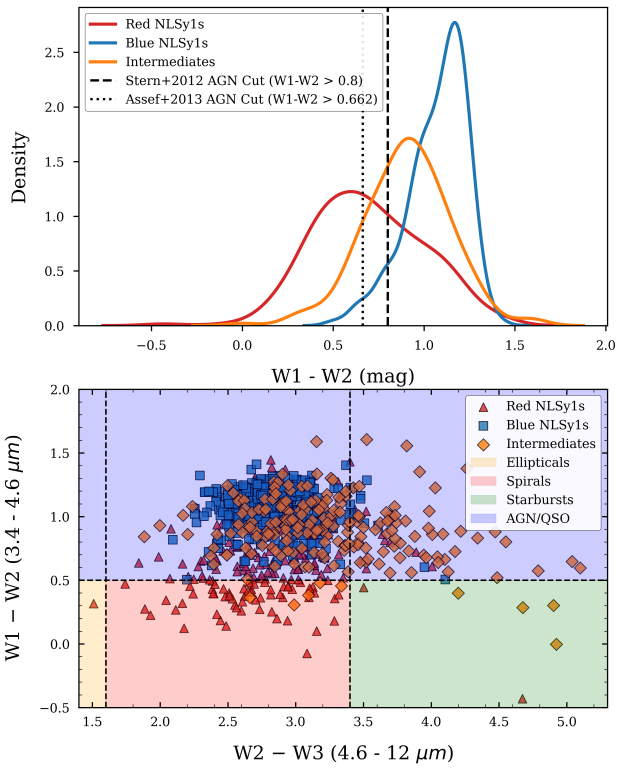

Figure 11. Top: kernel density distribution of W1 − W2 colors for red NLSy1s, blue NLSy1s, and intermediate Seyferts. Vertical lines indicate the AGN selection thresholds from D. Stern et al. (2012) and R. J. Assef et al. (2013). Bottom: WISE color–color diagram showing W1 − W2 vs. W2 − W3 with shaded regions denoting different galaxy and AGN types (adapted from B. Mingo et al. 2016).

We do not have enough samples with emission-line parameters for the blue NLSy1s (their high redshift limits coverage of the Hα line), due to which it would be impossible to make a comment about the nature of these objects using the BPT diagram. For the intermediate Seyferts and red NLSy1s, a closer inspection reveals intriguing distinctions. The red NLSy1s, while also occupying the Seyfert and composite regions, are mostly concentrated in the LINER region. They show a broader spread along the horizontal axis and appear slightly shifted toward lower [O III]/Hβ ratios, suggesting a somewhat weaker ionizing continuum or enhanced Balmer emission (in case of star-forming galaxies), consistent with their observed lower AGN luminosity and host-dominated continua (as discussed in Section 4.2). Intermediate Seyferts cluster completely in the Seyfert boundary and even extend toward the region with the highest [O III]/Hβ ratios. This shift toward elevated ratios along with the higher velocity dispersion (indicated by significantly higher [O III] EWs; Figure 10) could indicate a stronger AGN ionizing flux (see L. Ulivi et al. 2024). Their location suggests a higher incidence of either intermediate AGN activity with harder continua or orientation effects leading to diluted ionizing radiation fields. The number of galaxies in each class is given in Table 3.

Together, the BPT analysis confirms the AGN identity of all three types while unveiling systematic differences in their ionization properties and central engine strength. These differences are likely tied to a combination of intrinsic AGN power, orientation, and possibly host galaxy conditions.

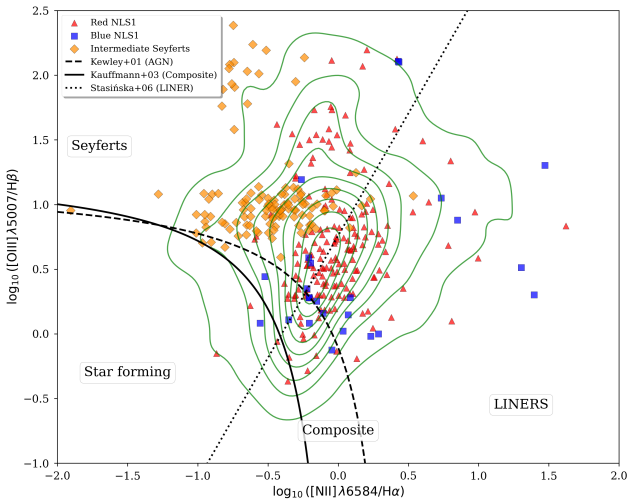

Figure 12. BPT diagram for the three detected anomaly groups along with all the NLSy1s in the V. S. Paliya et al. (2024) catalog (depicted by green contours). The respective regions of the BPT diagram are annotated.

Table 3
Number of Galaxies in Each Class of the BPT Diagram

Type	Star Forming	Composite	LINERS	Seyfert
Red	4	17	69	156
Blue	1	5	12	4
Intermediate	0	5	6	236

4.6. Eigenvector 1 Diagram

The quasar main sequence was formulated by P. Marziani et al. (2001), where they made the EV1 diagram by plotting Hβ FWHM versus Fe II strength. They divided the quasar population into two major distinctions, i.e., population A quasars with an Hβ FWHM less than 4000 km s^{-1}and population B quasars with an Hβ FWHM greater than 4000 km s^{-1}. All NLSy1s by definition fit into population A as their Hβ FWHM is defined to be less than 2000 km s^{-1}. Additionally, population A is binned on the basis of Fe II strength (R4570) into four subpopulations namely A1 (0−0.5), A2 (0.5−1), A3 (1−1.5), and A4 (1.5−2), with A4 being the strongest accretion sources and A1 being the weakest. A3 and A4 comprise the "extreme" population A (xA) population, indicating their strong Fe II emission.

Figure 13 shows an EV1 diagram for the anomalous NLSy1s detected in this work. The intermediate Seyferts marked in yellow scatter points are present in the lowermost region of Fe II strength (R4570), covering mostly the A1 region, indicating the lowest accretion rate of the three classes. On the other hand, the red NLSy1s and blue NLSy1s are majorly present in the A2 region and are spread all the way to A4 indicating high accretion rates evident by their strong Fe II strength. As discussed, blue NLSy1s are mostly "quasar-like" with strong AGN activity, while the red NLSy1s can be considered as perfect candidates for am NLSy1. This is because red NLSy1s can be thought of as galaxies where the AGN activity has just began but is not yet strong enough to wash out the host galaxy signature in the spectra.

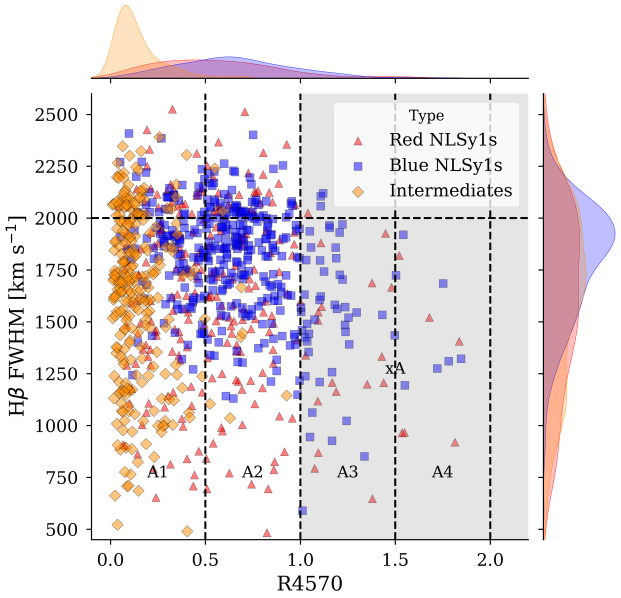

Figure 13. EV1 diagram (Hβ FWHM vs. R4570) for red and blue NLSy1s, and intermediate Seyferts, represented as a scatter plot. The histograms on the top and right axes represent the distribution of R4570 and Hβ FWHM values, respectively, for all three classes. The gray region indicating an Fe II strength greater than one (R4570 > 1) represents the xA population.

5. Conclusion

We applied the unsupervised machine learning technique SQuAD to the spectra of 22,656 NLSy1s obtained from SDSS DR17 to identify two groups of anomalous NLSy1s and a population of intermediate Seyferts (misclassified as NLSy1s in the parent sample of V. S. Paliya et al. 2024). Table 4 provides a sample list, including the SDSS name, R.A., decl., redshift, *r*-band magnitude, and their final classification. We then characterized the spectroscopic, photometric, and physical properties of these three populations—red and blue NLSy1s and intermediate Seyferts—by combining multi-wavelength and multiapproach diagnoses. Below, we summarize the key findings for each group and highlight the broader implications of this work.

5.1. Red Narrow-line Seyfert 1 Galaxies

We conclude them to be host-dominated, low-luminosity young AGNs with a smaller black hole mass. They exhibit a steep red continuum with weak Fe II emission and moderate [O III] emission, dominated by host galaxy starlight (evidenced by strong D4000 breaks and Ca II absorption). The modest average $E(B-V) \sim 0.54$ values rule out heavy nuclear obscuration. Reddening is instead attributed to host galaxy dilution of a faint AGN continuum. In the WISE color distributions, the red NLSy1s scatter into starburst and spiral regions in mid-IR color space, confirming host dominance over AGN emission. Hence, the physically red NLSy1s likely represent low-accretion-rate NLSy1s where the AGN activity has just began and hence is intrinsically weak, allowing the host galaxy to dominate the optical spectrum. Their prevalence at low redshifts suggests selection biases against detecting them at higher z due to sensitivity limits. It was noted that nearly all the red NLSy1s for which the SDSS image exhibited an extended identifiable source featured a bar. This goes in

Table 4
Sample Table Listing the Anomalous Narrow-line Seyfert 1 Galaxies Identified in This Work

Name	R.A. (deg)	Decl. (deg)	Redshift (z)	r-band magnitude (mag)	Anomaly type
SDSS J092438.88+560746.8	141.1620	56.1296	0.025	13.87	red
SDSS J103922.32-003404.0	159.8430	−0.5677	0.771	18.71	blue
SDSS J032723.13+004422.4	51.8463	0.7395	0.138	17.80	intermediate

Note. A complete version of the table is available in machine-readable format.
(This table is available in its entirety in machine-readable form in the online article.)

accordance with the findings of D. M. Crenshaw et al. (2003) who report that the NLSy1 fueling is bar driven.

5.2. Blue Narrow-line Seyfert 1 Galaxies

These are highly luminous NLSy1s with larger black hole masses and high-Eddington ratios. The blue NLSy1s exhibit a blue continuum with strong Fe II (see bottom right panel of Figure 4) and weak [O III] emission, consistent with high-Eddington rate accretion. In the WISE color distribution plot, they are concentrated in the AGN region ($W1-W2 \geqslant 0.8$), indicating hot-dust emission and minimal host contamination. Their population exhibits a maximum at $z \sim 0.8$, aligning with the cosmic peak of quasar activity, and are rare at low z due to the declining AGN luminosity density function. Physically, they can be thought as classic high-$L/L_{\rm Edd}$ NLSy1s, possibly undergoing rapid black hole growth. Their bluer colors and Fe II strength suggest a radiatively efficient accretion flow with low dust extinction. As compared to a typical NLSy1, they exhibit a much steeper blue continuum and higher Fe II emission with mean $R4570 = 0.66$ as compared to a mean R4570 of 0.46 for the entire V. S. Paliya et al. (2024) catalog. Additionally, blue NLSy1s have higher bolometric luminosities with a mean $\log L_{\rm bol} = 45.51$ as compared to an $\log L_{\rm bol}$ of 44.74 to 44.89 for a typical NLSy1 (V. K. Jha et al. 2022; V. S. Paliya et al. 2024). The blue NLSy1s also exhibit a larger black hole mass with $\log M_{\rm BH} = 7.32 \pm 0.02$ as compared to 6.98 ± 0.04 for a typical NLSy1 (S. Rakshit et al. 2017; S. Rakshit & C. Stalin 2017; V. S. Paliya et al. 2024). The blue NLSy1s have a mean Eddington ratio of 1.13 as compared to a mean of 0.44 for the entire V. S. Paliya et al. (2024) catalog. The distributions of their absolute B-band magnitudes (Figure 5) and WISE colors (Figure 11) suggest that these sources are more like narrow-line quasars than NLSy1s.

5.3. Intermediate Seyfert Galaxies

This group is distinct from the NLSy1s and represents a different subset of Seyfert galaxies called intermediate Seyferts. These are primarily identified by the presence of composite emission-line profiles (D. E. Osterbrock & A. Koski 1976; R. D. Cohen 1983), i.e., a narrow component (characteristic of a Seyfert 2 galaxy) superimposed on a broad component (typical of a Seyfert 1 galaxy). This group is distinguished by their spectra, which are almost entirely devoid of continuum emission and dominated by strong atomic emission lines (see Figure 14). Their spectral appearance resembles that of objects with an exposed NLR and highly suppressed or obscured accretion disk emission. As shown in Figure 10, these objects exhibit the largest [O III] EWs among the anomaly types—an indicator often associated with edge-on viewing angles or optically thick circumnuclear material, as

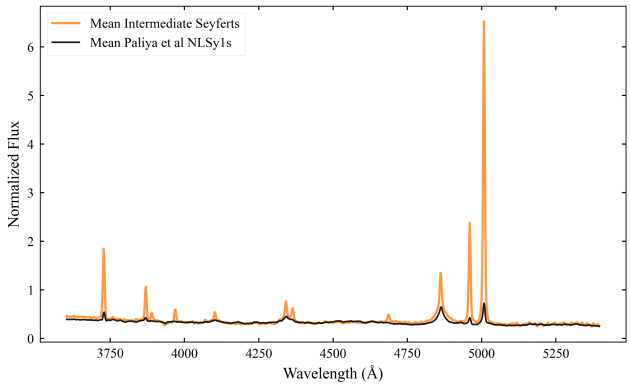

Figure 14. Mean composite spectrum for the intermediate Seyferts (orange) as compared to the mean composite of all NLSy1s in the V. S. Paliya et al. (2024) catalog (black). The V. S. Paliya et al. (2024) composite spectrum is scaled arbitrarily to match the continuum of the intermediate Seyferts for better comparison.

suggested by G. Risaliti et al. (2011). This interpretation aligns with the lack of observable continuum emission, supporting the idea that these sources may be obscured by orientation or structural effects. Moreover, the intermediate Seyferts display elevated values of the $W1-W2$ IR color, likely due to thermal dust reemission, consistent with their inferred intermediate inclination geometry. However, their $E(B-V)$ distribution suggests little to no effect of dust reddening in their optical spectra, contradicting the edge-on geometric interpretation. Their positions on the WISE color–color diagram (see bottom panel of Figure 11) span both the AGN and starburst–composite regions, implying a diversity in dust geometries or evolutionary phases. These combined features suggest that the intermediate Seyferts may represent a transitional or structurally unique phase of AGN activity. Overall, the physical properties of intermediate Seyferts—as suggested by our analysis—are, in general, contrasting to those of a typical NLSy1. The intermediate Seyferts represent a contaminant class in the V. S. Paliya et al. (2024) catalog, wrongfully labeled as NLSy1s and instead are a class of intermediate Seyferts. This interpretation is consistent with the detailed reanalysis of SDSS J21185.96-073227.5 by E. Järvelä et al. (2020), who revised its classification from a γ-ray-emitting NLSy1 to an intermediate Seyfert.

The properties of intermediate Seyferts such as hard-ionization continua, narrow and strong emission lines, and the virtual lack of a continuum is remarkably similar to the "C IV peakers" presented by A. Tiwari & M. Vivek (2025). These objects exhibit rare emission lines such as Ne V λ3345, [Ne V] λ3427, and [Ne VI] λ3869 (R. D. Cohen 1983), which are absent in a typical quasar or NLSy1, providing us with an opportunity to explore new areas of quasar astrophysics. These

high-ionization lines have been used to determine physical parameters such as the inclination of the sources (J. Scholtz et al. 2025), B. Trakhtenbrot et al. (2025) use the [Ne V] λ3426 line to calculate appropriate scaling to calculate AGN luminosity and SMBH mass in high-redshift AGNs. They also use the ratios of these high-ionization lines to run photoionization modeling, which was then used to plot BPT diagrams curated for high-redshift sources. M. J. Temple et al. (2023) have provided a framework to study the coevolution of the C IV and He II emission lines relative to gas, black hole, and accretion properties. We plan to use this framework in an upcoming work for the analysis of these sources.

By integrating spectral decomposition, dust extinction analysis, mid-IR diagnostics, and emission-line kinematics, this work provides a unified framework for understanding NLSy1 and Seyfert diversity. The anomalies identified here—whether host dominated, accretion driven, or geometrically obscured—highlight the complex interplay between AGN physics and host galaxy environments. Future surveys (e.g., the Large Synoptic Survey Telescope and SDSS-V) will expand this taxonomy, probing fainter and rarer AGN subpopulations to refine our picture of black hole growth and feedback. Additionally, given the distinct and extreme properties of intermediate Seyferts, we are conducting a deeper spectroscopic and multiwavelength studies of these intriguing objects to further constrain their physical nature and evolutionary context.

Acknowledgments

We thank the anonymous referee for the feedback, which has significantly helped to improve the paper. A.T. acknowledges and thanks the Indian Institute of Astrophysics (IIA) for the financial and infrastructural support provided, and to Aarya H. for helping with the PyQSOFit fittings of the red NLSy1s. M.V. acknowledges support from Department of Science and Technology, India—Science and Engineering Research Board (DST-SERB) in the form of a core research grant (CRG/2022/007884). Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is www.sdss4.org.

SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics | Harvard & Smithsonian (CfA), the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

ORCID iDs

Arihant Tiwari https://orcid.org/0009-0007-5545-0518
Rachana https://orcid.org/0000-0003-4592-447X
M. Vivek https://orcid.org/0000-0001-5937-331X
Suvendu Rakshit https://orcid.org/0000-0002-8377-9667

References

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2009, ApJ, 707, L142
Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35
Andonie, C., Alexander, D. M., Greenwell, C., et al. 2025, MNRAS, 539, 2202
Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, ApJ, 772, 26
Baldi, R. D., Capetti, A., Robinson, A., Laor, A., & Behar, E. 2016, MNRAS, 458, L69
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Berton, M., Björklund, I., Lähteenmäki, A., et al. 2020, CoSka, 50, 270
Blustin, A. J., Page, M. J., Fuerst, S. V., Branduardi-Raymont, G., & Ashton, C. E. 2005, A&A, 431, 111
Boller, T., Brandt, W., & Fink, H. 1996, A&A, 305, 53
Boroson, T. A. 2002, ApJ, 565, 78
Boroson, T. A., & Green, R. F. 1992, ApJS, 80, 109
Boyle, B. J., Shanks, T., Croom, S., et al. 2000, MNRAS, 317, 1014
Brinchmann, J., Charlot, S., White, S. D., et al. 2004, MNRAS, 351, 1151
Buat, V., Giovannoli, E., Heinis, S., et al. 2011, A&A, 533, A93
Caccianiga, A., Severgnini, P., Braito, V., et al. 2004, A&A, 416, 901
Calderone, G., Ghisellini, G., Colpi, M., & Dotti, M. 2013, MNRAS, 431, 210
Calvi, V., Stiavelli, M., Bradley, L., Pizzella, A., & Kim, S. 2014, ApJ, 796, 102
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Chaudhary, S. C., & Prince, R. 2025, ApJ, 996, 118
Cohen, R. D. 1983, ApJ, 273, 489
Constantin, A., & Shields, J. C. 2003, PASP, 115, 592
Crenshaw, D. M., Kraemer, S. B., & Gabel, J. R. 2003, AJ, 126, 1690
Dalla Bontà, E., Peterson, B. M., Bentz, M. C., et al. 2020, ApJ, 903, 112
Fawcett, V., Alexander, D., Rosario, D., et al. 2020, MNRAS, 494, 4802
Fernández-Ontiveros, J., López-López, X., & Prieto, A. 2023, A&A, 670, A22
Fitzpatrick, E. L. 1999, PASP, 111, 63
Foschini, L. 2020, Univ, 6, 136
Foschini, L., Colpi, M., Gallo, L. et al. (ed.) 2011, in Narrow-Line Seyfert 1 Galaxies and their Place in the Universe
Garn, T., & Best, P. N. 2010, MNRAS, 409, 421
Goodrich, R. W. 1989, ApJ, 342, 224
Gordon, K. D. 2024, JOSS, 9, 7023
Gordon, K. D., Clayton, G. C., Misselt, K., Landolt, A. U., & Wolff, M. J. 2003, ApJ, 594, 279
Grossan, B., Remillard, R. A., Bradt, H. V., et al. 1996, ApJ, 457, 199
Grupe, D. 2004, AJ, 127, 1799
Grupe, D., Komossa, S., Leighly, K. M., & Page, K. L. 2010, ApJS, 187, 64
Guo, H., Shen, Y., & Wang, S. 2018, PyQSOFit: Python code to fit the spectrum of quasars, Astrophysics Source Code Library, ascl:1809.008
Ho, L. C. 2008, ARA&A, 46, 475
Ho, L. C., Filippenko, A. V., & Sargent, W. L. 1997, ApJS, 112, 315
Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, ApJ, 654, 731
Hotelling, H. 1933, Journal of Educational Psychology, 24, 417
Järvelä, E., Berton, M., Ciroi, S., et al. 2020, A&A, 636, L12
Jha, V. K., Chand, H., Ojha, V., Omar, A., & Rastogi, S. 2022, MNRAS, 510, 4379
Kaspi, S., Maoz, D., Netzer, H., et al. 2005, ApJ, 629, 61
Kaspi, S., Smith, P. S., Netzer, H., et al. 2000, ApJ, 533, 631
Kauffmann, G., Heckman, T. M., White, S. D., et al. 2003a, MNRAS, 341, 33
Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003b, MNRAS, 346, 1055
Kewley, L. J., Dopita, M. A., Sutherland, R., Heisler, C., & Trevena, J. 2001, ApJ, 556, 121
Klindt, L., Alexander, D. M., Rosario, D. J., Lusso, E., & Fotopoulou, S. 2019, MNRAS, 488, 3109
Leighly, K. M. 1999a, ApJS, 125, 297
Leighly, K. M. 1999b, ApJS, 125, 317

Leighly, K. M. 2004, ApJ, 611, 125
Leighly, K. M., & Moore, J. R. 2004, ApJ, 611, 107
Li, Z., Zhou, H., Hao, L., et al. 2015, ApJ, 812, 99
Lipari, S., Tsvetanov, Z., & Macchetto, F. 1993, ApJ, 405, 186
Lupton, R. H., Ivezic, Z., Gunn, J. E., et al. 2002, SPIE, 4836, 350
MacQueen, J. 1967, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics (Univ. California Press), 281
Maiolino, R., Marconi, A., Salvati, M., et al. 2001, A&A, 365, 28
Marziani, P., Sulentic, J., Zwitter, T., Dultzin-Hacyan, D., & Calvani, M. 2001, ApJ, 558, 553
Mingo, B., Watson, M., Rosen, S., et al. 2016, MNRAS, 462, 2631
Myers, A. D., Palanque-Delabrouille, N., Prakash, A., et al. 2015, ApJS, 221, 27
Nadarajah, S., & Kotz, S. 2004, Mathematical Problems in Engineering, 2004, 323
Nagar, N. M., Oliva, E., Marconi, A., & Maiolino, R. 2002, A&A, 391, L21
Netzer, H. 1990, Active Galactic Nuclei (Springer), 57
Nikutta, R., Hunt-Walker, N., Nenkova, M., Ivezić, Ž., & Elitzur, M. 2014, MNRAS, 442, 3361
Orienti, M., D'Ammando, F., Larsson, J., et al. 2015, MNRAS, 453, 4037
Osterbrock, D. E., & Koski, A. 1976, MNRAS, 176, 61P
Osterbrock, D. E., & Pogge, R. W. 1985, ApJ, 297, 166
Paliya, V. S., Stalin, C., Domínguez, A., & Saikia, D. 2024, MNRAS, 527, 7055
Pan, X., Zhou, H., Ge, J., et al. 2017, ApJ, 835, 218
Peterson, B. M. 2011, arXiv:1109.4181
Peterson, B. M. 2015, The Physics of Accretion onto Black Holes (Springer), 253
Poggianti, B. M., & Barbaro, G. 1997, A&A, 325, 1025
Pounds, K. A., Done, C., & Osborne, J. P. 1995, MNRAS, 277, L5
Rakshit, S., & Stalin, C. 2017, ApJ, 842, 96
Rakshit, S., Stalin, C., Chand, H., & Zhang, X.-G. 2017, ApJS, 229, 39
Rakshit, S., Johnson, A., Stalin, C. S., Gandhi, P., & Hoenig, S. 2019, MNRAS, 483, 2362
Reeves, J. N., Braito, V., Porquet, D., et al. 2023, ApJ, 952, 52
Richards, G. T., Strauss, M. A., Fan, X., et al. 2006, AJ, 131, 2766
Risaliti, G., Salvati, M., & Marconi, A. 2011, MNRAS, 411, 2223
Savitzky, A., & Golay, M. J. E. 1964, AnaCh, 36, 1627
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schmidt, E. O., Ferreiro, D., Neme, L. V., & Oio, G. A. 2016, A&A, 596, A95
Schnorr-Müller, A., Storchi-Bergmann, T., Robinson, A., Lena, D., & Nagar, N. M. 2016, MNRAS, 457, 972
Scholtz, J., Maiolino, R., D'Eugenio, F., et al. 2025, A&A, 697, A175
Shen, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2020, MNRAS, 495, 3252
Shen, Y., & Ho, L. C. 2014, Natur, 513, 210
Shlosman, I., Begelman, M. C., & Frank, J. 1990, Natur, 345, 679
Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, AJ, 146, 32
Stasińska, G., Fernandes, R. C., Mateus, A., Sodré, L., & Asari, N. V. 2006, MNRAS, 371, 972
Steinhardt, C. L., & Silverman, J. D. 2013, PASJ, 65, 82
Stern, D., Assef, R. J., Benford, D. J., et al. 2012, ApJ, 753, 30
Temple, M. J., Matthews, J. H., Hewett, P. C., et al. 2023, MNRAS, 523, 646
Tiwari, A., & Vivek, M. 2025, A&A, 699, A132
Tombesi, F., Cappi, M., Reeves, J. N., et al. 2013, MNRAS, 430, 1102
Trakhtenbrot, B., Ricci, C., Treister, E., et al. 2025, ApJ, 992, L27
Tukey, J. W. 1977, Exploratory Data Analysis, 2 (Springer)
Ulivi, L., Venturi, G., Cresci, G., et al. 2024, A&A, 685, A122
Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, AJ, 122, 549
Varglund, I., Järvelä, E., Ciroi, S., et al. 2023, A&A, 679, A32
Véron-Cetty, M. P., Véron, P., & Gonçalves, A. C. 2001, A&A, 372, 730
Vivek, M., Srianand, R., Petitjean, P., et al. 2012, MNRAS, 423, 2879
Whittle, M. 1992, ApJS, 79, 49
Wu, X.-B., Hao, G., Jia, Z., Zhang, Y., & Peng, N. 2012, AJ, 144, 49
Yang, H., Yuan, W., Yao, S., et al. 2018, MNRAS, 477, 5127
Yao, S., & Komossa, S. 2023, MNRAS, 523, 441
Yu, Q., & Tremaine, S. 2002, MNRAS, 335, 965
Zhang, S., Zhou, H., Shi, X., et al. 2017, ApJ, 845, 126
Zhang, W., Shu, X., Sheng, Z., et al. 2022, A&A, 660, A119