Erasmus Mundus Master
in Astrophysics and Space Science

Master Thesis

Comparison of the spectral properties of a sample of near-by type 1 AGNs measured from Gaia XP and SDSS spectra

Supervisors

Prof. Dragana Ilić
University of Belgrade, Faculty of Mathematics

Prof. Francesco Tombesi
Physics Dept., Tor Vergata University of Rome

Prof. Andjelka B. Kovačević
University of Belgrade, Faculty of Mathematics

Prof. Luka Č. Popović
Astronomical Observatory Belgrade

Author

Nandita Das

Academic Year 2024/2025

UNIVERZITET U BEOGRADU
MATEMATIČKI FAKULTET

Nandita Das

COMPARISON OF THE SPECTRAL PROPERTIES OF A SAMPLE OF NEAR-BY TYPE 1 AGNS MEASURED FROM GAIA XP AND SDSS SPECTRA

master rad

Beograd, 2025.

Mentor:

dr Dragana Ilić
Univerzitet u Beogradu, Matematički fakultet

Članovi komisije:

dr Andjelka B. Kovačević
Univerzitet u Beogradu, Matematički fakultet

dr Francesco Tombesi
Università degli Studi di Roma "Tor Vergata", Dipartimento di Fisica

dr Luka Č. Popović
Astronomska Opservatorija, Beograd

Datum odbrane: _____

This Master thesis is submitted in partial fulfillment of the requirements for the degree "MASTER ASTROFIZICAR" as part of a multiple degree awarded in the framework of the Erasmus Mundus Joint Master in Astrophysics and Space Science -- MASS jointly delivered by a Consortium of four Universities: Tor Vergata University of Rome, University of Belgrade, University of Bremen, and Université Cote d'Azur, regulated by the MASS Consortium Agreement and Co-funded by the EU under the call ERASMUS-EDU-2021-PEX-EMJM-MOB.

"The path from dreams to success does exist. May you have the vision to find it, the courage to get on to it, and the perseverance to follow it."

— **Kalpana Chawla**

Contents

Contents v

List of Figures vii

List of Tables x

Acknowledgements xi

Abstract xiii

1 Introduction 1

1.1 Active Galactic Nuclei . . . 1

1.1.1 What is an Active Galactic Nucleus (AGN) . . . 1

1.1.2 The Unified Model . . . 2

1.1.3 Broad-Line Region (BLR) . . . 3

1.1.4 Type 1 AGNs . . . 5

1.2 Optical Emission Lines of AGNs . . . 6

1.3 Motivation and Thesis Objectives . . . 7

2 Datasets and Sample Selections 9

2.1 Datasets . . . 9

2.1.1 Sloan Digital Sky Survey (SDSS) . . . 9

2.1.2 Gaia Mission Overview and Scientific Context . . . 10

2.2 Sample Selection . . . 11

2.2.1 Selection Criteria for Type 1 AGNs in SDSS . . . 12

2.2.2 Source Cross-Match Between SDSS and Gaia . . . 13

2.3 Properties of Gaia XP Data . . . 14

2.3.1 Gaia XP Data Retrieval . . . 14

2.3.2 Calibration of Continuous XP Spectra . . . 15

2.3.3 Preparation of Calibrated XP Spectra for Fitting . . . 15

2.3.4 Testing XP Spectra Corrections . . . 17

2.3.5 Comparison of Gaia and SDSS Spectra . . . 18

2.4 Light Curves . . . 19

2.4.1 SDSS Stripe 82 Light Curves . . . 19

2.4.2 Gaia Epoch Photometry Overview . . . 20

3 Spectral Analysis 23

3.1 Spectral Fitting Approach with FANTASY . . . 23

3.1.1 Framework . . . 23

3.1.2 SDSS Line-fitting Model . . . 24

3.1.3 Gaia Line-fitting Model . . . 25

3.2 Example Fits for Representative Objects . . . 26

3.2.1 SDSS Fits 26
3.2.2 Gaia Fits 30
3.2.3 Comparison of SDSS and Gaia Fits 31
3.3 Error Treatment and Uncertainty Estimates 33

4 Physical Properties Measurements of the Sample 35
4.1 Emission Line and Continuum Flux Measurements 35
4.1.1 Fluxes and FWHM 36
4.1.2 Balmer Decrement 36
4.2 Luminosities Measurement 38
4.3 Correlations of Spectral Parameters 40
4.3.1 Line Luminosity Correlations (H , H) 40
4.3.2 Line Widths (FWHM) 42
4.3.3 Continuum Luminosities at 5100 Å and 6200 Å 43
4.3.4 Line Continuum Correlations 45
4.4 Black Hole Mass Estimation 46
4.5 Eddington Ratio 48

5 Time-Domain Analysis 50
5.1 Light Curve Sample 50
5.2 Preprocessing 51
5.2.1 SDSS Stripe 82 51
5.2.2 Gaia Epoch Photometry 51
5.3 Computation of Variability Metrics 52
5.3.1 Common Sample: SDSS and Gaia 53
5.3.2 Variability Analysis of only Gaia Light Curves 57

6 Discussion 62

7 Conclusion 64

References 66

Appendix A: read_gaia Loader Function 72

Appendix B: Full Sample of Spectra 75

List of Figures

1 Uni ed Model of AGNs 3
2 BPT diagram for AGN classi cation 6
3 Layout of the Gaia focal plane 11
4 Comparison of GaiaXPy calibrated XP spectrum from this work (top) and the independent ARI XP spectrum (bottom) for the same AGN (Gaia DR3 source ID: 3990830907900291584). The excellent match in continuum and Balmer line features supports the accuracy of the GaiaXPy calibration. 16
5 Gaia XP spectrum of a representative source (source ID: 6342859972491701 76) calibrated using GaiaXPy with two di erent wavelength samplings: 2 nm (solid blue line) and 0.2 nm (dashed red line with markers). The two samplings produce indistinguishable spectral shapes, but the ner sampling provides more points across emission lines, improving integrated ux accuracy. 17
6 Comparison between the original Gaia XP spectrum (green) and the corrected spectrum using the (Huang et al., 2024) code (orange) for a representative AGN. Only minor di erences are observed in the blue continuum; broad emission lines remain una ected. 18
7 Comparison of Gaia DR3 XP spectrum (red) and SDSS spectrum (green) for one representative source. The Gaia spectrum is shifted vertically for clarity while ux calibration is preserved. 19
8 Example of an SDSS Stripe 82 light curve from our sample (source ID: LC_3861327). Multi-band observations over nearly a decade are shown in the ve SDSS lters, illustrating irregular but seasonally clustered cadence and wavelength dependent variability amplitude. 21
9 Example of an Gaia light curve from our sample (source ID: 3267488664984941184), same source as for SDSS light curve Figure 8. 22
10 Example of an SDSS spectrum with a well-constrained t (Case 1: Good Fit). The observed spectrum (gray) is reproduced by the global model (red), with contributions from the continuum (dark blue dashed), Fe II template (yellow-green), broad Balmer components (blue), and narrow lines including [O III] and [N II] (magenta). Residuals are low across the wavelength range, highlighting the robustness of the t. 26
11 Example of an SDSS spectrum with strong host galaxy contamination. Top: PCA-based decomposition of the observed spectrum (blue) into AGN (green) and host galaxy (orange) components, with the total model shown in red. Bottom: Final spectral t after host subtraction, including continuum, Fe II emission, and broad/narrow line components. 27
12 Example of an SDSS spectrum with strong Fe II emission features. The broad Fe II multiplets produce extensive blends across the optical range, particularly blueward of H , complicating the identi cation of the Balmer line wings. Incorporating empirical Fe II templates into the t allows reliable recovery of both the continuum and the broad-line pro les. 28

13 Example of an SDSS spectrum with an exceptionally broad H emission line. The model includes two broad Gaussian components (solid and dashed blue) to capture the extended wings in addition to the narrow-line contribution. The t demonstrates how extreme BLR velocities can still be constrained within the FANTASY framework. 29
14 Example of an SDSS spectrum where the broad Balmer component is weak and narrow, making it di cult to separate from the strong narrow-line emission. The t includes continuum, Fe II, narrow forbidden lines, and Balmer components, with the residuals indicating that the broad component is only marginally detected. 29
15 Example of an SDSS spectrum where the narrow forbidden lines ([O III] and [N II]) are exceptionally strong, partially blending with the broad Balmer emission. The t includes continuum, Fe II, narrow forbidden lines, and Balmer components. While the decomposition reproduces the observed spectrum well, the broad-line uxes remain uncertain due to the dominance of the narrow features. 30
16 Example of a well-constrained Gaia XP spectrum. The continuum is well de ned across the tting range, and the principal Balmer lines (H , H , H) are clearly recovered with minimal blending. The residuals remain small and structureless, indicating that the simpli ed broad-line only model provides a robust representation of the data despite the limited spectral resolution. . . . 31
17 Example of a Gaia XP spectrum where the continuum could not be consistently tted across both H and H . Top: Cropped t around H and H . Bottom: Cropped t around H . Splitting the spectrum allows better continuum placement in each region and improves the reliability of line ux measurements. 32
18 Example of a challenging Gaia XP spectrum with strong Fe II contamination. The smoothed Fe II blends overlap with the Balmer wings, preventing robust decomposition of the broad-line pro les. The resulting uxes are agged as unreliable in the analysis. 32
19 Broad-line H vs H luminosity correlations within SDSS (left) and within Gaia XP (right). 40
20 Comparison of broad-line luminosities between SDSS and Gaia XP: H (left) and H (right). 41
21 Comparison of broad-line Balmer decrement $F_H = F_H$ between SDSS and Gaia XP. Grey line marks the one-to-one relation, while the blue line shows the linear t. 42
22 Comparison of FWHM measurements between SDSS and Gaia XP: H (left) and H (right). The dashed line indicates the one-to-one relation. 43
23 Internal comparison of FWHM(H) and FWHM(H) within SDSS (left) and within Gaia XP (right). The dashed line indicates the one-to-one relation. . . 43

24 Continuum luminosity comparison at 5100 Å and 6200 Å within each survey. Left: SDSS; right: Gaia XP. The grey line shows the one-to-one relation; the blue line is the best- t regression. 44
25 Cross-survey comparison of continuum luminosities. Left: SDSS vs Gaia XP at 5100 Å; right: SDSS vs Gaia XP at 6200 Å. The grey line shows the one-to-one relation; the blue line is the best- t regression. 44
26 Correlation between H line luminosity and 5100 Å continuum luminosity. Left: SDSS; right: Gaia XP. The grey line shows the one-to-one relation, and the blue line the linear t. 45
27 Correlation between H line luminosity and 6200 Å continuum luminosity. Left: SDSS; right: Gaia XP. The grey line shows the one-to-one relation, and the blue line the linear t. 45
28 Comparison of black hole mass estimates derived from the broad H line using SDSS (x-axis) and Gaia XP (y-axis) spectra. The dashed line indicates the one-to-one relation, while the shaded region marks 0:4 dex, corresponding to the typical scatter of single-epoch virial mass estimators. Points are color-coded by redshift, with error bars showing the propagated calibration uncertainties. The median o set between the two estimates is ~ ' +0 :76 dex, with a Pearson correlation coe cient of r 0:71. 47
29 Comparison of Eddington ratios $L=L_{Edd}$ derived from SDSS (x-axis) and Gaia XP (y-axis) spectra. The dashed line indicates the one-to-one relation, while the o set represents the median $\log(L=L_{Edd})$ ' 0:56 dex. Points are color-coded by redshift, with error bars omitted for clarity. The scatter could possibly be due to challenges in measuring the continuum ux in low-luminosity AGNs. This illustrates the impact of spectral resolution and calibration differences on the recovered accretion ratios. 49
30 Composite light curves for the rst three AGN observed in both SDSS Stripe 82 (multi-band u; g; r; i; z ; colored points) and Gaia DR3 (epoch photometry in G, BP , RP ; green/blue/red points). The plots illustrate the complementary strengths of the two surveys: long-term, multi-band ground-based coverage from SDSS and high-precision, space-based photometry from Gaia. 54
31 The same is in Figure 30 but for the remaining four AGNs. 55
32 Histograms with KDE overlays showing the distribution of F_{var} values for the BP , G, and RP bands (top to bottom) across the 48 AGN with Gaia-only light curves. 58
33 Cumulative distribution of F_{var} for the BP , G, and RP bands. The G band shows a steep rise, indicating low variability for most AGN. BP and RP exhibit longer tails, revealing greater diversity in photometric variability. 59
34 Cross-survey Pearson correlations (in $\log_{10}$ space) between SDSS spectral features (y-axis) and Gaia XP-derived features (x-axis). Continuum and luminosity quantities show high correlations (r > 0:9), while emission-line ux ratios and FWHM exhibit much weaker agreement. 63

List of Tables

1 Observed line and continuum uxes (SDSS and Gaia) for the total sample of 64 Type 1 AGNs. 37
2 Observed line and continuum luminosties (SDSS and Gaia) for the total sources of 60 Type 1 AGNs. 39
3 Variability features of seven objects with light curves both in SDSS and Gaia surveys. 56
4 Summary statistics of F_{var} for the 7 AGNs with SDSS and Gaia. 57
5 Summary statistics of F_{var} for the 48 AGNs with Gaia-only light curves. . . . 57
6 Variability features of 48 objects with light curves in Gaia surveys. 60
7 Rest of 48 objects with light curves in Gaia surveys. 61

Acknowledgements

With this master thesis, a de ning journey in the pursuit of knowledge and inner growth is coming to an end point with a new beginning towards the mysteries of the universe. The last three years with MASS full of Memories, Achievements, Struggles, Successes not only reshaped my academic path but also my perspective in life in such a way that I will carry forever within me. It was a time of questions and awakenings, of being broken and rebuilt, of losing myself and nding anew. Reaching this point would not have been possible without the enormous support and encouragement of the people who have been part of this journey.

At the heart of this acknowledgment, I would like to express my deepest gratitude to my mentor and supervisor, Prof. Dragana Ili¢, for her continuous support and guidance. Her mentorship has been invaluable for the direction of this thesis and my growth as a researcher. With remarkable patience explaining concepts as many times as needed, her availability, whether during long days or late hours, was just an email away. I would like to thank her for enlightening me on methods of research, values of perseverance, discipline, and integrity. I am further grateful for her psychological support, encouragement, and the life lessons she imparted, which I will carry with me throughout my life. I am equally grateful to my supervisors, Prof. Andjelka B. Kova£evi¢, Prof. Francesco Tombesi and Prof. Luka ƒ. Popovi¢ for their valuable guidance and feedback throughout this project.

I am deeply thankful to the MASS project o ce, especially Dr. Adalia Caroli, for their unwavering support and assistance with all logistics. I am also grateful to the entire MASS consortium universities for establishing this global platform, which provides a remarkable foundation for academic and scienti c collaboration. Special thanks to the MASS Teaching Committee and Selection Committee, who believed in me and gave me this wonderful opportunity to pursue my dream.

I am sincerely grateful to the students of MASS Edition 1 for their support and encouragement at the very beginning of this journey. I am equally thankful to the students of MASS Edition 2, who welcomed me with warmth and kindness when I returned, with full of doubts about myself. I would like to thank Marina, whom I regard as a friend, philosopher, and guide. Her strength, positivity, and calm resilience have been a constant source of courage and inspiration, and I am truly grateful to have shared this unforgettable experience with her. I also wish to thank Juthy and Shoaib for their care and support in challenging days. Their presence has made this journey far easier than I could have imagined.

I would like to thank my family for their unconditional love, support, and encouragement. I am especially grateful to my husband, who has given me the wings to y and the reassurance that whenever I fall, there will always be someone to catch me.

Last but not the least, I like to thank myself for an indomitable spirit - for holding on and ghting back, no matter what life placed before me. From a place where there was neither chance nor reality to dream of this path, I stand here because of that ambitious child within me, who dared to dream about space.

Nandita Das acknowledges support through an Erasmus Mundus Joint Master (EMJM) scholarship or co-funded by the European Union in the framework of the Erasmus+, Erasmus Mundus Joint Master in Astrophysics and Space Science MASS. Views and opinions expressed are however those of the author(s) only and do not necessarily re ect those of the European Union or granting authority European Education and Culture Executive Agency (EACEA). Neither the European Union nor the granting authority can be held responsible for them.

Funding for the SDSS has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

Abstract

Active Galactic Nuclei (AGNs) are among the most luminous astrophysical sources, powered by accretion of matter onto supermassive black holes (SMBHs) at the centers of galaxies. In type 1 AGNs, observed broad emission lines originating from the broad-line region (BLR) provide essential diagnostics for SMBH mass estimation and accretion studies. While ground-based spectroscopic surveys such as the Sloan Digital Sky Survey (SDSS) have been extensively used for such measurements, the recent availability of space-based low-resolution spectra from the Gaia DR3 XP instrument o ers a new opportunity for AGN research. In this work, we perform a comparative analysis of optical spectra for a sample of nearby type 1 AGNs observed by both SDSS and Gaia XP. The spectra are decomposed using the open-source Python package FANTASY, enabling multi-component tting of the AGN continuum, broad and narrow emission lines. We focus on measuring the continuum and uxes, line widths, and ratios of the broad Ha and Hb emission lines in both datasets, evaluating the consistency and limitations of Gaia XP for BLR studies. Additionally, we compare variability patterns using SDSS multi-band light curves and Gaia G/BP/RP photometry. Our results demonstrate the potential of Gaia XP spectra for broad emission-line diagnostics and measurements of the SMBH parameters, such as the mass and accretion rate, while highlighting systematic di erences relative to SDSS that are mostly due to spectral resolution in the XP data. This work serves as a rst step toward leveraging Gaia's massive spectroscopic dataset for future AGN studies and for complementing ground-based observations in large-scale surveys.

Chapter 1

1 Introduction

1.1 Active Galactic Nuclei

1.1.1 What is an Active Galactic Nucleus (AGN)

Active Galactic Nuclei (AGNs) are compact, highly luminous regions at the centers of some galaxies where the dominant power source is accretion onto a supermassive black hole (SMBH) with masses 10^6 10^{10} M (Kormendy and Ho, 2013). Gas and dust from the host galaxy lose angular momentum and spiral inward, forming an accretion disk whose viscous dissipation converts gravitational potential energy into radiation from the radio to the -ray regime (Rees, 1984; Netzer, 2013). In this sense, an AGN is better understood as a long-lived phenomenon - sustained energy in ow and release - rather than a single static object.

The immediate environment of the SMBH comprises several components that imprint characteristic spectral features: a photoionizing continuum from the accretion ow; a dusty, molecular structure that absorbs and re-emits in the infrared (Krolik and Begelman, 1988; Elitzur, 2008); fast, dense gas producing broad permitted lines; and more extended, lower-density gas responsible for narrow permitted and forbidden lines (Osterbrock and Ferland, 2006). Some AGNs also launch collimated relativistic jets, which are mostly observed in the radio band (Blandford et al., 2019). The out ows and radiation of the nucleus can interact with the interstellar medium, a process often referred to as AGN feedback (Fabian, 2012; Heckman and Best, 2014). Thus, in observation, AGNs can be distinguished from Galaxies by all these key signatures: extremely high luminosities, broad-band continuum emission from gamma-rays to radio, strong and often broad emission lines, rapid variability across all wavelengths, and, in some cases, signi cant polarization and radio jets.

A key unifying framework for AGNs is the so-called AGN uni cation model, which explains much of their observed diversity in terms of orientation and obscuration e ects (Antonucci, 1993; Urry and Padovani, 1995; Netzer, 2015). In this paradigm, Type 1 AGNs show both broad and narrow emission lines because the broad-line region (BLR) is directly visible, while in Type 2 AGNs, the BLR is hidden by the dusty torus and only narrow lines remain observable. This model has been re ned to incorporate factors such as accretion rate, host-galaxy environment, and evolutionary stage, but orientation and obscuration remain central to interpreting the phenomenology of AGN.

AGNs span a wide range of luminosities and host-galaxy properties and are observed across cosmic time. Their extreme energetics and variability make them useful laboratories for accretion physics and black hole growth, and they are central to studies of the coevolution

of SMBHs and galaxies (Kormendy and Ho, 2013; Padovani et al., 2017).

1.1.2 The Uni ed Model

The Uni ed Model of AGNs proposes that the diverse observed classes arise mainly from orientation e ects applied to a common physical structure, rather than fundamentally different kinds of nuclei (Antonucci, 1993; Urry and Padovani, 1995; Beckmann and Shrader, 2012). At the center lies a supermassive black hole surrounded by an accretion disk emitting strongly from optical to X-rays. Encircling the disk is the broad-line region (BLR), a collection of high-velocity, dense gas clouds producing Doppler-broadened permitted lines. This is, in turn, embedded within a dusty, molecular region most likely having an equatorial geometry, which can obscure the BLR depending on the viewing angle. Beyond the so-called dusty torus, the narrow-line region (NLR) contains a more extended lower-density gas that produces narrow permitted and forbidden lines.

In this framework, one can introduce a general classi cation of AGNs mainly based on the optical properties.

- Type 1 AGNs - Viewed closer to the face-on, allowing a direct line of sight to both the BLR and NLR; broad and narrow emission lines are observed. Type 1 AGNs often include intermediate types (e.g., Seyfert 1.5) that show a combination of broad and narrow lines, due to partial obscuration or intrinsic changes in the BLR.
- Type 2 AGNs - Viewed closer to the edge-on, where the dusty torus obscures the BLR; only narrow lines are visible in the optical spectrum.

Relativistic jets, when present, can be oriented at various angles to our line of sight; if closely aligned, they dominate the observed emission, producing the properties seen in one speci c class of AGNs called blazars. Although the uni ed model accounts for much of the variety in AGN appearance, additional factors, such as the accretion rate, evolutionary stage, and host-galaxy environment, also in uence the observed properties.

The AGN classi cation and nomenclature are generally more complex and are based on di erent observed properties from radio to high energies. In this thesis, we focus on the AGNs obtained through optical surveys, thus more detailed classi cation beyond Type 1 and Type 2 AGNs will be omitted.

The geometry of AGNs is commonly summarized by the Uni ed Model, which highlights how orientation and obscuration shape the observed properties of AGNs. A schematic view is shown in Fig. 1, where the central SMBH is surrounded by an accretion disk and BLR, embedded within a dusty torus that obscures the BLR at high inclinations. Beyond this, the NLR extends to larger scales, while relativistic jets may emerge along the polar axis. The gure further illustrates how radio-quiet Seyferts and radio-loud AGNs (FR I/II radio

galaxies, BLRGs, blazars) can be understood within the same framework, with additional observational e ects from absorption, scattering, and re ection along di erent sightlines.

Figure 1: Schematic illustration of the Uni ed Model, showing the SMBH, accretion disk, BLR, dusty torus, NLR, and relativistic jets. Adapted from Beckmann and Shrader (2012).

1.1.3 Broad-Line Region (BLR)

The Broad-Line Region is a compact zone of dense gas located within a subparsec to a few parsecs of the central SMBH. It is characterized by high-velocity motions (v 10^3 10^4 km s ^{1}), inferred from the Doppler broadening of permitted emission lines such as H, H , Mg II , and C IV (e.g., Peterson, 2004; Popovi¢, 2008). These velocities are consistent with gravitationally bound gas in the potential well of the SMBH, suggesting that the BLR is dynamically dominated by the central mass.

The ionizing continuum from the accretion disk photo-ionizes the BLR gas. The line strengths and ratios observed depend on the local gas density (typically in the range n_H 10^9 10^{11} cm ^{3}), the ionization parameter, and the chemical composition. Because the BLR is unresolved spatially with current telescopes, its geometry and kinematics are probed indirectly through techniques such as reverberation mapping (RM), which measures the time delay between continuum variations and the corresponding response in broad emission lines

(e.g., Peterson, 2004). This delay gives the characteristic BLR radius:

$$R_{BLR} \quad c \; ; \tag{1}$$

where c is the speed of light.

Under the assumption that the BLR clouds are in virialized motion around the SMBH (i.e., gravitationally bound to the SMBH), the black hole mass can be estimated via virial method:

$$M_{BH} = f \frac{R_{BLR} \quad V^2}{G} ; \tag{2}$$

where V is a measure of the width of the emission line (e.g., full width at half maximum, FWHM), G is the gravitational constant and f is a scale factor that accounts for the unknown geometry and inclination of the BLR (e.g., Peterson, 2004; Kollatschny et al., 2014; Fausnaugh et al., 2017).

Direct reverberation mapping is conceptually robust but observationally demanding because it requires long-term, high-cadence spectroscopic monitoring of individual AGNs to measure reliable time delays. Such campaigns are therefore limited to relatively small samples (e.g., Peterson, 2004). To extend black hole mass estimates to large surveys, an empirical radius luminosity (R L) relation has been established, linking the BLR size to the AGN continuum luminosity (Bentz et al., 2009). This relation makes it possible to use single-epoch spectra to estimate M_{BH} with the virial method, and it has become a key tool of AGN demographic studies and the scaling relation (Greene and Ho, 2005; Popovi¢, 2012). In this thesis, we follow this approach to derive black hole masses from SDSS and Gaia XP spectra, using the broad H line as the most robust and consistently measurable tracer.

The BLR is strati ed: high-ionization lines (e.g. C IV) are typically emitted closer to the SMBH, while low-ionization lines (e.g. H) originate farther out. This strati cation provides clues to the structure and dynamics of the BLR, which can range from a at disk-shaped con guration to more complex geometries involving in ows, out ows, or turbulence (Kuhn et al., 2020).

Understanding the BLR properties in terms of kinematical and physical properties is important for AGN physics because:

- It enables direct black hole mass measurements for Type 1 AGNs (via Eq. 2).
- Its variability encodes information about accretion physics and radiative transfer in extreme environments.
- It acts as a diagnostic of the ionizing continuum shape and the chemical enrichment history of the host galaxy's central regions.

In Type 1 AGNs, the BLR is directly visible, making them essential laboratories for BLR studies. In contrast, in Type 2 AGNs the BLR is hidden in the optical/UV due to obscuration by the dusty torus, although it may be detected in polarized light.

1.1.4 Type 1 AGNs

Type 1 Active Galactic Nuclei are those in which the line of sight provides an unobstructed view of both the BLR and the NLR, as predicted by the Uni ed Model (1.1.2). Their optical and ultraviolet spectra are typically dominated by broad permitted lines such as H , H , and Mg II , with full widths at half maximum (FWHM) ranging from about 1;000 to 10;000 km s ^{1}, produced by high-velocity BLR gas. Narrow permitted and forbidden lines, for example [O III] 5007, arise in the more distant, low-velocity NLR (v . 1;000 km s ^{1}). In addition, a strong continuum from the accretion disk spans from the infrared to the X-rays, often showing a prominent 'big blue bump' in the optical-UV.

The unobscured BLR in Type 1 AGNs enables black hole mass estimation through the virial method (Eq. 2), detailed studies of BLR geometry and kinematics through line pro le analysis, and reverberation mapping to measure continuum line time delays. Within this class, Seyfert 1 galaxies represent generally lower-luminosity systems in the nearby Universe, quasars are more luminous examples often found at higher redshift and capable of outshining their host galaxies, and broad-line radio galaxies are radio-loud counterparts of Type 1 AGNs.

In addition to the broad division into Type 1 and Type 2 AGNs, several important subclasses re ne this scheme. Intermediate Seyfert galaxies (types 1.2 1.9) display a combination of broad and narrow Balmer components, with H typically stronger and broader than H . Such intermediate features may result from partial obscuration of the BLR or intrinsic variations in line emission (Osterbrock and Ferland, 2006). Another subclass is Narrow-Line Seyfert 1 galaxies (NLSy1s), which belongs to the Type 1 family but is characterized by unusually narrow broad Balmer lines (FWHM H < 2000 km s ^{1}), strong Fe II emission, and frequently high Eddington ratios (Osterbrock and Pogge, 1985; Komossa, 2008). Several of the objects in our sample fall into this category, making NLSy1s especially relevant for the present work.

Another class often included in AGN demographics, is Low-Ionization Nuclear Emission-line Regions (LINERs). These nuclei show weak ionizing continua and spectra dominated by strong low ionization lines such as [O I], [N II], and [S II] (Heckman, 1980; Ho, 2008). Although they overlap phenomenologically with Seyferts, their excitation mechanism remains debated, and possible explanations include low-luminosity accretion, shocks, or stellar contributions. In population studies, LINERs are frequently grouped with low-luminosity AGNs, and we include them in Fig. 2 for completeness.

Together, these subclasses illustrate the diversity of optical AGN spectra and provide important test cases for orientation-based uni cation, accretion physics, and host nucleus interactions.

Although the term 'quasar' is sometimes used interchangeably with Type 1 AGN, it is strictly a luminosity-based designation, while the latter is a spectroscopic classi cation. This thesis focuses on Type 1 AGNs because their visible BLR provides the observational basis

for comparing broad-line measurements obtained from spectra of two very important optical sky surveys: Gaia space mission and ground-based Sloan Digital Sky Survey (SDSS) surveys. More details on the surveys and data sample will be provided later.

1.2 Optical Emission Lines of AGNs

Optical spectroscopy is one of the widely used tools for studying the physical conditions, kinematics, and ionization mechanisms in AGNs. By dispersing the nuclear light into its constituent wavelengths, it reveals diagnostic information from both the continuum and the emission line spectrum, which serve as diagnostics of the central engine and its surrounding gas. Typical AGN spectra display a strong non-stellar continuum overlaid with broad and narrow emission lines and, in some cases, stellar absorption features from the host galaxy.

Figure 2: Baldwin Phillips Terlevich (BPT) diagnostic diagram showing the separation between star-forming galaxies and AGNs based on the line ratios [O III] 5007/H and [N II] 6583/H . The solid line is the empirical division from Kau mann et al. (2003), while the dashed line is the theoretical maximum starburst line from Kewley et al. (2001). AGNs are further separated into Seyferts and low-ionization nuclear emission-line regions (LINERs). [Adopted from Porqueres et al. (2018)]

Emission lines are powerful probes of the ionization source and gas properties. The ratios of speci c diagnostic lines, such as [O III] 5007, [N II] 6583, [S II] 6716,6731, H , and H , are commonly used in diagnostic diagrams, such as the Baldwin Phillips Terlevich (BPT) diagrams (Baldwin et al., 1981) to separate AGNs from star-forming galaxies, i.e., H II regions dominated by star formation. In AGNs, the harder ionizing radiation eld, extending into the X-ray regime, produces higher ionization and stronger forbidden lines than those found in star-forming galaxies (Figure 2). High-ionization lines, such as [Ne V] 3426 and

[Fe VII] 6087, indicate the presence of an intense nonthermal continuum, whereas low-ionization lines like [O II] 3727 trace more distant or shielded gas.

The classi cation into broad and narrow emission lines is fundamental to the AGN taxonomy. Broadly permitted lines (e.g., H , H , Mg II, C IV) originate in the BLR, where high-velocity motions (FWHM 1;000 10;000 km s ^{1}) are driven by the gravitational pull of the supermassive black hole. Narrow permitted and forbidden lines (e.g., [O III] 5007, [N II] 6583) form in the NLR, which lies farther out and has smaller velocity dispersions (. 1;000 km s ^{1}). The relative strengths and velocity widths of these lines provide insight into the orientation, gas geometry, and ionization conditions of the AGN. As discussed earlier (Section 1.1.2), Type 1 AGNs show both broad and narrow lines, while in Type 2 AGNs the broad lines are obscured and only narrow lines are observed.

In addition to line ratios, optical spectra also reveal kinematic signatures: velocity shifts in emission lines can indicate bulk motions such as in ows or out ows, while asymmetric or broadened pro les may signal turbulence or rotation in a disk-like con guration. Time-resolved spectroscopy further enables reverberation mapping, where emission-line variability follows continuum changes with a measurable lag. This technique, introduced in Section 1.1.3, provides a powerful means to link the line-emitting gas to the central engine.

Although the AGN continuum often dominates, spectra may also contain stellar absorption features from the host galaxy, such as the Ca H&K lines, G-band, and Mg b triplet. These features are particularly evident in low-luminosity AGNs or in high signal-to-noise data. Accurate modeling and subtraction of host galaxy light is essential for precise emission-line measurements, especially when the nucleus-to-host contrast is low.

1.3 Motivation and Thesis Objectives

A direct comparison between SDSS spectroscopy and Gaia BP/RP spectrophotometry o ers a unique opportunity to evaluate the complementary strengths of two of the most in uential astronomical surveys of the past two decades. SDSS has set the benchmark for optical spectroscopic studies of AGNs, delivering ux-calibrated spectra covering 3800 9200 Å with a spectral resolution of (R 2000) (ranging from R 1500 at 3800 Å to R 2500 at 9200 Å) (Almeida and et al., 2023), providing su cient detail to resolve the shapes, widths, and substructures of broad and narrow emission lines.

In contrast, Gaia provides homogeneous, all-sky, low-resolution (R 20 100) slitless spectra across a broader wavelength range (330 1050 nm) with exceptional photometric stability and uniform ux calibration. Although Gaia is best known for its transformative impact on astrometry and photometry, Data Release 3 (DR3) has made spectrophotometric data available for hundreds of millions of sources, most of which remain largely unexplored in the context of AGN broad-line studies.

The motivation for this comparison arises from the complementary nature of the two datasets. SDSS delivers high- delity spectral information for a limited set of targets, whereas Gaia o ers uniform, multi-epoch spectrophotometry for an unprecedented number of sources across the entire sky. By focusing on Type 1 AGNs observed by both surveys, we can assess how well Gaia's low-resolution spectra reproduce key broad-line measurements such as integrated uxes and Balmer line ratios, as well as use them for the estimate of the mass of SMBH, and study the impact of spectral resolution, calibration di erences, and observational conditions. Ultimately, this will help us to understand under what circumstances Gaia XP spectra can serve as a practical substitute or complement to medium-resolution spectroscopy in large-scale AGN research.

The primary goal of this work is to evaluate the suitability of Gaia BP/RP (XP) spectra for quantitative broad-line measurements in Type 1 AGNs, using SDSS spectroscopy as the reference standard. Speci cally, our goal is to:

- Directly compare the integrated H and H broad-line and continuum uxes derived from Gaia DR3 XP sampled and continuous spectra with those measured from medium resolution SDSS spectra for a cross-matched set of Type 1 AGNs.
- Evaluate the agreement in the Balmer line ratios (H /H) between the two data sets and investigate systematic o sets due to spectral resolution, ux calibration, or aperture di erences. Emphasize how poor spectral resolution impacts measurements of the spectral parameters.
- Assess whether Gaia spectra could be use to measure the spectral features as well as to estimate important physical parameters of AGN, such as the SMBH mass and accretion rate.
- Test if Gaia XP spectra can serve as a viable complement or substitute for traditional medium-resolution spectroscopy in large-scale AGN studies.

This thesis is organized into seven chapters, each addressing a speci c component of the study. Chapter 1 provides the theoretical background on AGNs, and introduce the research objectives of this work. Chapter 2 describes the studied observational datasets, namely the SDSS spectroscopic and the Gaia DR3 BP/RP spectrophotometric data, as well as multi-epoch photometry. Chapter 3 outlines the methodology of spectral analysis and decomposition, which is mainly based on the open-source FANTASY package. Chapter 4 presents the core results of the thesis, whereas the Chapter 5 extends the study into the time domain by analyzing AGN variability using the SDSS Stripe82 and the Gaia photometry. Chapter 6 discusses the broader implications of the results in the context of AGN studies, and Chapter 7 concludes the thesis by summarizing the main ndings and their signi cance. The last part consists from the extensive list of references and Appendices.

Chapter 2

2 Datasets and Sample Selections

2.1 Datasets

2.1.1 Sloan Digital Sky Survey (SDSS)

The SDSS[1] is one of the most in uential astronomical projects of the past two decades, fundamentally transforming our understanding of the Universe through its combination of wide- eld imaging, spectroscopy, and public data releases (York et al., 2000). The rst light was achieved in 2000 at the Apache Point Observatory in New Mexico, using a dedicated 2.5 m wide- eld Ritchey Chrétien telescope (Gunn et al., 2006). From its inception, SDSS was designed to operate as a survey machine, with a highly e cient drift-scan imaging camera and a ber-fed, multi-object spectroscopic system capable of observing hundreds of targets simultaneously.

The original spectrographs were equipped with 640 bers of diameter 3^{00}, feeding light to blue and red channels that together covered wavelengths from approximately 3800 to 9200 Å at a spectral resolution of R 2000 (Gunn et al., 1998). Spectrophotometric calibration was performed using standard F-type subdwarfs observed on each plate, and automated pipelines provided wavelength calibration, ux calibration, and sky subtraction with remarkable uniformity. An intrinsic limitation of plate design was the radius of 55^{00} ber collision, which prevented bers from being placed closer than this on a single plate, but the overlap of plate coverage helped mitigate the e ect.

Over time, SDSS has gone through several major upgrades. The Baryon Oscillation Spectroscopic Survey (BOSS) in SDSS-III introduced a redesigned pair of spectrographs with 1000 bers of aperture 2^{00}, volume-phase holographic gratings, and new CCD detectors optimized for blue and red sensitivity (Smee et al., 2013). This extended the wavelength coverage to 3600 10 000 Å while maintaining R 2000 and signi cantly improving throughput. Later survey phases expanded target selection to include quasars, galaxies, stars, and time-domain sources, each supported by re ned calibration strategies.

The scale and impact of SDSS have been extraordinary. By the end of SDSS-IV, the survey had produced more than four million optical spectra of galaxies, quasars, and stars, along with a legacy of uniformly processed and publicly accessible data (Ahumada et al., 2020). The current SDSS-V phase marks a new era of robotic ber positioning and includes the Black Hole Mapper (BHM), Milky Way Mapper (MWM), and Local Volume Mapper (LVM),

[1] https://www.sdss.org/

o ering rapid, exible target acquisition and integral eld spectroscopy for unprecedented mapping of the local Universe (SDSS Collaboration et al., 2025).

The combination of medium spectral resolution, broad wavelength coverage, and exceptional spectrophotometric calibration has made SDSS one of the primary resources for extragalactic spectroscopy. In particular, it has been widely used in large-scale studies of AGNs using both photometry and spectral data (Vanden Berk et al., 2006; Richards et al., 2002). In addition, it has provided reliable measurements of broad emission lines such as Hand H , continuum luminosities, and derived physical parameters across diverse and large samples of AGNs (Shen et al., 2011; Rakshit et al., 2017; Ili¢ et al., 2023). SDSS also provided monitoring of a small sample of Type 1 AGNs for reverberation mapping (Shen et al., 2015). In general, its consistent data quality, long time baseline, and vast archival coverage make it an ideal foundation for selecting high-quality Type 1 AGN spectra for detailed analysis in this work.

2.1.2 Gaia Mission Overview and Scienti c Context

The Gaia[2] mission, launched by the European Space Agency (ESA) in December 2013, is one of the most ambitious space-based surveys in the history of astronomy (Gaia Collaboration et al., 2016). Operating from the Sun Earth Lagrange point L2, the spacecraft continuously scans the sky, observing each source multiple times over the course of its mission. Its primary objective is to construct a precise three-dimensional map of the Milky Way, allowing transformative studies of the structure, formation, and evolution of the Galaxy.

Gaia is equipped with two telescopes that share a single focal plane, which hosts 106 CCD detectors, one of the largest focal planes Fig. 3 ever own in space. The payload includes three main instruments: the astrometric instrument, which delivers microarcsecond-level measurements of stellar positions, parallaxes, and proper motions; the photometric instrument, which provides broad-band G photometry and low-resolution spectrophotometry through the Blue Photometer (BP; 330 680 nm) and Red Photometer (RP; 640 1050 nm); and the Radial Velocity Spectrometer (RVS), which obtains medium-resolution spectra over 845 872 nm for radial velocity measurements of bright sources. The BP and RP channels employ slitless prism dispersion optics, recording low-resolution spectra for nearly every detected object. Figure 3 illustrates this focal plane layout, showing how the di erent CCD strips are assigned to astrometric, photometric, and spectroscopic instruments through which each source successively passes.

Since its rst data were released in 2016, Gaia has revolutionized many elds of astrophysics. The mission has mapped the detailed structure and kinematics of the Milky Way's disk, halo, and bulge; discovered and characterized stellar streams, satellite galaxies, and tidal debris; provided stellar parameters for hundreds of millions of stars; re ned the cosmic distance scale through highly accurate parallaxes; and identi ed thousands of transient and

[2] https://www.cosmos.esa.int/gaia

Figure 3: Layout of the Gaia focal plane, which contains 106 CCD detectors divided among the Sky Mapper (SM), Astrometric Field (AF), Blue Photometer (BP), Red Photometer (RP), and Radial Velocity Spectrometer (RVS) instruments. The scanning motion moves stars from left to right across the focal plane over a period of approximately 10 seconds.

variable phenomena. Each successive data release has expanded the scope and precision of available measurements. The most recent, Data Release 3 (DR3) in 2022, includes astrometry for 1.8 billion sources, photometry for the same number, radial velocities for 33 million, and for the rst time low-resolution BP/RP spectra for over 220 million objects.

Despite having the main objective to map the stars of our Galaxy, Gaia has detected millions of quasar candidates. Recently, the all-sky quasar Quaia catalog has been released (Storey-Fisher et al., 2024) with 6,649,162 quasar candidates identi ed by the Gaia mission.

Gaia's photometry and astrometry have been widely recognized as groundbreaking, and its datasets have become a cornerstone of modern astrophysics. However, the low-resolution BP/RP (so-called XP) spectrophotometry remains far less explored despite its enormous scienti c potential. Where each XP spectrum is provided as a mean spectrum which basically consists of an array of coe cients to a set of basis functions (Carrasco et al., 2021). This underuse, combined with the unprecedented scale of available spectra, provides a compelling motivation to investigate their applicability for AGN studies, as pursued in this work.

2.2 Sample Selection

This study aims to evaluate the reliability of Gaia XP spectra for measuring broad emission line properties in Type 1 AGNs, using SDSS spectroscopy as a benchmark. Therefore, a common set of AGN targets observed by both surveys was required. The selection process began with identifying suitable SDSS spectra, chosen for their robust calibration, medium spectral resolution, and extensive legacy in AGN research, followed by cross-matching with

Gaia Data Release 3 (DR3) to obtain the corresponding XP spectra. The resulting data set comprises high-con dence matches that meet both the SDSS selection criteria and the availability of calibrated Gaia XP data, enabling direct, object-by-object spectral comparison.

2.2.1 Selection Criteria for Type 1 AGNs in SDSS

SDSS spectra formed the primary reference data set to evaluate the performance of Gaia XP spectrophotometry. The selection process followed established practices from large-scale spectroscopic investigations of Type 1 AGNs (e.g., Shen et al., 2011; Rakshit et al., 2017; Raki¢, 2022), which place particular emphasis on redshift constraints, data quality, and robust spectral classi cation to ensure reliable broad-line measurements.

Sources were selected from the SDSS DR18[3] based on the following criteria:

- Spectroscopic classi cation: Objects classi ed as QSO by the SDSS pipeline (Bolton et al., 2012), ensuring the inclusion of broad permitted emission lines characteristic of Type 1 AGNs.
- Redshift constraint: z . 0:4, to ensure that both H and H fall within the SDSS spectral range (3800 9200 Å).
- Continuum signal-to-noise ratio: S/N > 50 per pixel in the vicinity of both Balmer lines, following Raki¢ (2022), to allow accurate continuum placement and reliable emission-line tting.
- Spectral quality ag: ZWARNING = 0, excluding spectra with unreliable redshift solutions, strong sky residuals, or calibration artifacts.

To implement the above selection criteria in practice, we used the SDSS SkyServer SQL Search tool[4]. The following SQL query was executed to retrieve all Type 1 AGNs meeting the de ned constraints. The resulting table provided the necessary identi ers, and the corresponding calibrated spectra were then downloaded from the SDSS Science Archive Server (SAS)[5] using wget for subsequent analysis:

```
SELECT s.specobjid, s.bestobjid, s.z AS redshift,
       s.class, s.subclass, s.ra, s. dec,
       s.plate, s.mjd, s.fiberid, s.snMedian
FROM SpecObjAll AS s
JOIN PhotoObj AS p ON s.bestobjid = p.objid
WHERE s.class = 'QSO'
  AND s.z < 0.4
  AND s.zwarning = 0
  AND s.snMedian > 50
```

[3] https://www.sdss.org/dr18/
[4] https://skyserver.sdss.org/dr18/SearchTools/sql
[5] https://dr18.sdss.org/optical/spectrum/search?tab=bulk&run2d=any

These criteria result in a limited but clean, high-quality sample of low-redshift Type 1 AGNs that provides a solid basis for direct spectral comparisons between SDSS and Gaia.

2.2.2 Source Cross-Match Between SDSS and Gaia

The initial parent sample consisted of 77 SDSS spectra that met the selection criteria described in Section 2.2.1. This number is smaller compared to previous studies such as Raki¢ (2022); Ili¢ et al. (2023), primarily due to our adoption of a more stringent continuum signal-to-noise ratio threshold (S/N > 50), whereas Raki¢ (2022), for example, used a lower threshold of S/N > 30. Our somewhat smaller sample of objects was not relevant for our case study, since the motivation was not to study the complete population of type 1 AGN in SDSS and Gaia, but rather to assess the feasibility of AGN spectroscopy with Gaia.

To identify Gaia DR3 counterparts, we used the precomputed sdssdr13_best_neighbour table[6] (Marrese et al., 2017, 2019), which provides optimal cross-matches between SDSS DR13 and Gaia DR3 sources. The cross-match procedure is based on the o cial Gaia pipeline developed by (Marrese et al., 2017, 2019), which combines angular separation with the full astrometric covariance matrix, positional uncertainties, and local source density to assign a robust gure of merit and select the best neighbour.

The cross-match yielded 72 reliable Gaia sources whereas for only 60 sources spectral data was available. For each matched source, we retrieved both versions of the BP / RP spectra through Gaia@AIP service[7] an external data center hosted by the Leibniz Institut für Astrophysik Potsdam (AIP) that provides alternative access to Gaia DR3 data, since the mean spectrum data is not available in Gaia Archieve [8]. Continuous spectra, stored in the xp_continuous_mean_spectrum table, are available for all sources and are delivered in the form of basis-function coe cients. Sampled spectra, stored in the xp_sampled_mean_spectrum table, are provided only for a restricted subset of bright sources (typically G . 15) that additionally satisfy internal quality requirements, such as su cient on-board ux, multiple transits, and absence of problematic processing ags (De Angeli et al., 2023; Montegri o et al., 2023). In this work, we relied primarily on the continuous BP/RP spectra reconstructed from basis-function coe cients, which are available for all objects and thus ensure uniform coverage of the sample, while sampled spectra were used only for cross-checks when available.

In the nal matched set, 60 sources possessed continuous XP spectra, and two of these also had sampled spectra. These corresponded to 64 SDSS spectra in total, as four AGNs were observed at two distinct spectroscopic epochs. Our nal sample of 64 SDSS spectra was used in this analysis to preserve multi-epoch information and enable direct, object-by-object comparisons with Gaia XP data. The nal sample studied in this thesis, therefore, consists of 64 objects that were further analyzed in detail.

[6] https://gea.esac.esa.int/archive/documentation/GDR3/Gaia_archive/chap_datamodel/sec_dm_cross-matches/ssec_dm_sdssdr13_best_neighbour.html
[7] https://gaia.aip.de/
[8] https://gea.esac.esa.int/archive/

2.3 Properties of Gaia XP Data

The Gaia Blue and Red Photometers (BP and RP) provide low-resolution slitless spectrophotometry, collectively known as XP spectra. In Data Release 3 (DR3), these spectra are available in two primary formats: sampled and continuous.

The sampled spectra contain ux measurements at discrete wavelength points, derived by calibrating the raw prism observations onto a common wavelength grid and applying corrections for instrumental e ects such as the line spread function (LSF) and wavelength dispersion (Carrasco et al., 2021; Montegri o et al., 2023). This representation o ers a direct view of the observed spectral energy distribution (SED) and is well-suited for applications that require ux values at xed wavelengths.

In contrast, the continuous spectra are stored as coe cients of orthonormal Hermite basis functions for the BP and RP channels, providing a smooth and compact representation of the SED over roughly 330 1050 nm (De Angeli et al., 2023). This format accommodates variations in the resolution of the dispersed photometry of Gaia and allows spectra to be reconstructed at any desired sampling resolution.

2.3.1 Gaia XP Data Retrieval

For this study, Gaia XP spectra were obtained programmatically via the Gaia@AIP[9] TAP service using Virtual Observatory (VO) standards. After cross-matching the SDSS parent list with Gaia DR3 through the precomputed sdssdr13_best_neighbour table (Section 2.2.2), we assembled the corresponding set of source_id s and accessed the XP spectrophotometry with pyvo (Graham and Group, 2014) by issuing ADQL queries and managing results as VOTables and CSV les.

Sampled XP spectra. For sources with entries in the sampled product, we queried the xp_sampled_mean_spectrum table to obtain uxes and associated uncertainties on Gaia's native pixel grid. These spectra are already provided in calibrated ux units by the DR3 spectrophotometric pipeline, with corrections applied for wavelength solution, throughput, and the line-spread function (Carrasco et al., 2021; Montegri o et al., 2023; De Angeli et al., 2023). Data were retrieved via the TAP service using Python, running synchronous queries for small lists and asynchronous jobs for larger batches, and saved directly as CSV les. Only light formatting was applied before analysis, as no reconstruction was required.

Continuous XP spectra. For targets with continuous BP/RP spectra, the matched source_id list stored as a pandas Series was uploaded to the Gaia@AIP TAP service as a temporary VO table. An asynchronous ADQL query was then executed against the

[9] https://gaia.aip.de/

xp_continuous_mean_spectrum table, joining on the uploaded IDs. Once the job completed, the resulting table was retrieved into a pandas DataFrame. Using the DataLink service, the corresponding per source BP and RP basis coe cient les were downloaded as CSVs. These individual les were saved locally and later merged into a single consolidated dataset for subsequent reconstruction and analysis.

2.3.2 Calibration of Continuous XP Spectra

The merged Gaia XP continuous dataset, containing BP and RP basis-function coe cients for all matched sources, was converted into calibrated ux spectra using the GaiaXPy Python package[10]. The calibrate function reconstructed each spectrum on a uniform vacuum wavelength grid, applying the DR3 spectrophotometric calibration model (De Angeli et al., 2023; Montegri o et al., 2023).

Before proceeding to resampling, we validated the output by comparing our reconstructions to the independent Gaia DR3 XP spectra published by the ARI's Gaia Services library[11]. The comparison, shown in Figure 4, demonstrates near-perfect agreement in spectral shape and continuum level for the same source (Gaia DR3 source ID: 3990830907900291584), con- rming the delity of our calibration pipeline.

By default, GaiaXPy outputs spectra sampled at 2 nm intervals. While this preserves the continuum shape, the coarse pixel scale reduces the number of data points across broad emission lines, which can bias integrated ux measurements. Direct comparison with SDSS showed that such under-sampling introduced systematic o sets in line uxes, even though the spectra appeared visually consistent.

Through iterative testing, the sampling was re ned to 0.2 nm (2 Å), closely matching the SDSS pixel scale (1 2 Å per pixel). This provided a comparable number of points per line pro le in the Gaia spectra, resulting in much closer agreement in integrated line uxes. Tests at 2 nm and even coarser 20 nm con rmed that while the overall spectral shape remained unchanged (see Figure 5), di erences in pixel size produced measurable discrepancies in integrated ux values.

The calibrated, high-resolution Gaia spectra produced through this process served as the input for subsequent preparation steps prior to spectral tting (Section 2.4).

2.3.3 Preparation of Calibrated XP Spectra for Fitting

After calibration (Section 2.3.2), the Gaia XP spectra were preprocessed for analysis with FANTASY. This step consolidated survey metadata, enforced a consistent wavelength/ ux convention, and produced spectrum objects consumable by the tting pipeline.

[10] https://gaia-dpci.github.io/GaiaXPy-website/
[11] https://gaia.ari.uni-heidelberg.de/

(a) This work (GaiaXPy+ Fantasy)

(b) ARI Gaia DR3 XP Library

Figure 4: Comparison of GaiaXPy calibrated XP spectrum from this work (top) and the independent ARI XP spectrum (bottom) for the same AGN (Gaia DR3 source ID: 3990830907900291584). The excellent match in continuum and Balmer line features supports the accuracy of the GaiaXPy calibration.

First, each calibrated XP record was augmented with astrometry and spectroscopic redshift from the matched SDSS entry (RA, Dec, z), enabling a one-to-one association across surveys and ensuring that all spectra carry the same reference information. To streamline ingestion, a small Python loader, read_gaia, was developed (full listing in Appendix 7). The loader operates on the combined XP CSV export and performs the following operations for a selected source_id (or row index): it parses the serialized arrays of flux and flux_error from the CSV (robust to minor formatting artifacts), drops malformed rows, and selects the requested source. Flux densities are then converted from SI to cgs and scaled to the conventional SDSS numeric unit: the code applies the exact relation $1\ \mathrm{W\,m^{-2}\,nm^{-1}} = 100\ \mathrm{erg\,s^{-1}\,cm^{-2}\,Å^{-1}}$ and multiplies by 10^{17} so that values are expressed in units of $10^{-17}\ \mathrm{erg\,s^{-1}\,cm^{-2}\,Å^{-1}}$, matching typical SDSS spectral products. The corresponding uncertainties are transformed identically.

For the wavelength axis, read_gaia constructs a uniform vacuum-wavelength grid by linearly sampling the calibrated XP bandpass from 336 to 1020 nm and converting to Ångströms.

Figure 5: Gaia XP spectrum of a representative source (source ID: 6342859972491701 76) calibrated using GaiaXPy with two di erent wavelength samplings: 2 nm (solid blue line) and 0.2 nm (dashed red line with markers). The two samplings produce indistinguishable spectral shapes, but the ner sampling provides more points across emission lines, improving integrated ux accuracy.

The number of samples is taken from the length of the ux vector in the GaiaXPy output, so the wavelength step re ects the chosen reconstruction sampling (e.g. 0.2 nm). The loader also carries forward the original_ext_source_id (used here as bestobjid) to preserve the cross-survey identi er, and assigns a human-readable name (Gaia_$source_id$) for plotting and bookkeeping.

The outputs of read_gaia (wave, flux , err , RA/Dec, z, IDs) are then wrapped into FANTASY Spectrum objects via make_spec, which standardizes metadata and arrays for downstream continuum and emission-line modeling. Wavelengths are treated in vacuum Å at this stage; any optional vacuum-to-air transformation (if desired to mirror speci c SDSS reductions) is applied consistently within the tting pipeline to avoid mixing conventions. This preparation ensures that the Gaia XP and SDSS spectra enter the analysis with compatible units, sampling, and metadata, allowing direct, object-by-object comparison of broad-line uxes and ratios in Section 2.4.

Before nalizing our use of the GaiaXPy-calibrated spectra for tting, we also tested an alternative empirical correction approach from Huang et al. (2024)), described in the following subsection.

2.3.4 Testing XP Spectra Corrections

While GaiaXPy provides calibrated XP spectra using the DR3 spectrophotometric model, recent work by Huang et al. (2024) identi ed residual systematics particularly ux discrepancies and continuum slope biases correlated with color and magnitude. To assess whether

these e ects impact our AGN sample, we tested their publicly released correction code[12] on all Gaia XP spectra used in this study.

Huang et al. (2024) developed their empirical corrections by comparing Gaia XP spectra of stars and galaxies with external high-quality spectrophotometric libraries. They constructed correction functions as a function of color and magnitude, reporting wavelength-dependent discrepancies up to 10%, especially in the blue end (. 4000Å) and for faint sources (G & 17.5). Their corrections signi cantly improve color indices and continuum shapes when benchmarked against template libraries.

We adapted their Python code to our AGN dataset, applying it to the XP spectra after ux calibration with our standard GaiaXPy pipeline. A representative comparison between the original and corrected spectra is shown in Figure 6. The corrections slightly modi ed the continuum slope in the blue but did not a ect the shape or integrated ux of broad emission lines such as H or H . Given this minimal impact, we chose not to use the corrected spectra in our nal analysis.

All subsequent results are based on the original GaiaXPy-calibrated spectra with optimized wavelength sampling (Section 2.3.2).

Figure 6: Comparison between the original Gaia XP spectrum (green) and the corrected spectrum using the (Huang et al., 2024) code (orange) for a representative AGN. Only minor di erences are observed in the blue continuum; broad emission lines remain una ected.

2.3.5 Comparison of Gaia and SDSS Spectra

To illustrate the comparison between Gaia XP and SDSS spectra, we present in Figure 7 one representative case from our sample. The SDSS spectrum (green) provides high-resolution measurements of the continuum and emission lines (e.g., H, H , [O III]), while the Gaia XP spectrum (red) reproduces the overall continuum shape and broad features but lacks the ne detail due to its lower spectral resolution. For clarity, the Gaia spectrum is vertically shifted, while ux calibration is preserved.

[12] https://github.com/Jiaxuan-Huang/XP-correction

Figure 7: Comparison of Gaia DR3 XP spectrum (red) and SDSS spectrum (green) for one representative source. The Gaia spectrum is shifted vertically for clarity while ux calibration is preserved.

It is important to note that this gure serves only as a visual illustration. In the analysis presented in this thesis, no direct spectrum-to-spectrum comparison was performed. Instead, the measured spectral properties such as continuum uxes, emission-line uxes, line widths, and Balmer line ratios were extracted from both datasets and compared quantitatively. The complete set ofGaia SDSS spectra for the full sample is provided in Appendix B (Figure B1 onward).

2.4 Light Curves

2.4.1 SDSS Stripe 82 Light Curves

The SDSS carried out repeated imaging along the celestial equator in the Southern Galactic Cap, in a 2:5 -wide region known asStripe 82 ($20^h < \quad < \quad 4^h$, j j < 1:26 ; York et al. 2000; Abazajian and et al. 2009). Between 1998 and 2007, this region was observed 60 90 times in each of the ve broad optical lters (u; g; r; i; z), producing one of the deepest and most uniform multi-epoch optical datasets available. Stripe 82 contains light curves for several million sources, including stars, galaxies, and quasars, and has become a benchmark for time-domain astronomy.

Each epoch provides calibrated point spread function magnitudes with typical photometric precision of 0:01 mag for bright sources, anchored to a dense network of standard stars (Ivezi¢ et al., 2007; Sesar et al., 2007). The cadence is highly irregular: many observations are separated by 1 days, but long seasonal gaps of 300 days are also common. This combination of intra-night sampling, long-term monitoring (up to a decade), and high photometric accuracy makes Stripe 82 uniquely valuable for variability studies.

AGN in particular bene t from these properties: they exhibit stronger variability at shorter wavelengths, with larger amplitudes in the bluer u and g bands compared to the redder i and z bands (Ivezi¢ et al., 2004). Repeated observations across multiple lters also

improve the separation of quasars from stars in the color-color and variability space. Stripe 82 has since served as a reference data set for calibrating and validating subsequent surveys such as Pan-STARRS and Rubin Observatory Legacy Survey in Space and Time (LSST).

In this work, we used the publicly available Stripe 82 variable source catalog[13] light curves for a subset of our AGN sample, which was directly downloaded from the catalog archive. The catalog contains 67,507 sources with a master le for summary of the sources and for each light curve a separate dat le naming by speci c ID including mjd (modi ed Julian date), band, mag and mag_err. To nd out speci c ID of our target sources, we cross-match our sources with the master le by coordinates(ra,dec) using python, which were further converted into uxes for quantitative variability analysis (see Section 5).

An example of an SDSS Stripe 82 light curve from our AGN sample is shown in Fig. 8. The source LC_3861327 was observed repeatedly for nearly a decade across the ve SDSS lters. The light curve demonstrates the irregular but seasonally clustered cadence of Stripe 82, with dense sampling during observing seasons followed by long annual gaps. Intense intrinsic variability is detected in all bands. The amplitude is strongest in the bluer lters (u and g), while the redder bands (and z) show smaller but still signi cant variations. This behavior, with wavelength dependent variability strength, is consistent with AGN emission processes and highlights the reliability of Stripe 82 for multi-band variability studies.

Only 7 AGN in our spectroscopic sample have Stripe 82 coverage. This limited overlap arises because Stripe 82 is con ned to a narrow equatorial strip, while most of our AGN fall outside this region. In contrast, Gaia provides nearly all-sky coverage, which explains the much larger number of Gaia light curves available for our analysis.

2.4.2 Gaia Epoch Photometry Overview

The third Gaia Data Release (DR3) substantially expanded Gaia's time-domain capabilities by publishing per-transit epoch photometry in the G, BP, and RP bands (Vallenari et al., 2023). These data are derived from the rst 34 months of operations (mid-2014 to late 2017) and include calibrated magnitudes, uxes, and uncertainties for nearly 10 million variable sources in 24 variability classes (Eyer et al., 2023). For the rst time, Gaia provided light curves that combine the exceptional stability of space-based photometry with dense, homogeneous all-sky coverage.

Unlike SDSS, which followed a survey schedule designed for imaging campaigns, Gaia observes according to its scanning law. This results in an irregular cadence: sequences of transits can occur within a few days, followed by gaps of weeks to months. Although this non-uniform sampling di ers signi cantly from the seasonal pattern of Stripe 82, the individual Gaia epochs achieve subpercent photometric precision. This precision, combined with the space-based uniformity, makes Gaia epoch photometry particularly powerful for

[13] https://faculty.washington.edu/ivezic/sdss/catalogs/S82variables.html

Figure 8: Example of an SDSS Stripe 82 light curve from our sample (source ID: LC_386132). Multi-band observations over nearly a decade are shown in the five SDSS filters, illustrating irregular but seasonally clustered cadence and wavelength dependent variability amplitude.

tracing variability in AGN, even when the intrinsic signal is subtle.

For this work, we retrieved Gaia epoch photometry for 48 AGN sources matched to our spectroscopic sample from the table gaiadr3.vari_summary . For this we uploaded a VO table for our 64 target sources to check the availability. A sql query was done by python to find out our target sources and downloaded later creating datalink for epoch_photometry and saved as a separate csv file for each sources. The following SQL query was used to do the job,

```
SELECT s.source_id, s.ra, s.dec, s.parallax, s.has_epoch_photometry
FROM gaiadr3.gaia_source AS s
JOIN tap_upload.gaia_sourceids AS w ON w.source_id = s.source_id
WHERE s.has_xp_sampled = 'True' OR s.has_xp_continuous = 'True'
```

This is a considerably larger sample than the 7 AGN with Stripe 82 coverage, reflecting both Gaia's all-sky footprint and the extensive reach of DR3 epoch tables. Each Gaia light curve was downloaded from the DR3 archive and converted from magnitudes to flux densities to ensure consistency with the spectroscopic measurements presented in this thesis. These data provide an essential complement to the SDSS Stripe 82 light curves: while Stripe 82 offers a decade-long baseline with multiband coverage, Gaia contributes high-precision space-based monitoring over shorter baselines but for a far larger fraction of the sample. Together, the two data sets allow the investigations of the variability of AGN across different time scales and observational regimes, which will be addressed in Chapter 5.

Figure 9: Example of anGaia light curve from our sample (source ID: 3267488664984941184), same source as for SDSS light curve Figure 8.

Chapter 3

3 Spectral Analysis

3.1 Spectral Fitting Approach with FANTASY

3.1.1 Framework

For the spectral fitting in this work, we employed the open-source code FANTASY[14] (Fully Automated pythoN tool for AGN Spectra analYsis; (Rakić, 2022; Ilić et al., 2023), a Python-based package designed for simultaneous multi-component fitting of AGN spectra. The software is optimized for the optical rest-frame range (3600–8000 Å), but is also applicable to the ultraviolet (1800–3600 Å) and near-infrared (8000–11000 Å) ranges. It is built on the Sherpa Python package (Burke et al., 2022) and is installable via pip, ensuring accessibility and reproducibility.

The framework performs simultaneous fitting of the underlying continuum, modeled with a broken power law, together with sets of emission lines that can be chosen from predefined lists or specified by the user. This flexibility allows the fitting procedure to be adapted to different wavelength ranges, spectral resolutions, signal-to-noise ratios, and object types—a crucial advantage for heterogeneous datasets such as the SDSS and Gaia XP spectra used in this study.

A distinctive feature of FANTASY is its treatment of the complex Fe II emission, which can produce thousands of blended transitions overlapping with key AGN lines such as H, H, and in some cases H. The Fe II model is based on atomic parameters and grouping schemes from Kovačević et al. (2010) and Shapovalova et al. (2012), assuming that lines sharing the same lower energy level have fixed intensity ratios determined by oscillator strengths. Multiplets covering the 3800–11000 Å range are included, along with specific templates such as that of I Zw 1 (Kovačević et al., 2010). For strong Fe II emitters, e.g., narrow-line Seyfert 1 galaxies, additional forbidden Fe II transitions can be incorporated following Véron-Cetty et al. (2004). These templates are part of the broader predefined line lists provided by FANTASY, which also include the hydrogen Balmer and Paschen series, helium lines (He I and He II), narrow forbidden lines (e.g., [O III], [N II]), other AGN narrow and broad lines, and coronal lines. The framework also allows users to create new lists, modify existing ones, and constrain fitting parameters such as the maximum allowed line width (given with the FWHM parameter) or the minimum permitted wavelength shift.

Another important capability is the subtraction of the host-galaxy contribution, ensuring

[14] https://fantasy-agn.readthedocs.io/en/latest//

that the ts focus on the AGN's nuclear emission. This is achieved via principal component analysis (PCA), which models the observed spectrum as a linear combination of galaxy and quasar eigenspectra. In the SDSS-optimized approach (Yip et al., 2004b,a; Vanden Berk et al., 2006), xed sets of components derived from large SDSS samples are used. A more exible variant employs the full set of available eigenvectors, typically 10 for the stellar population and 15 for the quasar component, and determines the optimal number by minimizing the ^{2} statistic. In both methods, weighted tting is applied to reduce the in uence of strong emission lines, and if the host contribution is negligible, the process terminates automatically. Reliable decomposition is achieved when tting over a wide wavelength range (3600 8000Å), which maximizes the recovery of stellar absorption features.

This combination of exibility, physically motivated templates, and automated handling of both continuum and emission-line components makes FANTASY particularly suitable for this thesis, where spectra from di erent instruments and resolutions must be modeled in a consistent manner.

3.1.2 SDSS Line- tting Model

The SDSS spectra were analyzed following an approach optimized for isolating broad- and narrow-line region (BLR/NLR) features while accounting for possible host-galaxy contamination. Out of the 64 spectra in our sample, 19 exhibited a measurable stellar continuum contribution, which was removed using the PCA method described in 3.1.1, which is also available as one of the procedures in FANTASY. To avoid noisy edges and focus on key diagnostic lines, all ts were performed over the wavelength range 4000 7500Å, ensuring coverage of H , H , helium lines, and major forbidden lines such as [O III] and [S II].

The continuum was tted using a broken power-law anchored in emission-line free intervals. When Fe II emission was present, empirical templates were included to better isolate other spectral components, in particular the H line where the Fe II contamination is the strongest.

For the broad hydrogen Balmer lines (H , H , and, when detected, H and H), a mode 2 con guration was applied: two Gaussian components were used for each line to represent the velocity-dispersed core and the extended wings. This approach captures the complex BLR kinematics, including high-velocity tails. Helium lines (He I and He II) were tted with single Gaussian components, allowing physically motivated exibility in widths and centroid shifts.

The [O III] 4959, 5007 and [S II] 6716, 6731 doublets were modeled with tied Gaussian pro les, xing their ux ratios to theoretical values. In cases where the [O III] 5007 pro le displayed a pronounced blueward wing indicative of out owing ionized gas an additional Gaussian component was introduced, independently parameterized in width and velocity shift to separate it from the systemic NLR emission.

This targeted strategy provided robust separation of continuum, broad-line, and narrow-line contributions, enabling reliable measurements of emission-line uxes, widths, and velocity o sets. These measurements form the basis for subsequent physical interpretation and direct comparison with the lower-resolution Gaia XP ts.

3.1.3 Gaia Line- tting Model

The Gaia XP spectra were modeled using the FANTASY framework with a simpli ed, line-by-line approach tailored to much lower resolution (R 50 100) and modest signal-to-noise ratio of the data. The ts were generally performed over the 4000 7500Å range; however, for some objects it was not possible to achieve a satisfactory continuum placement across the full interval. In such cases, the spectrum was divided into two cropped regions (around H and H respectively) and tted independently. This strategy improved continuum determination and line recovery, though at the cost of losing a globally consistent continuum model.

The principal focus was on the broad hydrogen Balmer lines, with H and H modeled in all cases, and H and H included when su ciently prominent. Each line was tted with its own single broad component rather than enforcing a tied multi-line con guration. This approach provided the exibility needed to accommodate pro le di erences between lines, which may arise from BLR kinematics or local continuum uncertainties ampli ed by the low resolution.

Helium lines, most often He I 4471 and He I 5877, were added where detectable and modeled with single Gaussian components. Unlike the SDSS spectra, narrow emission lines were not explicitly modeled, as the XP resolution does not allow reliable decomposition of narrow versus broad components.

An additional complexity arises from Fe II emission, which produces broad blended features overlapping with the Balmer wings and often shaping the pseudo-continuum in the optical range. At the spectral resolution of SDSS, these blends can be modeled using empirical templates and models (e.g., Kova£evi¢ et al., 2010; Shapovalova et al., 2012; Ili¢ et al., 2023), allowing their contribution to be separated from the underlying continuum and the Balmer components. In contrast, in Gaia XP spectra the limited resolution and smoothing of features make Fe II emission di cult to isolate, as its contribution becomes degenerate with the continuum slope or with the broad wings of H and H . For this reason, Fe II was not explicitly modeled in our XP ts, though its potential presence should be kept in mind as a source of systematic uncertainty when interpreting Balmer line properties.

This simpli ed but exible approach enabled us to extract reliable measurements of the integrated uxes of the main broad Balmer lines, which are the key quantities required for comparison with SDSS results. The limitations inherent to the Gaia ts in particular the challenges of continuum placement, the treatment of Fe II blends, and the inability to model narrow lines are illustrated through representative examples in the following section.

3.2 Example Fits for Representative Objects

3.2.1 SDSS Fits

While all 64 SDSS spectra in our sample were tted with the FANTASY framework, here we present a set of representative cases that illustrate the range of challenges encountered during the analysis (Figures 10 - 15). These examples were selected to highlight di erent aspects of the tting strategy, including high-quality baseline ts, spectra dominated by host-galaxy light, strong Fe II contamination (e.g., in NLSy1 type of objects), unusually broad and complex Balmer-line pro les, and sources with only weak or narrow broad-line components.

By showcasing these diverse scenarios, we demonstrate both the exibility of the tting procedure and the robustness of the measurements derived from heterogeneous data. Each case is discussed with reference to the tted spectrum, emphasizing the speci c modeling considerations and the physical interpretation of the results.

Case 1 (Good Fit): This spectrum represents an ideal example of a well-constrained SDSS t (Figure 10). The continuum is stably anchored using line-free regions, while Fe II emission is well reproduced by the template. The broad Balmer lines are clearly decomposed into two Gaussian components, capturing both the line cores and extended wings, and the narrow lines (e.g., [O III], [N II]) are consistent with their expected ratios. Residuals across the tted wavelength range remain small, indicating that the main physical components continuum, Fe II, broad and narrow emission are all reliably accounted for. This case demonstrates the level of precision achievable in spectra with good signal-to-noise and modest host-galaxy contribution.

Figure 10: Example of an SDSS spectrum with a well-constrained t (Case 1: Good Fit). The observed spectrum (gray) is reproduced by the global model (red), with contributions from the continuum (dark blue dashed), Fe II template (yellow-green), broad Balmer components (blue), and narrow lines including [O III] and [N II] (magenta). Residuals are low across the wavelength range, highlighting the robustness of the t.

Case 2 (Strong Host Galaxy Contribution): In some SDSS spectra, the stellar continuum from the host galaxy dominates over the AGN emission, signi cantly altering the apparent spectral shape. This e ect is especially strong in objects at lower luminosities or with prominent absorption features from the underlying stellar population. Figure 11 shows such a case, where the raw spectrum exhibits strong host contamination. The upper panel presents the decomposition into galaxy and quasar eigenspectra following the PCA approach, illustrating how stellar absorption features and the overall red slope are accounted for. After subtraction, the AGN continuum and broad emission-line wings become more evident. The lower panel displays the nal multi-component t, where the continuum, Fe II templates, narrow and broad emission lines are modeled consistently. The host subtraction is essential here, as without it the broad Balmer components, particularly H , would be underestimated due to dilution by the stellar continuum.

Figure 11: Example of an SDSS spectrum with strong host galaxy contamination. Top: PCA-based decomposition of the observed spectrum (blue) into AGN (green) and host galaxy (orange) components, with the total model shown in red. Bottom: Final spectral t after host subtraction, including continuum, Fe II emission, and broad/narrow line components.

Case 3 (Strong Fe II Emission): In some spectra, the tting is complicated by the presence of strong Fe II emission blends, such as typically seen in NLSy1 galaxies. These multiplets consist of thousands of transitions, many of which overlap with key AGN diagnostic lines, especially in the4400 5500Å region around H and He II. The example shown in Figure 12 illustrates this case, where the broad Fe II pseudo-continuum dominates the spectrum blueward of H , creating di culties in isolating the true wings of the Balmer line.

To address this, we incorporate empirical Fe II templates into the t, following the model from (Ili¢ et al., 2023), which is based on studies of (Kova£evi¢ et al., 2010) and (Shapovalova et al., 2012). This model constrain the relative intensities of multiplets, while allowing the overall normalization, width, and small velocity shifts to vary within physical limits. As a result, the Fe II contribution can be e ectively disentangled from both the AGN continuum and the Balmer lines, producing more reliable measurements of broad-line uxes and kinematic properties.

Although such spectra are more challenging to model, the combination of exible Fe II templates and simultaneous multicomponent tting ensures that the recovered line parameters remain consistent with the physical interpretation of the broad-line region.

Figure 12: Example of an SDSS spectrum with strong Fe II emission features. The broad Fe II multiplets produce extensive blends across the optical range, particularly blueward of H , complicating the identi cation of the Balmer line wings. Incorporating empirical Fe II templates into the t allows reliable recovery of both the continuum and the broad-line pro les.

Case 4 (Very Broad H Component): In some spectra, the H line exhibits an extremely broad component that extends well into the wings beyond the narrow-line contribution. Figure 13 illustrates this case, where the tting required two Gaussian components to adequately capture both the core and the extended wings of the broad H pro le. The continuum is relatively well constrained across the optical window, and narrow forbidden lines are modeled separately. Despite the challenges of isolating the broadest emission from the underlying continuum and nearby narrow lines, the tting successfully reproduces the observed pro le, allowing reliable ux and width measurements. Such spectra are representative of objects with particularly high-velocity gas in their broad-line regions.

Case 5 (Weak or Narrow Broad Component): In some spectra, the broad Balmer component is only marginally detected and appears unusually narrow compared to typical Type 1 AGN pro les. This is illustrated in Figure 14, where the broad H and H wings

Figure 13: Example of an SDSS spectrum with an exceptionally broad H emission line. The model includes two broad Gaussian components (solid and dashed blue) to capture the extended wings in addition to the narrow-line contribution. The t demonstrates how extreme BLR velocities can still be constrained within the FANTASY framework.

are weak relative to the strong narrow-line emission. Such cases pose a challenge for reliable decomposition, as the tting algorithm can recover only a modest, broad component that is easily confused with the narrow lines. The continuum and Fe II emission remain well constrained, but the Balmer decomposition requires careful interpretation, and the measured broad-line uxes should be regarded as lower limits. This case highlights the importance of distinguishing between genuine weak broad-line AGNs and sources where observational limitations reduce the detectability of the broad wings.

Figure 14: Example of an SDSS spectrum where the broad Balmer component is weak and narrow, making it di cult to separate from the strong narrow-line emission. The t includes continuum, Fe II, narrow forbidden lines, and Balmer components, with the residuals indicating that the broad component is only marginally detected.

Case 6 (Dominant Narrow Emission Lines): Some spectra in the sample exhibit the typical appearance of Type 2 AGNs, where the emission is dominated by narrow lines

with no evident broad Balmer component. Figure 15 illustrates such a case: the strong [O III] 4959, 5007 and [N II] 6548, 6583 lines are clearly resolved, along with narrow H and H . The continuum is weak, and the absence of detectable broad wings distinguishes these spectra from the Type 1 AGN population.

This behaviour is consistent with the AGN uni cation picture, in which the broad-line region is obscured from view while the narrow-line region remains visible. In such cases, the decomposition is dominated by the forbidden lines and the narrow Balmer components, and the resulting ts highlight the prominence of the narrow-line region contribution.

Figure 15: Example of an SDSS spectrum where the narrow forbidden lines ([O III] and [N II]) are exceptionally strong, partially blending with the broad Balmer emission. The t includes continuum, Fe II, narrow forbidden lines, and Balmer components. While the decomposition reproduces the observed spectrum well, the broad-line uxes remain uncertain due to the dominance of the narrow features.

3.2.2 Gaia Fits

The general tting approach for Gaia XP spectra was described in Ÿ3.1.3, here we present representative examples that illustrate the practical challenges and outcomes of the modeling (Figures 16 - 18). Because of the lower spectral resolution and strong line blending in XP data, the ts are often less constrained than for SDSS, and continuum placement can strongly in uence the recovered line uxes. The following cases demonstrate typical situations encountered in our sample, ranging from satisfactory ts to spectra requiring special treatment.

Case 1 (Good Fit): In this case, the Gaia XP t demonstrates a stable continuum placement and well-separated broad Balmer components, particularly the strong H feature and a resolved H . Although the lower resolution of the XP spectra smooths out narrow-line features and any potential Fe II emission, the main broad-line pro les are nonetheless reproduced reliably (Figure 16). The close agreement between the observed spectrum (gray) and

the total model (red) highlights the consistency of the t, making it an illustrative example of a good Gaia spectrum where integrated broad-line uxes can be measured with con dence.

Figure 16: Example of a well-constrained Gaia XP spectrum. The continuum is well de ned across the tting range, and the principal Balmer lines (H , H , H) are clearly recovered with minimal blending. The residuals remain small and structureless, indicating that the simpli ed broad-line only model provides a robust representation of the data despite the limited spectral resolution.

Case 2 (Separate Fits for H and H): In some spectra, it was not possible to achieve a consistent continuum t across both the H and H regions simultaneously. Figure 17 illustrates this case, where the spectrum was divided into two cropped intervals and each Balmer line was tted independently. This approach ensures a more accurate continuum placement in each region and prevents cross-contamination of line measurements, though at the expense of a fully global t.

Case 3 (Challenging Spectrum with Strong Fe II Blends): A subset of Gaia XP spectra is dominated by broad Fe II emission, which at the XP resolution appears as smooth blends overlapping with the Balmer wings. Unlike in SDSS spectra, where empirical templates can disentangle Fe II from the continuum and Balmer components, the lower resolution of Gaia makes a reliable decomposition infeasible. Figure 18 shows an example in which Fe II contamination severely hampers the recovery of H and H pro les.

For such spectra, integrated uxes were still measured for completeness. These cases highlight the intrinsic limitation of Gaia XP data for AGN spectral analysis and demonstrate the need for cautious interpretation.

3.2.3 Comparison of SDSS and Gaia Fits

The preceding sections illustrated representative cases from both SDSS and Gaia XP spectra, highlighting the range of spectral tting outcomes. While the SDSS spectra bene t from

Figure 17: Example of a Gaia XP spectrum where the continuum could not be consistently tted across both H and H . Top: Cropped t around H and H . Bottom: Cropped t around H . Splitting the spectrum allows better continuum placement in each region and improves the reliability of line ux measurements.

Figure 18: Example of a challenging Gaia XP spectrum with strong Fe II contamination. The smoothed Fe II blends overlap with the Balmer wings, preventing robust decomposition of the broad-line pro les. The resulting uxes are agged as unreliable in the analysis.

higher resolution, wider wavelength coverage, and reliable host-galaxy subtraction, the Gaia XP spectra are limited by their lower resolution and smoothed line pro les. Nevertheless, they provide valuable broad-line ux measurements when carefully modeled. Several key di erences are:

- Resolution e ects: SDSS resolves narrow components, asymmetric [O III] wings, and Fe II blends, while in XP spectra these features are merged into the continuum or broad Balmer pro les. As a result, Gaia ts are restricted to broad hydrogen and, where visible, helium lines.
- Continuum placement: In SDSS, continuum anchors are well de ned by line-free spectral windows. In XP, the absence of such anchors and the smoothing of Fe II emission make continuum tting more uncertain, often requiring separate ts for H and H regions.
- Line ux recovery: For strong broad lines (H , H), Gaia uxes are in reasonable agreement with SDSS, though scatter increases for weaker or blended lines. Cases with strong Fe II contamination or weak broad components yield less reliable XP results and are agged accordingly in our tables.
- Fe II treatment: SDSS allows empirical template tting, whereas in XP spectra Fe II is heavily smoothed and cannot be robustly modeled, often absorbed into the broad Balmer continuum.
- Narrow-line dominance: A small fraction of spectra are dominated by strong narrow forbidden lines such as [O III] and [N II], with little or no detectable broad Balmer emission. These cases resemble Type 2 AGN spectra, where the BLR is obscured, and highlight situations where narrow-line features drive the decomposition.

To provide a quantitative assessment, integrated line uxes and luminosities from both data sets are compared in Section 4, alongside correlation plots for H and H . This synthesis demonstrates where Gaia XP spectra can serve as a practical substitute for SDSS in broad-line studies, where their limitations must be explicitly acknowledged.

3.3 Error Treatment and Uncertainty Estimates

The tting uncertainties in this work are based on the assessment of the outputs of the Sherpa optimization framework, as implemented in the FANTASY package. For each tted parameter (amplitude, centroid, and width), Sherpa reports a best- t value and its 1 error based on the covariance matrix. These statistical errors quantify the sensitivity of the t to noise in the observed spectrum and are therefore su cient to characterize the internal precision of the tting procedure.

Across our sample, the formal uncertainties are typically at the level of a few percent: broad-line amplitudes and integrated uxes have errors of 3 7%, and line widths are restricted to within 5%. To account for additional uncertainties not captured by the formal covariance errors (e.g. ux calibration and continuum placement) and based on the empirical experiments from previous studies and AGN monitoring campaigns (e.g., Shapovalova et al., 2012), we adopt a conservative error oor of 5% for SDSS spectra and 10% for Gaia XP spectra. These values re ect the higher calibration accuracy and resolution of SDSS compared to the lower-resolution smoothed XP spectra.

In problematic cases, most notably spectra strongly contaminated by Fe II blends or with poorly de ned continua the formal Sherpa errors may underestimate the true uncertainties. And derived uxes from such ts should be considered less reliable. In general, however, the broad-line uxes reported here are robust to within the global error estimates given above.

A complete treatment of the systematic error budget (for example, template choice for Fe II, degeneracies between line and continuum, and aperture e ects) is beyond the scope of this thesis. Instead, we focus on providing consistent statistical errors from the tting framework, complemented by conservative global uncertainties that account for calibration di erences between SDSS and Gaia.

Chapter 4

4 Physical Properties Measurements of the Sample

A central goal of this work is to extract reliable physical quantities from the spectra of our AGN sample, enabling a direct comparison between the high-resolution SDSS data and the lower-resolution Gaia XP spectra. From spectroscopic tting, we derive fundamental observables such as line uxes, continuum luminosities, and emission-line widths. These parameters provide insights into the structure of the broad-line region (BLR), ionization mechanisms, and scaling relations relevant for black hole mass estimation and accretion studies (Peterson, 2004; Shen et al., 2011).

In AGN research, measurement of emission line and continuum properties is central to probing the physics of the BLR and the accretion disk. The Balmer lines, particularly H and H , trace the kinematics of the BLR through their widths, which re ect the velocity distribution of the emitting gas. Continuum luminosities measured at 5100 Å and 6200 Å serve as proxies for the ionizing output of the accretion disk, while the Balmer decrement provides a diagnostic of reddening and gas conditions. Together, these properties underpin widely used scaling relations for black hole mass estimation, accretion rate inference, and intrinsic variability assessments across surveys (e.g., Vanden Berk et al., 2006; Shen et al., 2011).

In the following subsections, we outline the methodology for measuring these quantities, starting with emission line and continuum uxes, followed by FWHM and luminosity determinations. The results are summarized in tabular form and complemented by correlation analyses. Where appropriate, direct comparisons between SDSS and Gaia measurements are highlighted to evaluate the reliability of Gaia XP spectra for AGN studies.

4.1 Emission Line and Continuum Flux Measurements

A fundamental step in extracting AGN physical properties is the measurement of emission-line uxes and adjacent continuum levels. In this work, we focus on the Balmer lines H and H , which dominate the optical spectra of type 1 AGN and provide the primary tracers of the broadline region (BLR). These lines are central to diagnostics such as black hole mass scaling and reddening corrections via the Balmer decrement.

Line uxes were measured directly from the best- t spectral models obtained with FANTASY Each emission pro le was integrated using the scipy.integrate.simpson method. In cases where multiple Gaussian components were required to reproduce the broad line shape, their contributions were summed to obtain the total BLR ux. Narrow components, present in

SDSS but generally unresolved in Gaia, were excluded from the BLR ux estimates.

Continuum ux densities were extracted at rest-frame 5100 Å and 6200 Å from the tted models. The 5100 Å point is widely adopted as a standard reference in AGN scaling relations, while 6200 Å provides a redder continuum anchor less a ected by H and Fe, II contamination.

All uxes are reported in cgs units. For consistency, SDSS spectra were read with uxes internally scaled by 10^{17}, whereas Gaia XP spectra were multiplied by 10^{17} during preprocessing. This rescaling ensures that both data sets are analyzed in comparable 'natural' ux units, with the factor of 10^{17} restored only at the luminosity calculation stage.

4.1.1 Fluxes and FWHM

For each object, we measured the integrated uxes of H and H and the corresponding velocity widths (FWHM). When multiple broad Gaussian values were used, the summed pro le was reconstructed before integration and width measurement. The FWHM values provide the principal kinematic parameter of the BLR, complementing the ux measurements that quantify the line strength.

Together, these quantities form the observational basis for the subsequent derivation of line luminosities, Balmer decrements, and black hole mass estimates. The complete results of the uxes and the FWHM for all 64 objects are reported in Table 1, where SDSS and Gaia values are listed side by side to allow for direct comparison.

4.1.2 Balmer Decrement

The ratio of the broad ux of H to H , commonly called the Balmer decrement, is a widely used diagnostic in AGN studies. Under standard Case B recombination conditions at T 10^4, K and low density, the intrinsic H /H ratio is expected to be 2:86 3:1 (Osterbrock and Ferland, 2006). Deviations from this value can indicate reddening by dust, optical depth e ects in the broad-line region (BLR), or additional physical processes such as collisional excitation.

For each source, we computed the Balmer decrement directly from the integrated broad-line uxes of H and H obtained from the spectral ts (see Section 4.1.1). Narrow-line components, when present in SDSS spectra, were excluded to ensure that the measurement traces only BLR gas.

The Balmer decrement provides a rst-order probe of internal extinction in the BLR. Ratios signi cantly larger than the intrinsic value may suggest the presence of dust or complex radiative transfer e ects, whereas lower values are rare and typically indicate issues with line deblending or continuum placement.

Table 1: Observed line and continuum uxes (SDSS and Gaia) for the total sample of 64 Type 1 AGNs.

SDSS Name	RA	Dec	z	H Flux		H Flux		cont5100 Å		cont6200 Å		H /H		H fwhm		H fwhm	
				SDSS	Gaia	SDSS	Gaia	SDSS	Gaia	SDSS	Gaia	SDSS	Gaia	SDSS	Gaia	SDSS	Gaia
J0857+5041	134.13	50.69	0.23	20707.8	23028.7	6950.5	8255.2	85.3	72.6	61.6	53.7	2.98	2.79	2416	6879	2693	10873
J0830+5131	127.49	51.51	0.34	21760.0	29060.9	6887.8	5673.3	91.8	80.8	65.1	48.4	3.16	5.12	7593	8630	7524	5401
J1019+6358	154.80	63.97	0.04	17183.5	21756.6	5111.1	7904.5	79.8	140.8	80.7	128.7	3.36	2.75	5177	5260	5246	9405
J1509+0902	227.29	9.04	0.04	3.0	7605.7	706.0	15844.6	5.3	67.4	4.1	73.7	0.00	0.48	1174	5244	1242	8103
J1356+2046	208.96	20.77	0.20	43638.5	53853.5	11985.6	13572.4	155.6	160.6	122.2	115.1	3.64	3.97	2554	6411	2692	10890
J1354+2326	208.53	23.43	0.06	22206.7	17647.8	3347.7	7069.4	107.1	163.9	100.0	161.6	6.63	2.50	5524	6065	2720	10291
J1451+2709	222.79	27.16	0.06	52468.7	60320.4	16055.7	28930.1	385.1	405.5	296.1	308.6	3.27	2.09	759	5061	1709	9606
J1701+2919	255.19	29.32	0.07	1768.1	4589.8	0.0	2184.2	17.3	45.6	15.9	61.4	nan	2.10	5591	7687	nan	6125
J1703+3737	255.83	37.62	0.06	13629.9	6290.8	3723.3	2706.0	30.6	40.0	31.8	48.9	3.66	2.32	4832	5996	2600	5928
J1619+4425	244.71	44.42	0.34	13804.4	17824.0	4221.5	5216.9	75.5	59.7	52.8	46.5	3.27	3.42	3659	7994	4350	7887
J1610+4902	242.47	49.03	0.04	6900.31	5297.28	990.85	1827.5	32.91	86.14	20.4	89.73	6.96	2.9	3451	7033	3451	9369
J1546+4846	236.38	48.77	0.40	16403.01	26266.88	5648.09	7284.35	69.08	88.96	49.07	70.35	2.9	3.61	2485	8272	3037	7453
J1720+5540	260.11	55.67	0.36	10349.29	11911.06	3280.68	3030.83	28.39	28.74	21.19	20.98	3.15	3.93	2623	8045	2831	5698
J1341+4037	205.16	40.61	0.16	17570.69	20818.18	5632.78	5758.43	56.89	64.89	44.79	47.42	3.12	3.62	2347	6213	2416	7861
J1250+3331	192.40	33.51	0.16	4470.77	5401.05	320.43	591.65	56.38	62.04	45.57	43.94	13.95	9.13	276	11882	207	7945
J1242+3317	190.54	33.28	0.04	33641.5	54138.64	6571.89	14203.71	176.08	236.05	188.2	249.75	5.12	3.81	1587	5247	1726	10895
J1221+3853	185.15	38.89	0.38	15822.47	19286.1	5120.17	4827.25	64.23	51.48	47.45	42.77	3.09	4.0	1933	8158	2071	4746
J1216+5442	183.96	54.71	0.15	6092.96	13583.3	1848.5	3734.81	75.91	73.4	60.08	55.64	3.3	3.64	966	5799	1104	9983
J1233+6625	188.16	66.41	0.05	22527.32	49334.38	3546.22	16845.72	21.05	154.82	24.46	146.44	6.35	2.93	2761	5668	2692	9856
J1230+6622	187.59	66.37	0.18	23950.48	27103.28	7737.31	9280.95	70.09	73.05	50.47	52.53	3.1	2.92	2761	6941	3452	9894
J0156-0857	28.88	-8.95	0.16	21316.39	29529.97	6579.01	9783.66	68.46	70.74	50.31	56.69	3.24	3.02	6075	7755	6903	10978
J0217-0325	34.24	-3.42	0.14	12952.74	17188.1	5685.59	9105.46	57.24	49.73	42.79	39.76	2.28	1.89	3245	6662	3383	9254
J0239-0001	39.85	-0.02	0.26	54512.98	60270.47	21092.06	16995.96	230.36	198.63	160.72	147.72	2.58	3.55	3382	7234	3521	9912
J0138-0109	24.54	-1.16	0.27	12927.02	17713.91	4928.16	3511.35	62.52	58	43.11	44.84	2.62	5.04	2209	7242	2278	9926
J0140-0050	25.07	-0.83	0.33	39769.09	46494.41	12104.48	12359.46	112.9	86.42	81.94	63.84	3.29	3.76	4833	8402	5457	7921
J0140-0050	25.07	-0.83	0.33	39799.10	46498.12	10753.36	12401.61	97.15	86.42	72.04	63.84	3.70	3.75	4765	8402	5595	7921
J0127+0327	21.70	3.45	0.22	19299.80	21583.39	6094.66	5048.77	63.35	62.25	50.06	45.98	3.17	4.27	3796	7199	4073	9913
J0140+0619	24.98	6.32	0.40	3281.37	9191.69	1564.15	1717.68	54.55	61.85	35.53	38.06	2.10	5.35	4901	9755	4970	3969
J0140+0619	24.98	6.32	0.40	3387.60	9191.69	1666.19	1717.68	58.80	61.85	38.07	38.06	2.03	5.35	4970	9755	4970	3969
J2340+0010	355.04	0.17	0.06	4629.23	6229.77	1445.98	3705.30	22.36	44.43	23.85	68.65	3.20	1.68	3037	4307	3244	8041
J2302-0107	345.48	-1.11	0.24	13168.03	18784.56	4149.96	3947.49	44.11	55.13	34.16	39.89	3.17	4.76	4006	7008	4491	9904
J2220+0110	335.10	1.16	0.21	28098.80	24046.39	7962.21	5159.08	99.08	85.43	83.98	68.83	3.53	4.66	4345	7073	5588	9943
J2141+0026	325.23	0.43	0.08	17967.23	24788.53	5933.61	10376.06	108.80	103.87	93.26	95.74	3.03	2.39	1864	5814	2071	10776
J0002+0529	0.39	5.49	0.25	10436.21	12921.37	3580.74	3936.05	40.56	40.67	31.67	29.72	2.91	3.28	2140	6863	2209	9915
J2315+1828	348.82	18.47	0.10	10540.98	14680.76	4783.34	6536.32	107.73	87.74	84.33	69.47	2.20	2.25	2692	6292	2831	10914
J2328+2605	352.01	26.08	0.33	12678.08	16966.85	4046.70	4132.69	36.91	41.38	28.53	31.93	3.13	4.11	2140	7272	2209	5088
J0141+2345	25.15	23.75	0.32	31696.06	37134.26	9545.16	8134.26	88.70	69.53	68.69	53.18	3.32	4.57	2554	7998	2692	5948
J0806+0735	121.44	7.59	0.05	12515.43	17185.66	2688.66	6026.66	43.55	76.88	46.80	125.94	4.65	2.85	4142	5199	4764	8103
J0310-0050	47.62	-0.83	0.08	18764.04	31061.74	8911.02	10352.13	141.61	116.33	101.25	108.82	2.11	3.00	2345	5666	3310	8947
J0322+0055	50.56	0.92	0.18	26346.75	40660.93	11345.84	11336.55	106.53	91.58	82.53	79.41	2.32	3.59	2278	6628	2416	9938
J1317-0216	199.26	-2.26	0.02	7126.11	9500.83	1157.14	3836.94	7.13	58.69	11.95	32.46	6.16	2.48	1104	4475	1724	7253
J1105+2124	166.15	21.40	0.19	13711.25	16379.00	4255.32	3144.93	46.54	50.61	38.67	36.81	3.22	5.21	4210	7526	5245	9921
J1053+2406a	163.16	24.11	0.40	8984.77	24092.21	5683.06	5626.30	82.94	60.73	60.87	45.85	1.58	4.28	3636	9369	7524	7876
J1053+2406b	163.16	24.11	0.40	20406.99	24092.21	5739.35	5626.30	79.92	60.73	55.45	45.85	3.56	4.28	5455	9369	6078	7876
J1119+2119	169.79	21.32	0.18	79634.12	160319.56	51495.01	52032.24	451.54	387.10	314.92	240.79	1.55	3.08	3728	6990	3244	10957
J1131+3114	172.79	31.23	0.29	35525.93	38356.63	20.96	8077.87	110.15	82.59	78.57	57.47	nan	4.75	3175	7357	4696	10928
J1609+1220	242.15	12.33	0.03	40279.03	54089.30	13118.71	20999.81	205.21	220.11	189.09	162.27	3.07	2.58	3452	4422	3936	8328
J0855+1741	133.66	17.69	0.07	18641.69	21860.31	4930.77	15453.73	148.10	130.55	137.47	127.17	3.78	1.41	2140	4283	2692	11356
J0926+1954	141.48	19.90	0.19	82279.12	77412.58	20982.47	19430.30	266.93	267.75	176.76	169.44	3.92	3.98	7661	8725	7386	9856
J0927+2301	141.83	23.02	0.03	16157.91	11179.75	3754.47	3974.86	21.82	106.60	33.53	70.14	4.30	2.81	6284	7117	6492	9480
J0805+2121	121.22	21.35	0.12	26727.87	23371.76	7975.49	8347.56	91.91	50.89	74.13	43.73	3.35	2.80	6559	7379	6627	9892
J0836+2500	128.90	24.99	0.33	35537.71	44785.21	9967.23	9964.83	111.80	115.89	79.23	78.08	3.57	4.49	2623	7750	2969	5915
J1004+2856	151.01	28.93	0.33	24547.14	38971.64	9853.25	10578.47	111.25	149.67	82.89	106.08	2.49	3.68	3797	7828	3658	5927
J1017+2915	154.33	29.24	0.05	7538.75	13443.11	3325.30	5274.50	50.26	72.35	33.78	78.03	2.27	2.55	2140	4345	2140	8166
J1006+4333	151.42	43.54	0.18	8889.28	12996.19	2958.53	3206.03	66.91	73.87	47.41	49.97	3.00	4.05	2485	6585	2347	7602
J0949+4335	147.25	43.59	0.23	16657.19	17793.66	5296.84	3325.88	86.45	81.81	69.28	61.92	3.14	5.35	3175	6922	3382	9860
J1000+5045	149.88	50.75	0.14	13172.45	16043.59	5000.87	4447.74	52.64	48.19	38.24	34.78	2.63	3.61	2072	6071	2140	8919
J1139+5912	174.79	59.20	0.06	55706.91	68133.70	19172.65	32532.39	151.59	210.06	120.22	149.22	2.91	2.09	3383	5765	3590	9816
J1106+5851	166.41	58.86	0.19	16541.79	24662.50	5397.75	5050.23	73.89	59.49	57.66	49.42	3.06	4.88	5246	7278	5453	9921
J0750+3455	117.45	34.91	0.13	11752.39	18428.63	5370.16	5452.38	40.81	70.34	30.46	47.84	2.19	3.38	3934	6052	3728	9870
J0741+3547	115.36	35.78	0.13	19125.90	24366.90	5605.55	8440.67	55.13	47.73	45.27	43.88	3.41	2.89	5867	5626	6213	8627
J0736+3926	114.10	39.44	0.12	46722.63	67645.84	12716.71	24621.44	115.77	135.81	96.92	116.76	3.67	2.75	2140	5798	2416	10965
J0749+4511	117.28	45.18	0.19	36854.68	38188.59	11814.15	9617.30	102.37	98.85	82.09	69.15	3.12	3.97	4349	7043	4280	9880
J0749+4511	117.28	45.18	0.19	24376.38	38023.54	9780.00	9584.88	84.19	98.85	58.27	69.15	2.49	3.97	5315	704	5316	9880

Note: Comparison of ux and line width measurements between SDSS and Gaia XP spectra. RA and Dec are in degrees, z is redshift. Fluxes are in units of erg s ^{1} cm ^{2} and FWHM values are in km s ^{1}. In this table, the Gaia ID was not included due to the lack of space, and it is provided later in Table 2.

In our sample, the decrements were systematically compared between SDSS and Gaia XP spectra. This allows us to assess both the physical interpretation of BLR reddening and the reliability of Balmer line ux ratios when derived from lower-resolution Gaia data. A summary of the Balmer decrements for all objects is included in Table 1, together with the ux and FWHM measurements.

4.2 Luminosities Measurement

A direct comparison of uxes between di erent AGN is not physically meaningful, since the observed ux depends strongly on the source redshift and luminosity distance. Two objects with identical intrinsic emission can appear with very di erent uxes simply because of their distance. To place all sources on an equal footing, the measured uxes must be converted into luminosities, which represent the intrinsic power output of the source. This step is therefore essential for any physical interpretation and for comparing SDSS and Gaia measurements across the full redshift range of our sample.

Luminosities were derived from the measured uxes using luminosity distances calculated with the astropy.cosmology package (Astropy Collaboration, 2013, 2018), adopting the (Planck Collaboration, 2020) cosmology ($H_0 = 67.4$ km s^{-1} Mpc^{-1}, $_M = 0.315$, $= 0.685$). These rest frame line and continuum luminosities are used consistently for both SDSS and Gaia.

For integrated emission lines, the luminosity was obtained as

$$L_{line} = 4 \; D_L^2 \, F_{line}; \tag{3}$$

where F_{line} is the total line ux measured from the best t model and D_L the luminosity distance at the source redshift.

For continuum luminosity densities at rest frame 5100 Å and 6200 Å one may write,

$$L \; (\; _{rest}) = 4 \; D_L^2 \, \frac{F \; (\; _{obs})}{1 + z}; \qquad _{obs} = (1 + \; z) \; _{rest}. \tag{4}$$

with $_{obs} = (1 + \; z) \; _{rest}$. The factor $(1 + z)$ accounts for bandwidth compression by converting the observed ux densities to the rest-frame values. In our analysis, the spectra are shifted to the rest frame during preprocessing by the FANTASY pipeline. As a result, both the tted continuum ux densities F (evaluated directly at $_{rest} = 5100$ and 6200 Å) and the integrated broad-line uxes (H , H) are already expressed in the rest frame.

Therefore, in our case, luminosities are computed using the following relation without any additional redshift correction,

$$L = 4 \; D_L^2 \, F \tag{5}$$

Table 2: Observed line and continuum luminosties (SDSS and Gaia) for the total sources of 60 Type 1 AGNs.

SDSS Name	Gaia DR3 ID	z	L_H		L_H		L (5100Å)		L (6200Å)	
			SDSS	Gaia	SDSS	Gaia	SDSS	Gaia	SDSS	Gaia
J0857+5041	1016834950437655552	0.23	43.56	43.61	43.09	43.16	36.88	36.81	36.83	36.77
J0830+5131	1028155389013411584	0.34	43.96	44.08	43.46	43.37	37.29	37.24	37.23	37.10
J1019+6358	1053351556957717376	0.04	41.87	41.97	41.34	41.53	35.24	35.49	35.33	35.54
J1509+0902	1166949486271776640	0.04	88.16	41.57	40.54	41.89	34.12	35.23	34.09	35.35
J1356+2046	1249947878136029824	0.20	43.71	43.80	43.15	43.20	36.97	36.99	36.95	36.93
J1354+2326	1251676310414862464	0.06	42.24	42.14	41.42	41.74	35.63	35.81	35.68	35.89
J1451+2709	1269203419114884608	0.06	42.75	42.81	42.24	42.49	36.32	36.35	36.29	36.31
J1701+2919	1309468737516333440	0.07	41.33	41.74		41.42	35.03	35.45	35.08	35.66
J1703+3737	1339705032401828224	0.06	42.17	41.83	41.60	41.46	35.23	35.34	35.33	35.51
J1619+4425	1385484951370252160	0.34	43.74	43.85	43.22	43.32	37.18	37.08	37.11	37.06
J1610+4902	1399448336726161152	0.04	41.44	41.32	40.60	40.86	34.83	35.24	34.70	35.35
J1546+4846	1401637979773304960	0.40	43.99	44.20	43.53	43.64	37.32	37.43	37.26	37.42
J1720+5540	1420458114147368064	0.36	43.68	43.75	43.19	43.15	36.83	36.84	36.79	36.78
J1341+4037	1500142058274857984	0.16	43.13	43.20	42.63	42.64	36.34	36.40	36.32	36.35
J1250+3331	1514279097747811456	0.16	42.54	42.62	41.40	41.66	36.35	36.39	36.34	36.33
J1242+3317	1514610016387783296	0.04	42.20	42.41	41.49	41.83	35.63	35.76	35.74	35.87
J1221+3853	1532557997522989184	0.38	43.91	44.00	43.42	43.40	37.23	37.13	37.18	37.14
J1216+5442	1572729067161813760	0.15	42.60	42.95	42.08	42.39	36.40	36.39	36.38	36.35
J1233+6625	1680733166644357760	0.05	42.09	42.43	41.29	41.97	34.77	35.64	34.92	35.70
J1230+6622	1680740828866041088	0.18	43.39	43.44	42.90	42.98	36.56	36.58	36.50	36.52
J0156-0857	2462859664867002112	0.16	43.23	43.37	42.72	42.89	36.44	36.46	36.40	36.45
J0217-0325	2492983083717300352	0.14	42.84	42.97	42.48	42.69	36.20	36.13	36.15	36.12
J0239-0001	2498737553915316480	0.26	44.09	44.13	43.68	43.58	37.42	37.36	37.35	37.31
J0138-0109	2508869686708806144	0.27	43.51	43.64	43.09	42.94	36.90	36.86	36.82	36.84
J0140-0050	2508952772851444736	0.32	44.20	44.26	43.68	43.69	37.36	37.24	37.30	37.19
J0127+0327	2560047421790955264	0.22	43.45	43.50	42.95	42.87	36.68	36.67	36.66	36.62
J0140+0619	2562478059747868160	0.40	43.28	43.73	42.96	43.00	37.21	37.27	37.11	37.14
J2340+0010	2642879508227304704	0.06	41.63	41.76	41.12	41.53	35.02	35.32	35.13	35.59
J2302-0107	2650710023881781248	0.24	43.38	43.53	42.88	42.85	36.61	36.71	36.58	36.65
J2220+0110	2679303506661773696	0.21	43.60	43.53	43.05	42.86	36.85	36.79	36.87	36.78
J2141+0026	2687420896067111040	0.08	42.52	42.66	42.04	42.28	36.01	35.99	36.03	36.04
J0002+0529	2745007370896353664	0.25	43.33	43.42	42.86	42.90	36.62	36.62	36.60	36.57
J2315+1828	2819126102559044096	0.10	42.49	42.63	42.15	42.28	36.21	36.12	36.19	36.10
J2328+2605	2841448525065809536	0.33	43.68	43.81	43.19	43.20	36.86	36.91	36.83	36.88
J0141+2345	290930265608514816	0.32	44.06	44.13	43.54	43.47	37.22	37.11	37.19	37.08
J0806+0735	3097763789516063104	0.05	41.96	42.09	41.29	41.64	35.20	35.45	35.32	35.75
J0310-0050	3265557372810761344	0.08	42.50	42.72	42.18	42.24	36.09	36.00	36.03	36.06
J0322+0055	3267488664984941184	0.18	43.43	43.62	43.07	43.07	36.75	36.68	36.72	36.70
J1317-0216	3684714604300891904	0.02	40.81	40.93	40.02	40.54	33.52	34.43	33.82	34.26
J1105+2124	3985601493159227392	0.19	43.16	43.24	42.66	42.52	36.40	36.44	36.41	36.39
J1053+2406	3990157457027790848	0.40	43.72	44.15	43.53	43.52	37.40	37.26	37.35	37.22
J1119+2119	3990830907900291584	0.18	43.86	44.17	43.68	43.68	37.33	37.26	37.25	37.14
J1131+3114	4023588108070402816	0.29	44.00	44.04	40.77	43.36	37.20	37.08	37.14	37.01
J1609+1220	4458012573911009664	0.03	42.05	42.17	41.56	41.76	35.46	35.49	35.51	35.44
J0855+1741	611893888898143488	0.07	42.31	42.38	41.73	42.23	35.92	35.86	35.97	35.94
J0926+1954	634285997249170176	0.19	43.96	43.94	43.37	43.34	37.18	37.18	37.09	37.07
J0927+2301	638374531236818048	0.03	41.44	41.28	40.81	40.83	34.27	34.97	34.55	34.87
J0805+2121	670950827385990912	0.12	43.06	43.01	42.54	42.56	36.31	36.05	36.30	36.07
J0836+2500	678613976755423488	0.12	44.13	44.23	43.58	43.58	37.34	37.35	37.27	37.27
J1004+2856	740685668749400064	0.33	43.97	44.17	43.57	43.60	37.33	37.46	37.29	37.40
J1017+2915	741343039264023552	0.33	41.66	41.91	41.31	41.51	35.19	35.35	35.11	35.47
J1006+4333	808775434561821440	0.05	42.93	43.09	42.45	42.48	36.51	36.55	36.45	36.47
J0949+4335	819907023423890304	0.18	43.43	43.46	42.93	42.73	36.85	36.83	36.84	36.79
J1000+5045	827078008196318592	0.23	42.89	42.98	42.47	42.42	36.20	36.16	36.15	36.10
J1139+5912	858200268935710208	0.14	42.73	42.81	42.26	42.49	35.87	36.01	35.85	35.95
J1106+5851	860383349272777344	0.06	43.26	43.44	42.78	42.75	36.62	36.53	36.60	36.53
J0750+3455	882564415815249280	0.19	42.76	42.96	42.42	42.43	36.01	36.25	35.97	36.17
J0741+3547	895654342263361152	0.13	42.98	43.09	42.45	42.63	36.15	36.09	36.15	36.14
J0741+3547	895654342263361152	0.13	43.26	43.42	42.69	42.98	36.36	36.43	36.37	36.45
J0749+4511	926435994915623552	0.12	43.62	43.63	43.12	43.03	36.77	36.75	36.76	36.68

Note: All luminosities are reported as $\log_{10}$ (erg s ^{1}). Line luminosities L_H and L_H are derived from broad-line uxes; continuum luminosities L are evaluated at 5100 Å and 6200 Å. Distances are from the astropy Planck18 cosmology. Duplicate SDSS spectra of the same source are omitted, as their luminosities are e ectively identical in log scale.

4.3 Correlations of Spectral Parameters

Correlation plots provide a straightforward way to test the consistency of spectral measurements obtained from di erent instruments. In this work, they are used to examine whether properties derived from Gaia XP spectra follow the same trends as those of SDSS, which serves as the established reference. This comparison is essential to evaluate the reliability of Gaia-based measurements of broad-line and continuum features and to determine whether they can be used in place of, or complementary to, higher-resolution spectroscopy.

Beyond serving as a validation, correlation analysis also highlights potential limitations. Di erences or increased scatter may point to the e ects of spectral resolution, calibration procedures, or intrinsic source variability arising from non-simultaneous observations. By systematically comparing luminosities, ux ratios, and continuum measures, this step establishes a baseline to assess the robustness of Gaia XP in AGN studies. The primary aim here is not interpretation of individual results, but to provide a methodological foundation for the detailed plot-by-plot discussion that follows.

4.3.1 Line Luminosity Correlations (H , H)

The correlation plots of the luminosities of H and H (Figure 19 and Figure 20) show a general strong consistency within SDSS and within Gaia, as well as between the two surveys. Most objects lie closely along the one-to-one relation, indicating that the broad-line ux measurements are robust across instruments. However, a small number of clear outliers stand out, which can be traced back to issues in the spectral tting. There are some issues with 2 sources that are particularly noticed.

Figure 19: Broad-line H vs H luminosity correlations within SDSS (left) and within Gaia XP (right).

- SDSS J1509+0902 In the SDSS spectra, this source does not exhibit an obvious broad Balmer emission, consistent with a type 2 AGN classi cation. The automated

tting procedure appears to have attempted to model narrow [O III] or other oxygen lines as broad H emission, leading to arti cially suppressed H luminosities. This explains its strong deviation in the SDSS H H and H continuum correlations.

- SDSS J1131+3114 The spectrum around the H region shows an almost at continuum with no clear broad emission feature. The tting procedure attempts to place a broad component where none is convincingly present, leading to unreliable ux estimates. This propagates into discrepant H luminosities, making the source appear as an outlier in the correlation plots.

Figure 20: Comparison of broad-line luminosities between SDSS and Gaia XP: H (left) and H (right).

In contrast, the Gaia XP spectra do not show extreme outliers of this magnitude. Although some points deviate slightly from the one-to-one line, they remain in the vicinity of the correlation and do not dominate the overall trend. This suggests that the major discrepancies are primarily driven by anomalous SDSS ts rather than intrinsic di erences between the two datasets.

These cases underscore the importance of visual inspection alongside automated tting: individual spectra with poor line pro les can bias correlations if not agged, while the bulk of the sample shows consistent line luminosities across instruments.

Balmer decrement Correlations: Figure 21 compares the broad-line Balmer decrement $F_H = F_H$ from SDSS and Gaia XP. The extreme SDSS outlier in a very large ratio is due to the object SDSS J1131+3114 its spectrum shows essentially no measurable broad H component, so the tted F_H is driven toward zero and the ratio becomes arbitrarily large. Once this object (and the tiny-ratio mirror case) is excluded, most objects cluster around consistent values.

The comparison shows a slight tilt between SDSS and Gaia, with SDSS-based ratios tending to appear higher. This arises mainly in cases where the broad H line is extremely

weak: in SDSS, the deblending around H can push the tted ux close to zero, in ating the F_H =F_H ratio. In Gaia, the lower spectral resolution smooths over such weak features, which suppresses the most extreme values, although moderate scatter remains. The clearest example is SDSS J1131+3114, where the tted H contributes negligibly, producing an outlier that is far above the main trend. For clarity, we exclude ratios larger than 10, since values above this threshold are non-physical for AGN and are typically associated with poor line ts rather than intrinsic properties.

Figure 21: Comparison of broad-line Balmer decrement F_H =F_H between SDSS and Gaia XP. Grey line marks the one-to-one relation, while the blue line shows the linear t.

4.3.2 Line Widths (FWHM)

Figures 22 and 23 compare the FWHM of the broad H and H components measured from SDSS and Gaia. The SDSS results show the expected consistency: the FWHM of H and H are strongly correlated, apart from the two outlier sources already discussed in Section 4.3.1. This re ects the fact that both lines originate in the same broad-line region and are tted with a precision comparable to that of the higher-resolution SDSS spectra.

In contrast, the Gaia XP measurements appear considerably more scattered. Two factors contribute to this behavior. First, in our tting procedure the FWHM of H and H is not tied to be equal, so the di erences between the two are exaggerated when the lines are weak or blended. Second, the lower spectral resolution of XP tends to broaden line pro les, systematically shifting the Gaia-based FWHM toward larger values compared to SDSS. This e ect is particularly pronounced for H , where the weaker line strength and stronger blending with neighboring features (e.g. [O iii], Fe ii) make the continuum level more di cult to determine. As a result, the scatter in Gaia FWHM(H) is substantially larger than in FWHM(H).

Figure 22: Comparison of FWHM measurements between SDSS and Gaia XP: H (left) and H (right). The dashed line indicates the one-to-one relation.

Figure 23: Internal comparison of FWHM(H) and FWHM(H) within SDSS (left) and within Gaia XP (right). The dashed line indicates the one-to-one relation.

4.3.3 Continuum Luminosities at 5100 Å and 6200 Å

Figures 24 and 25 show the comparison of continuum measurements from SDSS and Gaia XP at the two standard reference wavelengths, 5100 Å and 6200 Å. We distinguish between two related quantities: the continuum luminosity density, L (erg s ^{1} Å ^{1}), and the monochromatic luminosity, L (erg s ^{1}). These provide complementary information: the luminosity density is a direct measure of the local continuum ux level, while the monochromatic luminosity is more commonly used in scaling relations (e.g., black hole mass estimates) as it incorporates the wavelength factor.

The results for the continuum luminosity densities at both 5100 Å and 6200 Å show a very tight correlation between SDSS and Gaia, with scatter remaining modest across the full luminosity range. This indicates that despite the lower spectral resolution of XP, the broad continuum level is recovered with high delity. The agreement also suggests that any residual calibration o sets between the two surveys are small compared to the intrinsic measurement uncertainties.

Figure 24: Continuum luminosity comparison at 5100 Å and 6200 Å within each survey. Left: SDSS; right: Gaia XP. The grey line shows the one-to-one relation; the blue line is the best-fit regression.

For monochromatic luminosities, the correlations remain strong, but a small systematic offset is visible: SDSS values tend to be slightly higher on average than those from Gaia. This offset may arise from subtle differences in how the continuum is placed when neighboring features are present, for example, Fe II emission blends near 5100 Å or molecular absorption bands at longer wavelengths. Since SDSS spectra offer higher resolution and broader wavelength coverage, they may isolate the continuum windows more precisely, while the smoothing in XP spectra can lead to a mild underestimation.

Figure 25: Cross-survey comparison of continuum luminosities. Left: SDSS vs Gaia XP at 5100 Å; right: SDSS vs Gaia XP at 6200 Å. The grey line shows the one-to-one relation; the blue line is the best-fit regression.

Importantly, no clear outliers are seen in these comparisons. All objects lie close to the one-to-one relation, and the scatter is fairly uniform, without evidence of redshift-dependent biases. This gives confidence that the continuum measurements derived from Gaia XP can be used in the same way as SDSS for applications relying on monochromatic luminosities, though one should be mindful of the small systematic tilt. In Figures 24 and 25, the blue line shows the best-fit relation, while the gray line marks the reference of one-to-one.

4.3.4 Line Continuum Correlations

A further test of consistency is provided by comparing the luminosities of broad Balmer lines with the adjacent continuum windows that are commonly used as reference anchors in AGN scaling relations. Figures 26 and 27 show the relations between H and the 5100 Å continuum, and between H and the 6200 Å continuum, for both SDSS and Gaia. The dashed grey lines mark the one-to-one relation, while the blue dashed lines show the linear fits.

Figure 26: Correlation between H line luminosity and 5100 Å continuum luminosity. Left: SDSS; right: Gaia XP. The grey line shows the one-to-one relation, and the blue line the linear fit.

In the SDSS measurements, both H 6200 Å and H 5100 Å correlations are tight, with correlation coefficients above $r = 0.9$. This reflects the expected physical link between broad-line strength and the underlying ionizing continuum, as well as the higher spectral resolution and S/N that enable robust separation of continuum and line components. Outliers correspond to the same sources flagged earlier (e.g. SDSS J1509+0902 with weak or absent broad emission), where the fitted line luminosities are underestimated relative to the continuum.

Figure 27: Correlation between H line luminosity and 6200 Å continuum luminosity. Left: SDSS; right: Gaia XP. The grey line shows the one-to-one relation, and the blue line the linear fit.

The Gaia XP correlations preserve the same overall trends, with correlation coefficients similarly high (r 0.95 0.98). However, the scatter is modestly larger than in SDSS, espe-

cially for H , consistent with the di culty of constraining the continuum level near 5100 Å at XP's lower resolution. In particular, weak-line objects tend to cluster slightly below the one-to-one line, indicating that the tted line luminosity is suppressed relative to the continuum. By contrast, the H 6200 Å relation remains very tight even in XP, re ecting the greater robustness of H ts in lower-resolution data.

Overall, these comparisons demonstrate that Gaia XP reproduces the expected line continuum relations with high delity, although the scatter increases for H due to blending and continuum placement uncertainties. The results suggest that XP-based luminosities can be reliably used for ensemble studies of line continuum scaling, while caution is required for individual weak-line sources.

4.4 Black Hole Mass Estimation

The mass of the central black hole is one of the most fundamental quantities for understanding active galactic nuclei. It sets the characteristic scale of the broad-line region, regulates the accretion rate, and provides a direct connection between nuclear activity and host-galaxy evolution. Traditionally, the most reliable determinations of M_{BH} come from reverberation mapping, but such campaigns are observationally expensive and limited to small samples. In large surveys, single-epoch virial estimators calibrated against reverberation-mapped AGN are therefore the only feasible approach. Testing these estimators across di erent instruments and resolutions is essential for assessing their reliability. In particular, comparing M_{BH} derived from SDSS and Gaia spectra provides a benchmark for evaluating whether the low-resolution XP data can recover meaningful black hole mass estimates.

We estimated the black hole masses in our sample using the single-epoch virial mass estimator based on the broad H line, as calibrated by (Greene and Ho, 2005). This method builds on the empirical radius luminosity relation established through reverberation mapping (Bentz et al., 2009), and is among the most widely applied prescriptions for deriving M_{BH} from single-epoch spectra (see review by Popovi¢, 2012). The adopted relation (Equation 6 of Greene and Ho 2005) is,

$$M_{BH} = 2.0^{+0.4}_{-0.3} \quad 10^{6} \quad \frac{L_{H}}{10^{42}\,\mathrm{erg\,s^{-1}}}^{0.55 \quad 0.02} \quad \frac{FWHM_{H}}{10^{3}\,\mathrm{km\,s^{-1}}}^{2.06 \quad 0.06} \quad M \quad . \qquad (6)$$

Here, L_H denotes the broad H luminosity and $FWHM_H$ is its full width at half maximum. The quoted uncertainties in the exponents and normalization represent the calibration errors of the scaling relation, while additional scatter arises from the assumed BLR geometry and orientation. We adopted H as the reference line since it is the strongest broad Balmer feature, less a ected by host-galaxy contamination than H , and reliably measurable in both SDSS and Gaia XP spectra.

Figure 28: Comparison of black hole mass estimates derived from the broad H line using SDSS (x-axis) and Gaia XP (y-axis) spectra. The dashed line indicates the one-to-one relation, while the shaded region marks 0:4 dex, corresponding to the typical scatter of single-epoch virial mass estimators. Points are color-coded by redshift, with error bars showing the propagated calibration uncertainties. The median o set between the two estimates is ~ ' +0:76 dex, with a Pearson correlation coe cient of r 0:71.

Uncertainty treatment. For each source we computed M_{BH} and propagated the calibration uncertainties of Eq. (6) into logarithmic space to obtain errors on $\log M_{BH}$. Speci cally, the exponent errors (0:02 on luminosity and 0:06 on line width) and the uncertainty in the normalization were combined in quadrature, yielding symmetric error bars that are shown in Figure 28. The shaded 0:4 dex region represents the typical intrinsic scatter of single-epoch virial mass estimators due to uncertainties in BLR geometry and inclination.

A close correspondence between SDSS- and Gaia-based mass estimates would demonstrate the potential of XP spectra for reliable black hole mass studies in large statistical samples, whereas any systematic deviations highlight the limitations imposed by low spectral resolution. The resulting comparison is presented in Figure 28.

The comparison shows a clear correlation with a median o set of ~ ' +0:76 dex and Pearson r ' 0:71. This systematic o set is most likely due to the overestimated FWHM of H line (see Figure 22), which is clearly due to limited spectral resolution of Gaia XP spectra.

Nevertheless, this analysis demonstrates that Gaia XP spectra, while low-resolution, can still yield meaningful SMBH mass estimates.

4.5 Eddington Ratio

The Eddington ratio, de ned as the ratio of bolometric to Eddington luminosity, is a fundamental parameter describing the accretion state of AGN. It provides insight into how e ciently matter is accreted relative to the theoretical maximum set by the balance between radiation pressure and gravitational attraction (e.g. Peterson, 1993; Netzer, 2015). The Eddington ratio is expressed as

$$_{\text{Edd}} = \frac{L_{\text{bol}}}{L_{\text{Edd}}}; \tag{7}$$

where L_{bol} is the bolometric luminosity and L_{Edd} is the Eddington luminosity, given by

$$L_{\text{Edd}} = 1:26 \quad 10^{38} \quad \frac{M_{\text{BH}}}{M} \quad \text{erg s}^{\ 1}: \tag{8}$$

Here, M_{BH} is the black hole mass obtained from the single-epoch virial estimator (Sect. 4.4). For the bolometric luminosity, we adopt the prescription $L_{\text{bol}} = k L_{5100}$ with a bolometric correction factor of $k = 9$ (Kaspi et al., 2000; Vestergaard and Peterson, 2006), where L_{5100} denotes the monochromatic continuum luminosity at 5100 Å. This approach allows a consistent comparison of Eddington ratios between SDSS and Gaia XP spectra.

In practice, we computed $_{\text{Edd}}$ separately for each dataset (SDSS and Gaia), using the black hole mass estimates (Section 4.4) and the bolometric luminosity derived from the 5100 Å continuum. To avoid unphysical values and extreme leverage in the plots, we applied a conservative threshold of $_{\text{Edd}}$ 3, which removes only one object SDSS J1250+3331 with a spuriously large value. The comparison between the Eddington ratios derived from SDSS and Gaia XP is presented in Figure 29, where we show both the one-to-one relation and the scatter in $\log(L=L_{\text{Edd}})$.

The results indicate a systematic o set of ~ ' 0:56 dex, with Gaia-based values typically lower than those from SDSS, which is the e ect of systematically larger SMBH masses obtained from Gaia spectra. Noticeably, the Eddington ratio discrepancy increases for larger Eddington ratios, i.e., for less luminous AGNs, pointing to the possibly underestimate of the bolometric luminosity from Gaia spectra. This could be due to challenges in measuring the continuum ux for low luminosity AGNs, as shown in Figure 24. Nevertheless, a moderate positive correlation (r 0:40) suggests that, despite its lower resolution, Gaia XP spectra can recover accretion-related trends in a statistical sense. The observed o set is likely attributable to the e ect of spectral resolution in the XP data.

Figure 29: Comparison of Eddington ratios $L=L_{Edd}$ derived from SDSS (x-axis) and Gaia XP (y-axis) spectra. The dashed line indicates the one-to-one relation, while the o set represents the median $\log(L=L_{Edd})$ ' 0:56 dex. Points are color-coded by redshift, with error bars omitted for clarity. The scatter could possibly be due to challenges in measuring the continuum ux in low-luminosity AGNs. This illustrates the impact of spectral resolution and calibration di erences on the recovered accretion ratios.

Chapter 5

5 Time-Domain Analysis

5.1 Light Curve Sample

Variability is a de ning characteristic of AGN, arising from instabilities in the accretion ow and the response of the BLR. Flux variations on time scales from days to years encode information on the size and structure of the accretion disk (e.g., Krolik, 1999; Peterson, 2001) and the dynamical response of the BLR gas (Kaspi et al., 2000; Bentz et al., 2009). Time domain studies therefore provide an essential complement to single epoch spectroscopy by probing physical processes that cannot be captured in static spectra alone.

In this chapter, we preliminary explore another important feature of these two large scale surveys that deliver, in addition to spectra, the multi epoch optical data: the SDSS Stripe 82 (Ivezi¢ et al., 2007; Sesar et al., 2007) and the Gaia Data Release 3 (Vallenari et al. 2023; Eyer et al. 2023; Carnerero et al. 2023). Stripe 82 provides densely sampled, multicolor light curves covering nearly a decade of observations, but its narrow equatorial footprint limits our sample to only 7 AGN with spectroscopic coverage. By contrast, Gaia o ers all sky epoch photometry in the G, BP, and RP bands. Although its cadence is irregular due to the satellite's scanning law, the space based stability of Gaia ensures exceptionally precise single epoch photometry, enabling high quality light curves for 48 of our AGN. Together, these data sets provide complementary strengths: Stripe 82 contributes long baseline, multi band monitoring from the ground, while Gaia adds high precision space based measurements with uniform all sky coverage. The extraction of light curves from SDSS stripe82 and Gaia databases are described in Sections 2.4.1 and 2.4.2, respectivly.

The analyses was done in three stages. First, we describe the preprocessing required to harmonize the SDSS and Gaia light curves, since the two datasets di er substantially in data format, cadence, and photometric system. We then outline the methodology used to quantify variability, applying the same set of diagnostics, including cadence statistics, ux-based variability amplitudes, and wavelength-dependent trends, to both surveys. Finally, we compare the resulting variability properties across SDSS and Gaia, and place them in context with the spectroscopic correlations established in earlier chapters. This uni ed framework enables us to evaluate the robustness of variability diagnostics and to assess whether the time domain properties inferred from Gaia are consistent with those obtained from deeper but spatially restricted ground based monitoring.

5.2 Preprocessing

5.2.1 SDSS Stripe 82

The Stripe 82 variability catalog contains 67,507 candidate variable point sources (Ivezi¢ et al., 2007), each indexed by an internal integer identi er rather than by sky coordinates or SDSS object names. Recovering the subset corresponding to our AGN sample, therefore, required an explicit positional cross-match rather than a simple name-based query.

We rst compile the list of target sources from our spectroscopic sample (see Chapter 2) and convert their equatorial coordinates into SkyCoord objects in Python. These were matched against the master Stripe 82 catalog (see Section 2.4.1 for details) using a tolerance of 1^{00}, a value comfortably lower than the typical SDSS astrometric uncertainty. This ensured robust associations while minimizing the possibility of spurious matches in the crowded equatorial eld. The resulting integer IDs were then used to locate the corresponding per-object light-curve les.

Each light curve in Stripe 82 is stored in a separate le named by its catalog ID rather than the usual SDSS designation. Using the matched IDs, we retrieved the les corresponding to our 7 AGN located within the Stripe 82 region. Each le contains multi-epoch photometry in the ve SDSS lters (u; g; r; i; z), spanning nearly a decade of monitoring.

After retrieval, the data were reformatted into a standardized structure with columns for observation time (MJD), lter, magnitude, and magnitude uncertainty. All epochs were sorted chronologically to ensure accurate cadence statistics. Since SDSS Stripe82 catalogue provides only magnitudes and their uncertainties, for our comparison analysis these were converted into relative ux densities and ux errors according to the general photometry equation

$$f = 10^{\ 0:4\,m}; \qquad {}_f = f\ (0:4 \ln 10)\ {}_m; \tag{9}$$

where m and $_m$ denote the observed magnitude and its uncertainty. This formulation provides internally consistent ux units, su cient for variability and comparative analyses, though not tied to the absolute AB ux scale. Quality control was applied by removing epochs with missing values or non-physical uncertainties. The resulting lter-separated, relative ux-calibrated light curves form the basis for the variability diagnostics presented later in this chapter.

5.2.2 Gaia Epoch Photometry

For the Gaia sample, epoch photometry from DR3 was retrieved for all 48 AGN in our spectroscopic set (Vallenari et al., 2023; Eyer et al., 2023). The access was performed using the unique source_id assigned to each object, which is linked directly to the published per-transit tables (see Section 2.4.2 for details). These tables provide calibrated measurements

in the G, BP, and RP bands, including magnitudes, uxes, and associated uncertainties.

The downloaded data were reshaped from their native wide format into a uniform long format with standardized columns: observation time (MJD), band identi er, magnitude, magnitude error, ux and their associated uncertainties.

Quality control was applied by discarding epochs agged as unreliable in the Gaia tables, ensuring that only robust ux and error measurements were retained. The outcome of this procedure is a set of fully cleaned, band-separated light curves for the G, BP, and RP bands of the 48 sources. These data are placed in the same standardized structure as the Stripe 82 light curves described in Section 5.2.1, providing a harmonized foundation for the variability diagnostics presented in the following subsection.

5.3 Computation of Variability Metrics

After preprocessing, we quanti ed the variability properties of each light curve using a uniform set of statistical diagnostics. The same methodology was applied to both SDSS Stripe 82 and Gaia epoch photometry, allowing for direct comparison despite their di erences in cadence and bandpass coverage. For each light curve, the following steps were carried out:

Cadence statistics. Observation times were rst chronologically sorted and the total monitoring baseline $T_{baseline}$ was calculated as the di erence between the earliest and the latest epochs. The time intervals between consecutive observations, $t_i = t_{i+1} \quad t_i$, were then used to compute the median, mean, and 5th and 95th percentiles of the sampling distribution (t_{median}, t_{mean}, t_5, t_{95}, respectively). These parameters capture both the short-term clustering (e.g., intra-night visits in SDSS or Gaia scanning clusters) and long-term gaps (seasonal in SDSS, scanning-law driven in Gaia).

Flux normalization. Magnitudes and uncertainties were converted into relative uxes and ux errors. The uxes were normalized by their mean, so that all variability estimates are dimensionless and directly comparable between lters and surveys. This step also mitigates di erences in absolute brightness between sources.

Variability amplitude. The intrinsic variability of each light curve was quanti ed using the fractional variability statistic F_{var}:

$$F_{var} = \frac{\sqrt{S^2 \quad h \quad {}^2_{err} i}}{hF i}; \tag{10}$$

where S^2 is the variance of the observed uxes, $h \ {}^2_{err} i$ is the uncertainty of measurement by the mean squares, and $hF i$ is the mean ux. This de nition corrects for measurement noise and

yields a dimensionless measure of the intrinsic amplitude of variability. Whenever $S^2 \leq \langle \sigma_{err}^2 \rangle$, F_{var} was set to zero, indicating that there was no variability above the photometric noise floor.

Summary quantities. For each band of every source, we recorded the following:

- N: number of valid epochs,
- $T_{baseline}$: total time span covered by the observations,
- t_{median}, t_{mean}, t_5, t_{95}: characteristic cadence intervals,
- $\langle F \rangle$, σ_F: mean and standard deviation of the flux distribution,
- F_{var}: fractional variability amplitude.

Consistency across surveys. These diagnostics were computed identically for both SDSS and Gaia. This ensures that differences in measured variability amplitudes or cadence properties can be attributed to survey characteristics (e.g., Stripe 82's seasonal ground-based coverage vs. Gaia's irregular scanning law sampling) rather than to the analysis methodology. The resulting metrics are listed in Table 5 and provide a standardized framework for the comparison of properties in the time domain in the two datasets.

5.3.1 Common Sample: SDSS and Gaia

Figure 30 shows the composite light curves of the 7 AGN with coverage from both SDSS Stripe 82 and Gaia DR3. The corresponding variability metrics for these objects are listed in Table 3, including the number of epochs, baseline duration, cadence, and F_{var} for each band.

In SDSS Stripe 82 data, the decade-long multiband monitoring reveals strong flux variations, with amplitudes increasing toward the bluer filters (u; g) and diminishing in the redder bands (i; z). Gaia confirms the same wavelength dependence: variability is highest in the BP band and systematically weaker in RP. This behavior reflects the well-established AGN signature in which disk emission dominates at shorter wavelengths.

The fractional variability amplitudes (F_{var}) measured from SDSS and Gaia summarized in Table 4 are consistent within uncertainties. Both surveys agree on the wavelength-dependent ordering: stronger variability in the blue, weaker in the red. Some epochs of increased activity are recorded in both datasets, confirming that the variability is intrinsic to the AGN rather than a calibration artifact.

Outliers are rare in the SDSS light curves, which appear smooth and consistent. In contrast, Gaia data, especially in the RP band, show more scattered points and sometimes also in the BP band. It is not clear what is the origin of the outliers as Gaia photometry

is in general very precise, but it possible could be due to the fact that Gaia pipelines are focused mainly on stars. Overall, these outliers do not a ect the main result: both surveys show consistent AGN variability, with small di erences explained by survey-speci c e ects.

The scatter across bands of fractional variability (F_{var}), as shown in Tables 3 and 4, is largely due to obvious outliers (see Figures 30 and 31). This metric should be calculated only after the outliers have been removed, for example, using a sigma-clipping technique.

Figure 30: Composite light curves for the rst three AGN observed in both SDSS Stripe 82 (multi-band $u; g; r; i; z$; colored points) and Gaia DR3 (epoch photometry in G, BP, RP; green/blue/red points). The plots illustrate the complementary strengths of the two surveys: long-term, multi-band ground-based coverage from SDSS and high-precision, space-based photometry from Gaia.

Figure 31: The same is in Figure 30 but for the remaining four AGNs.

Table 3: Variability features of seven objects with light curves both in SDSS and Gaia surveys.

Object	Survey	Object ID	Band	N (days)	$T_{baseline}$ (days)	t_{median} (days)	F_{var}
SDSS_J0322+0055	SDSS	LC_3861327	u	72	3322	3.9	0.139± 0.002
			g	72	3322	3.9	0.124± 0.001
			r	72	3322	3.9	0.111± 0.001
			i	71	3322	4.0	0.085± 0.002
			z	71	3322	3.9	0.090± 0.002
	Gaia	3267488664984941184	G	26	800	22.7	0.104± 0.001
			BP	23	912	22.9	0.105± 0.002
			RP	23	800	22.9	0.102± 0.002
SDSS_J0138 0109	SDSS	LC_947097	u	26	2178	10.9	0.079± 0.003
			g	28	2167	9.01	0.088± 0.002
			r	31	2178	8.5	0.069± 0.002
			i	31	2178	8.5	0.175± 0.003
			z	31	2178	8.5	0.041± 0.003
	Gaia	2508869686708806144	G	58	913	0.2	0.072± 0.001
			BP	56	913	0.2	3.388± 0.001
			RP	56	913	0.2	3.889± 0.001
SDSS_J2220+0110	SDSS	LC_1125521	u	53	2948	6.0	0.183± 0.002
			g	56	2948	6.1	0.198± 0.002
			r	55	2948	6.1	0.169± 0.001
			i	54	2942	6.1	0.124± 0.001
			z	52	2942	6.1	0.138± 0.002
	Gaia	2679303506661773696	G	24	932	22.5	2.479± 0.001
			BP	22	932	22.5	0.225± 0.003
			RP	22	932	22.5	0.187± 0.002
SDSS_J0140 0050	SDSS	LC_1340680	u	72	3337	4.9	0.104± 0.002
			g	77	3337	3.9	0.119± 0.002
			r	76	3337	4.01	0.111± 0.002
			i	77	3337	3.9	0.117± 0.002
			z	77	3337	3.9	0.084± 0.002
	Gaia	2508952772851444736	G	62	913	0.2	0.097± 0.000
			BP	60	913	0.2	2.874± 0.001
			RP	60	913	0.2	2.750± 0.001
SDSS_J0239 0001	SDSS	LC_1992659	u	68	3337	4.9	0.090± 0.001
			g	66	3337	3.8	0.190± 0.001
			r	69	3337	3.9	0.083± 0.001
			i	69	3337	3.9	0.083± 0.001
			z	69	3337	3.9	0.071± 0.001
	Gaia	2498737553915316480	G	21	912	24.7	0.048± 0.001
			BP	22	912	23.1	0.051± 0.001
			RP	22	912	23.1	0.035± 0.001
SDSS_J2302 0107	SDSS	LC_524243	u	54	2573	4.9	0.040± 0.002
			g	56	2935	4.9	0.050± 0.002
			r	57	2935	4.9	0.049± 0.001
			i	58	2935	4.9	0.048± 0.001
			z	57	2935	4.9	0.057± 0.002
	Gaia	2650710023881781248	G	25	933	21.7	0.051± 0.001
			BP	22	932	22.1	0.067± 0.003
			RP	22	933	22.1	2.322± 0.002
SDSS_J0310 0050	SDSS	LC_3087679	u	71	3328	3.0	0.144± 0.001
			g	72	3328	3.0	0.128± 0.001
			r	73	3328	3.0	0.122± 0.001
			i	73	3328	3.0	0.094± 0.001
			z	72	3328	3.0	0.099± 0.001
	Gaia	3265557372810761344	G	58	913	24.9	0.124± 0.001
			BP	56	913	25.1	3.088± 0.001
			RP	56	913	25.1	0.831± 0.001

Summary of variability properties for the 7 AGN with both SDSS Stripe 82 and Gaia DR3 light curves. For each source, both the SDSS designation and the internal catalog identi er are listed.

Band	Count	Mean	Std. Dev.	Min	25%	Median	Max
BP	7	1.40	1.61	0.05	0.09	0.22	3.39
G	7	0.42	0.91	0.05	0.06	0.10	2.48
RP	7	1.45	1.54	0.04	0.14	0.83	3.89
g	7	0.13	0.05	0.05	0.10	0.12	0.20
i	7	0.10	0.04	0.05	0.08	0.09	0.18
r	7	0.10	0.04	0.05	0.08	1.11	0.17
u	7	0.11	0.05	0.04	0.08	0.10	0.18
z	7	0.8	0.03	0.04	0.06	0.08	0.14

Table 4: Summary statistics of F_{var} for the 7 AGNs with SDSS and Gaia.

5.3.2 Variability Analysis of only Gaia Light Curves

To evaluate the optical variability of the AGN sample in the Gaia bands, for which more data is available, we computed the fractional variability amplitude (F_{var}) for each source in the three available photometric lters: BP, G, and RP. A total of 48 AGNs had su cient epoch photometry to allow for this analysis, with the full set of variability features listed later in this chapter in Tables 6 and 7. It is important to note that outliers have not been removed before calculating the fractional variability, as it was outside the scope of this thesis.

From this preliminary analysis of fractional variability, summarized in Table 5, we noticed that the G band exhibits the lowest variability, with a median F_{var} of only 0.086 and a narrow spread, re ecting Gaia's higher precision and broader response ofGaia in this band. In contrast, the BP and RP bands show substantially higher variability. The median F_{var} increases to 0.23 inBP and 1.66 in RP, and both bands show a much wider spread and higher maximum values, revealing a subset of highly variable AGNs.

To visualize these distributions, we produced histograms with overlaid kernel density estimates (KDEs) for each band. These con rm strong clustering at low variability in G, while BP and RP show long tails, indicating strong variability in a subset of sources. However, once again we emphasize that these results may be signi cantly a ected by outliers; thus, a more profound analysis should include rst the procedure of outlier removal.

Band	Count	Mean	Std. Dev.	Min	25%	Median	Max
BP	48	1.47	1.65	0.05	0.09	0.23	4.62
G	48	0.14	0.35	0.00	0.07	0.09	2.48
RP	48	1.76	1.61	0.03	0.13	1.66	4.99

Table 5: Summary statistics of F_{var} for the 48 AGNs with Gaia-only light curves.

In addition, the cumulative distribution functions (CDFs) in Figure 33 highlight the contrast in the variability behavior between bands. Although nearly 90% of AGN in the G band have $F_{var} < 0:2$, the BP and RP distributions are signi cantly broader, con rming a greater fraction of moderate to highly variable AGN in these bands.

Figure 32: Histograms with KDE overlays showing the distribution of F_{var} values for the BP, G, and RP bands (top to bottom) across the 48 AGN with Gaia-only light curves.

Figure 33: Cumulative distribution of F_{var} for the BP, G, and RP bands. The G band shows a steep rise, indicating low variability for most AGN. BP and RP exhibit longer tails, revealing greater diversity in photometric variability.

These results demonstrate that, Gaia, the DR3 epoch photometry is capable of revealing low-level and high-amplitude photometric variability in Type 1 AGNs. The increased variability in BP and RP could re ect the variations in the AGN continuum slope and the accretion disk processes that manifest more strongly outside the central broad G band. However, further analysis should be performed to demonstrate is this the true observed property or the artifact of the systematics.

In summary, the Gaia photometric data reveal detectable variability in all target AGNs, where the ux originates predominantly from the accretion disk. Although variability amplitudes are generally lower in the G band - due to its wider spectral coverage - the distribution of F_{var} values throughout the sample con rms the presence of intrinsic uctuations in optical emission. However, consistent detection of variability across bands supports its utility in identifying and studying active nuclei in large-scale surveys.

Table 6: Variability features of 48 objects with light curves in Gaia surveys.

source_id	band	N (days)	$T_{baseline}$ (days)	t_{median} (days)	t_{mean} (days)	t_5 (days)	t_{95} (days)	$\langle F \rangle$	F	F_{var}
1016834950437655552	BP	40	983.425	0.250	25.216	0.074	98.929	9509.491	28021.500	2.947
1016834950437655552	G	40	983.425	0.250	25.216	0.074	100.546	4747.859	670.275	0.141
1016834950437655552	RP	40	983.425	0.250	25.216	0.074	98.929	14191.978	40098.097	2.825
1028155389013411584	BP	29	983.350	23.784	35.120	0.074	126.687	6783.412	19846.032	2.926
1028155389013411584	G	31	983.424	19.341	32.781	0.074	126.620	4921.049	390.500	0.079
1028155389013411584	RP	29	983.350	23.784	35.120	0.074	126.687	9688.219	34729.104	3.585
1249947878136029824	BP	46	906.028	0.176	20.134	0.074	96.853	5967.381	479.473	0.080
1249947878136029824	G	46	906.028	0.176	20.134	0.074	96.853	9727.225	651.568	0.067
1249947878136029824	RP	46	906.028	0.176	20.134	0.074	96.853	6519.610	411.872	0.063
1251676310414862464	BP	63	905.274	0.176	14.601	0.074	75.254	6546.051	532.939	0.081
1251676310414862464	G	67	905.274	0.176	13.716	0.074	75.226	5624.822	800.352	0.142
1251676310414862464	RP	63	905.274	0.176	14.601	0.074	75.254	9034.437	565.885	0.062
1269203419114884608	BP	41	901.267	10.969	22.532	0.074	97.071	18499.271	33835.079	1.829
1269203419114884608	G	44	901.267	4.827	20.960	0.074	93.553	19688.440	1346.209	0.068
1269203419114884608	RP	41	901.267	10.969	22.532	0.074	97.071	19431.155	34324.539	1.766
1385484951370252160	BP	42	950.279	26.501	23.178	0.074	60.727	2676.742	288.268	0.107
1385484951370252160	G	47	950.279	22.363	20.658	0.074	57.593	4132.185	438.992	0.106
1385484951370252160	RP	42	950.279	26.501	23.178	0.074	60.727	3064.284	456.557	0.149
1401637979773304960	BP	40	948.769	27.250	24.327	0.074	40.875	4140.100	236.593	0.056
1401637979773304960	G	41	948.769	27.250	23.719	0.074	40.875	6399.795	329.960	0.050
1401637979773304960	RP	40	948.769	27.250	24.327	0.074	40.875	3682.085	188.048	0.050
1420458114147368064	BP	34	907.822	31.061	27.510	0.074	47.318	5460.767	23893.226	4.375
1420458114147368064	G	34	907.822	31.163	27.510	0.074	60.407	1952.473	185.563	0.095
1420458114147368064	RP	34	907.822	31.061	27.510	0.074	47.318	4283.625	17596.652	4.108
1500142058274857984	BP	82	983.449	0.176	12.141	0.074	45.732	4151.908	15444.199	3.720
1500142058274857984	G	85	983.449	0.176	11.708	0.074	45.725	3292.201	445.023	0.135
1500142058274857984	RP	82	983.449	0.176	12.141	0.074	45.732	2947.216	286.742	0.097
1514610016387783296	BP	37	886.518	10.092	24.625	0.074	104.087	7073.106	523.325	0.074
1514610016387783296	G	37	886.518	3.590	24.625	0.074	117.770	12747.688	1154.456	0.090
1514610016387783296	RP	37	886.518	10.092	24.625	0.074	104.087	13738.539	915.951	0.067
1532557997522989184	BP	38	909.697	0.250	24.586	0.074	103.593	6407.765	24254.927	3.785
1532557997522989184	G	39	909.697	0.176	23.939	0.074	102.432	3954.874	118.108	0.030
1532557997522989184	RP	38	909.697	0.250	24.586	0.074	103.593	6693.898	25134.248	3.755
1572729067161813760	BP	65	961.659	0.176	15.026	0.074	54.685	5155.194	19134.727	3.712
1572729067161813760	G	67	961.659	0.176	14.571	0.074	54.535	3858.084	197.817	0.051
1572729067161813760	RP	65	961.659	0.176	15.026	0.074	54.685	4352.028	9848.155	2.263
1680733166644357760	BP	40	968.403	25.930	24.831	0.074	54.648	5495.279	514.044	0.093
1680733166644357760	G	45	968.403	24.245	22.009	0.074	44.471	7980.844	822.550	0.103
1680733166644357760	RP	40	968.403	25.930	24.831	0.074	54.648	8637.218	457.804	0.053
1680740828866041088	BP	45	968.329	25.587	22.007	0.074	52.289	2875.669	207.020	0.071
1680740828866041088	G	49	968.329	22.834	20.174	0.074	50.729	4208.359	244.797	0.058
1680740828866041088	RP	45	968.329	25.587	22.007	0.074	52.289	3499.343	2143.937	0.613
2462859664867002112	BP	50	910.031	0.176	18.572	0.074	105.452	3092.439	948.011	0.306
2462859664867002112	G	52	910.031	0.176	17.844	0.074	103.963	4378.252	375.598	0.086
2462859664867002112	RP	50	910.031	0.176	18.572	0.074	105.452	5707.946	16175.770	2.834
2498737553915316480	BP	22	911.793	23.067	43.419	0.074	142.985	8884.971	459.828	0.051
2498737553915316480	G	21	911.793	24.745	45.590	0.074	144.329	13563.989	650.139	0.048
2498737553915316480	RP	22	911.793	23.067	43.419	0.074	142.985	8020.711	283.247	0.035
2508869686708806144	BP	56	912.043	0.176	16.583	0.074	102.207	4627.319	15677.442	3.388
2508869686708806144	G	58	912.043	0.176	16.001	0.074	100.248	3736.483	271.291	0.072
2508869686708806144	RP	56	912.043	0.176	16.583	0.074	102.207	5016.070	19506.504	3.889
2508952772851444736	BP	60	912.293	0.176	15.463	0.074	98.933	9506.454	27319.255	2.874
2508952772851444736	G	62	912.293	0.176	14.956	0.074	97.081	6959.864	677.619	0.097
2508952772851444736	RP	60	912.293	0.176	15.463	0.074	98.933	8330.638	22905.931	2.750
2560047421790955264	BP	72	912.124	0.176	12.847	0.074	93.689	4379.753	13686.971	3.125
2560047421790955264	G	75	912.124	0.176	12.326	0.074	92.851	4030.793	268.130	0.066
2560047421790955264	RP	72	912.124	0.176	12.847	0.074	93.689	4215.557	12141.012	2.880
2562478059747868160	BP	40	914.128	0.250	23.439	0.074	134.499	2597.841	240.038	0.092
2562478059747868160	G	43	914.305	0.176	21.769	0.074	133.045	3790.452	130.522	0.034
2562478059747868160	RP	40	914.128	0.250	23.439	0.074	134.499	2203.136	658.467	0.299
2650710023881781248	BP	22	932.736	22.172	44.416	0.074	164.329	1923.997	132.638	0.067
2650710023881781248	G	25	932.736	21.713	38.864	0.074	139.709	3091.387	158.767	0.051
2650710023881781248	RP	22	932.736	22.172	44.416	0.074	164.329	4411.187	10244.249	2.322
2679303506661773696	BP	22	931.481	22.501	44.356	0.074	141.626	2627.324	593.052	0.225
2679303506661773696	G	24	931.481	22.501	40.499	0.074	141.618	8759.913	21712.340	2.479
2679303506661773696	RP	22	931.481	22.501	44.356	0.074	141.626	3307.250	618.866	0.187
2687420896067111040	BP	47	931.157	0.176	20.243	0.074	121.212	3518.864	669.805	0.190
2687420896067111040	G	51	931.157	0.176	18.623	0.074	115.226	4424.756	142.518	0.032
2687420896067111040	RP	47	931.157	0.176	20.243	0.074	121.212	4976.133	955.961	0.192
2745007370896353664	BP	64	1019.064	0.176	16.176	0.074	95.814	5916.460	25558.549	4.320
2745007370896353664	G	62	1019.064	0.176	16.706	0.074	100.737	2458.511	131.119	0.053
2745007370896353664	RP	64	1019.064	0.176	16.176	0.074	95.814	4177.386	20834.862	4.988
2819126102559044096	BP	40	1019.895	0.176	26.151	0.074	126.703	11453.581	39959.504	3.489

Table 7: Rest of 48 objects with light curves in Gaia surveys.

source_id	band	N	$T_{baseline}$	t_{median}	t_{mean}	t_5	t_{95}	$\langle F \rangle$	$_F$	F_{var}
2819126102559044096	G	41	1019.895	0.176	25.497	0.074	126.625	3610.604	372.331	0.103
2819126102559044096	RP	40	1019.895	0.176	26.151	0.074	126.703	11255.785	35722.384	3.174
2841448525065809536	BP	59	1013.064	0.176	17.467	0.074	96.478	4168.335	19245.517	4.617
2841448525065809536	G	59	1013.064	0.176	17.467	0.074	96.478	2632.438	218.520	0.083
2841448525065809536	RP	59	1013.064	0.176	17.467	0.074	96.478	6520.069	26017.510	3.990
290930265608514816	BP	23	920.581	16.254	41.845	0.074	166.708	2618.254	472.875	0.180
290930265608514816	G	23	920.581	16.166	41.845	0.074	166.708	4240.555	680.611	0.160
290930265608514816	RP	23	920.581	16.254	41.845	0.074	166.708	9697.133	32251.564	3.326
3265557372810761344	BP	22	910.866	25.176	43.375	0.074	143.486	11333.020	34993.669	3.088
3265557372810761344	G	23	910.866	24.922	41.403	0.074	142.687	6524.835	811.850	0.124
3265557372810761344	RP	22	910.866	25.176	43.375	0.074	143.486	6075.595	5049.293	0.831
3267488664984941184	BP	23	799.881	22.958	36.358	0.074	142.720	2881.932	303.635	0.105
3267488664984941184	G	26	911.368	22.744	36.455	0.074	140.426	4890.343	509.094	0.104
3267488664984941184	RP	23	799.881	22.958	36.358	0.074	142.720	3751.410	383.671	0.102
3985601493159227392	BP	16	940.271	26.171	62.685	0.074	176.873	1959.932	333.258	0.169
3985601493159227392	G	17	940.271	25.296	58.767	0.074	174.241	2491.215	288.754	0.115
3985601493159227392	RP	16	940.271	26.171	62.685	0.074	176.873	2297.751	306.740	0.133
3990157457027790848	BP	36	941.199	0.250	26.891	0.074	134.231	3242.923	303.295	0.093
3990157457027790848	G	36	941.199	0.250	26.891	0.074	134.231	4714.768	377.499	0.080
3990157457027790848	RP	36	941.199	0.250	26.891	0.074	134.231	2803.766	678.048	0.242
3990830907900291584	BP	26	779.410	19.754	31.176	0.074	137.626	23718.406	33431.696	1.410
3990830907900291584	G	27	800.160	20.252	30.775	0.074	137.286	25114.740	1060.121	0.042
3990830907900291584	RP	26	779.410	19.754	31.176	0.074	137.626	21365.277	33253.427	1.556
4023588108070402816	BP	36	945.212	0.176	27.006	0.074	131.800	3972.346	349.074	0.087
4023588108070402816	G	35	945.212	0.176	27.800	0.074	132.516	5902.463	509.905	0.086
4023588108070402816	RP	36	945.212	0.176	27.006	0.074	131.800	4572.225	4799.778	1.050
4458012573911009664	BP	41	906.284	0.213	22.657	0.074	120.278	6570.424	1026.310	0.156
4458012573911009664	G	42	906.284	0.250	22.104	0.074	111.941	8127.089	1632.933	0.201
4458012573911009664	RP	41	906.284	0.213	22.657	0.074	120.278	10180.951	911.005	0.089
6342859972491701 76	BP	44	937.683	0.250	21.807	0.074	131.525	9752.863	721.518	0.074
6342859972491701 76	G	45	937.683	0.213	21.311	0.074	129.430	14798.218	1053.257	0.071
6342859972491701 76	RP	44	937.683	0.250	21.807	0.074	131.525	8777.520	542.670	0.062
670950827385990912	BP	35	969.135	9.755	28.504	0.074	135.827	2192.660	382.030	0.174
670950827385990912	G	36	969.135	8.079	27.690	0.074	135.484	2884.654	651.330	0.226
670950827385990912	RP	35	969.135	9.755	28.504	0.074	135.827	10284.597	32932.283	3.202
678613976755423488	BP	44	941.156	0.176	21.887	0.074	130.906	4830.699	494.706	0.102
678613976755423488	G	44	918.410	0.176	21.358	0.074	130.906	7552.866	720.666	0.095
678613976755423488	RP	44	941.156	0.176	21.887	0.074	130.906	4711.241	377.199	0.080
740685668749400064	BP	48	941.698	0.250	20.036	0.074	94.203	6970.315	827.623	0.119
740685668749400064	G	50	941.698	0.176	19.218	0.074	94.016	10485.120	719.944	0.069
740685668749400064	RP	48	941.698	0.250	20.036	0.074	94.203	6215.182	953.093	0.153
741343039264023552	BP	28	942.023	13.254	34.890	0.074	140.130	9145.152	32611.112	3.566
741343039264023552	G	30	942.023	13.005	32.484	0.074	139.987	2864.219	232.286	0.080
741343039264023552	RP	28	942.023	13.254	34.890	0.074	140.130	10146.908	32647.699	3.217
808775434561821440	BP	54	979.598	0.176	18.483	0.074	93.798	3067.009	216.381	0.070
808775434561821440	G	55	979.598	0.176	18.141	0.074	93.774	4090.739	299.812	0.073
808775434561821440	RP	54	979.598	0.176	18.483	0.074	93.798	3018.298	884.987	0.293
819907023423890304	BP	74	979.847	0.176	13.423	0.074	93.788	3912.251	5110.572	1.306
819907023423890304	G	78	979.847	0.176	12.725	0.074	93.357	4885.961	467.331	0.096
819907023423890304	RP	74	979.847	0.176	13.423	0.074	93.788	8478.842	30728.032	3.624
827078008196318592	BP	61	985.180	0.176	16.420	0.074	74.451	2095.424	849.594	0.405
827078008196318592	G	61	985.180	0.176	16.420	0.074	74.451	2872.877	332.522	0.114
827078008196318592	RP	61	985.180	0.176	16.420	0.074	74.451	4389.015	17057.863	3.886
858200268935710208	BP	56	996.697	11.005	18.122	0.074	53.213	10771.959	25432.388	2.361
858200268935710208	G	53	996.623	12.791	19.166	0.074	58.106	11978.418	1686.396	0.140
858200268935710208	RP	56	996.697	11.005	18.122	0.074	53.213	16919.325	68344.610	4.039
860383349272777344	BP	66	994.296	0.250	15.297	0.074	44.667	7340.762	28333.680	3.860
860383349272777344	G	69	994.296	0.176	14.622	0.074	44.657	3194.750	218.485	0.063
860383349272777344	RP	66	994.296	0.250	15.297	0.074	44.667	8661.361	29951.324	3.458
882564415815249280	BP	29	974.077	18.503	34.788	0.074	137.034	8569.905	33035.484	3.855
882564415815249280	G	33	974.151	11.002	30.442	0.074	136.980	3324.480	391.986	0.000
882564415815249280	RP	29	974.077	18.503	34.788	0.074	137.034	2748.356	183.621	0.066
895654342263361152	BP	39	974.578	0.176	25.647	0.074	141.100	2516.643	585.487	0.232
895654342263361152	G	41	974.578	0.176	24.364	0.074	138.570	3701.850	790.197	0.213
895654342263361152	RP	39	974.578	0.176	25.647	0.074	141.100	5968.145	19333.267	3.239
899509642346667392	BP	50	976.156	0.176	19.922	0.074	133.110	10780.116	27920.963	2.590
899509642346667392	G	52	976.156	0.176	19.140	0.074	133.053	7663.647	671.393	0.086
899509642346667392	RP	50	976.156	0.176	19.922	0.074	133.110	13275.966	31916.578	2.404
926435994915623552	BP	30	979.090	19.175	33.762	0.074	136.126	4097.931	596.288	0.145
926435994915623552	G	30	979.090	19.175	33.762	0.074	136.126	6044.171	746.219	0.123
926435994915623552	RP	30	979.090	19.175	33.762	0.074	136.126	7292.509	18704.537	2.565

Chapter 6

6 Discussion

A central goal of this thesis was to evaluate how reliably AGN spectroscopic properties can be recovered from low-resolution, time-averaged BP / RP spectra provided by Gaia DR3, compared with the higher resolution spectra from SDSS. The analysis focused on a controlled sample of 60 Type 1 AGNs for which both datasets are available. Using a uniform tting framework, diagnostics extracted from both surveys were systematically compared across a range of spectral features. The results clearly point to a dichotomy in Gaia XP's performance: continuum luminosities and integrated broad-line uxes are recovered with signi cant delity, while parameters sensitive to velocity structure or spectral decomposition show higher scatter. Speci cally, we nd that the continuum uxes at 5100 Å and 6200 Å measured from Gaia spectra correlate very strongly with those from SDSS, with minimal systematic o sets. This is encouraging, as these uxes are directly tied to accretion-disk emission and are widely used for deriving monochromatic luminosities and bolometric corrections in AGN studies. Similarly, the broad H and H line uxes exhibit moderate to strong correlation despite the coarse resolution of the XP, suggesting that Gaia is capable of capturing most of the output from the line-emitting region. However, the ability to resolve individual components, particularly narrow lines or complex blends, remains limited. Parameters such as FWHM and the Balmer decrement (H /H) show weak to no correlation, as expected for low-resolution spectra unable to distinguish ne velocity structures. In case of fwhm Gaia fwhm are larger than SDSS, especially in H due to di culties in determining the underlying continuum. When propagating these spectral features into physical quantities, we observe a continued trend; parameters relying primarily on luminosity are well recovered, while those depending on both luminosity and velocity widths started scattering more. The calculated M_{BH} shows a strong correlation with an increasing shift towards Gaia, which is re ecting the systematically larger FWHM values in Gaia. Nevertheless, the correlation indicates that Gaia XP can still provide reasonably consistent black hole mass measurements even in the absence of high-resolution spectra. On the other hand, $_{Edd}$, which depends both on the bolometric luminosity and the black hole mass, shows a weaker correlation with Gaia values typically lower than SDSS.

This nding highlights a key limitation: while Gaia XP is suitable for ux-based diagnostics, it cannot replace classical spectroscopy for spectral decomposition and detailed line-shape analyses or very precise black hole mass estimates that rely on velocity widths. However, these ndings demonstrate that Gaia's low-resolution spectra can still be valuable for large-scale studies. With millions of available AGN spectra, Gaia opens the door to statistical and population-wide studies that are otherwise limited by the smaller coverage of traditional spectroscopic surveys.

To better contextualize the global agreement between SDSS and Gaia-derived parameters, we constructed a Pearson correlation heatmap that captures the cross-survey consistency of all measured quantities. The matrix reveals a clear strati cation in performance: quantities like continuum luminosities, broad-line uxes, and line ratios involving strong lines show the strongest correlation, while those involving widths or weak features fall behind. In particular, the luminosity correlations exceed $r = 0.9$, highlighting the stability of the spectrophotometric calibration of Gaia XP and its ability to trace the intrinsic AGN power between sources. The H ux is particularly robust, consistently aligning with SDSS across the sample, while H shows slightly more scatter due to its lower intrinsic strength and blending with nearby features. On the other hand, the FWHM values and Balmer decrements exhibit low correlation ($r < 0.4$), driven primarily by Gaia's inability to resolve broad versus narrow components or measure pro le shapes reliably. Interestingly, while the H /H ratio is not tightly correlated, it does not display strong systematic bias either, implying that Gaia may still o er ensemble-level constraints on dust extinction if analyzed statistically. In general, the heat map serves not just as a diagnostic tool, but as a summary of the strengths and limitations of Gaia XP: excellent for measuring total ux and luminosity indicators, but less suited for any parameter dependent on resolving sharp spectral features or velocity structures. This distinction will be crucial in designing future population-level AGN studies that aim to combine Gaia XP data with targeted high-resolution follow-up.

Figure 34: Cross-survey Pearson correlations (in $\log_{10}$ space) between SDSS spectral features (y-axis) and Gaia XP-derived features (x-axis). Continuum and luminosity quantities show high correlations ($r > 0.9$), while emission-line ux ratios and FWHM exhibit much weaker agreement.

Chapter 7

7 Conclusion

This thesis aimed to evaluate the reliability of Gaia DR3 XP spectra for AGN investigations by comparing them directly with SDSS spectra. The central motivation was to test whether the low-resolution optical spectrophotometry from Gaia can recover BLR properties and black hole mass estimates with su cient accuracy for large statistical studies, complementing ground-based surveys.

The main tasks carried out were:

- Spectral decomposition of SDSS and Gaia XP spectra using the FANTASY package, tting continuum, Fe II templates, and multiple broad and narrow emission-line components.
- Measurement and comparison of continuum luminosities (5100 Å, 6200 Å), broad H and H luminosities, and FWHM values for the broad Balmer lines.
- Correlation analysis of SDSS vs Gaia luminosities and line widths, with results presented as redshift-color-coded scatter plots.
- Calculation of black hole masses using the single-epoch H virial estimator (Eq. 6 of Greene and Ho 2005), and comparison of SDSS- and Gaia-based estimates.
- Comparison of light curves from SDSS Stripe 82 multi-band data and Gaia epoch photometry, with variability characterized using fractional variability (F_{var}), ux amplitudes, and cadence analysis.

The main conclusions drawn from this work are:

- Gaia XP spectra reproduce continuum and emission-line luminosities with strong correlation to SDSS values, con rming the reliability of ux calibration in XP data.
- Line-width (FWHM) measurements from Gaia show increased scatter relative to SDSS, re ecting the limitations imposed by low spectral resolution.
- Black hole masses estimated from H correlate well between SDSS and Gaia, but Gaia-based masses are systematically higher by $\tilde{}$ ' +0:76 dex, with a Pearson correlation coe cient of r 0:7.
- The propagated calibration uncertainties from the virial scaling relation provide consistent error bars in both datasets, while the 0:4 dex scatter observed is consistent with the intrinsic dispersion of single-epoch virial mass estimators.

- Variability studies reveal complementary strengths: SDSS Stripe 82 provides a long multi-band baseline, while Gaia adds stable space-based photometry with broad sky coverage but irregular cadence. In the case of Gaia light curves, however, outliers should be carefully detected and removed before variability metrics are interpreted.
- Overall, Gaia XP spectra are well-suited for statistical AGN studies of luminosities and black hole masses, but systematic e ects due to spectral resolution must be carefully accounted for in precision applications. Importantly, the ability to recover continuum luminosities with high delity also enables robust estimates of accretion rates (via Eddington ratios), thereby extending the utility of Gaia XP to studies of SMBH growth and AGN demographics.

This study reveals the utility of low-resolution, all-sky spectrophotometry in AGN science. From the studies of Gaia XP spectra, we tested the feasibility of extracting key AGN diagnostics in the absence of traditional long-slit spectroscopy. Although a detailed comparison with SDSS provides a validation benchmark, the broader implication is signi cant. XP spectra though not tailored for AGN research retain enough spectral information to recover global properties such as continuum luminosities and integrated line uxes. This paves the way for population-level AGN studies, even among sources without access to high-resolution ground-based data.

From a validation point, the reliable recovery of the broad line and continuum uxes indicates that Gaia XP spectra, despite their lower resolution, can still support physically meaningful results in a statistical sense. This opens up new avenues for ensemble studies targeting AGN demographics, bolometric luminosity functions, and accretion-related scaling relations. While the time-averaged nature and limited resolution of XP spectra preclude detailed line decomposition in individual cases, there is great potential in population studies. For instance, when combined with redshift and variability information, XP-derived features could serve as priors in Bayesian frameworks or provide inputs for probabilistic classi cation of AGNs in wide-area surveys.In this way, Gaia XP contributes not only to luminosity measurements but also to deriving fundamental SMBH parameters such as black hole mass and accretion rate, which are central to understanding AGN evolution.

Looking ahead, the future release of single-epoch Gaia XP spectra could open new opportunities for AGN time-domain studies. Combined with upcoming surveys such as Rubin Observatory LSST, Gaia XP provides a complementary spectrophotometric layer, covering millions of AGNs particularly faint or high-Galactic-latitude sources with limited spectroscopic follow-up. Single-epoch XP spectra can serve as a coarse spectral baseline for transient or variable AGNs, helping to ag candidates for reverberation mapping or periodicity searches. As methodology for spectral extraction, calibration, and tting continues to evolve, Gaia XP may emerge as a powerful statistical tool for the next generation of extragalactic science.

References

Abazajian, K. N. and et al. (2009). The seventh data release of the sloan digital sky survey. Astrophysical Journal Supplement Series, 182:543 558.

Ahumada, R., Allende Prieto, C., Anderson, S. F., Andrews, B. H., Anguiano, B., Aubourg, , and et al. (2020). The sixteenth data release of the sloan digital sky survey. The Astrophysical Journal Supplement Series, 249(3):3.

Almeida, A. and et al. (2023). The eighteenth data release of the sloan digital sky surveys: Targeting and rst spectra from sdss-v. The Astrophysical Journal Supplement Series, 267(2):44.

Antonucci, R. (1993). Uni ed models for active galactic nuclei and quasars. Annual Review of Astronomy and Astrophysics, 31:473 521.

Astropy Collaboration (2013). Astropy: A community python package for astronomy. A&A , 558:A33.

Astropy Collaboration (2018). The astropy project: Building an open-science project and status of the v2.0 core package. AJ, 156:123.

Baldwin, J. A., Phillips, M. M., and Terlevich, R. (1981). Classi cation parameters for the emission-line spectra of extragalactic objects. Publications of the Astronomical Society of the Paci c, 93:5 19.

Beckmann, V. and Shrader, C. R. (2012). Active Galactic Nuclei. Wiley-VCH, Weinheim, Germany.

Bentz, M. C., Peterson, B. M., Netzer, H., Pogge, R. W., and Vestergaard, M. (2009). The radius luminosity relationship for active galactic nuclei: The e ect of host-galaxy starlight on luminosity measurements. The Astrophysical Journal, 697(1):160 181.

Blandford, R. D., Meier, D. L., and Readhead, A. C. S. (2019). Relativistic jets from active galactic nuclei. Annual Review of Astronomy and Astrophysics, 57:467 509.

Bolton, A. S., Schlegel, D. J., Aubourg, E., et al. (2012). Spectral classi cation and redshift measurement for the sdss-iii baryon oscillation spectroscopic survey. The Astronomical Journal, 144(5):144.

Burke, D. J., Doe, S., Siemiginowska, A., Eadie, G., Refsdal, B., Laurino, O., and et al. (2022). Sherpa: Modeling and tting in python. Astrophysics Source Code Library ascl:2203.001.

Carnerero, M. I., Raiteri, C. M., Rimoldini, L., Busonero, D., Licata, E., Mowlavi, N., Lecoeur-Taïbi, I., Audard, M., Holl, B., Gavras, P., Klioner, S., Jevardat de Fombelle, G., Carballo, R., Clementini, G., Delchambre, L., Eyer, L., and others (Gaia Collaboration) (2023). Gaia data release 3: The rst gaia catalogue of variable agn. Astronomy & Astrophysics, 674:A24.

Carrasco, J. M., Weiler, M., Jordi, C., Fabricius, C., De Angeli, F., Evans, D. W., Riello, M., and Montegri o, P. (2021). Internal calibration of gaia bp/rp low-resolution spectra. Astronomy & Astrophysics, 652:A86.

De Angeli, F., Weiler, M., Montegri o, P., Evans, D. W., Riello, M., Andrae, R., et al. (2023). Gaia data release 3: Processing and validation of bp/rp low-resolution spectral data. Astronomy & Astrophysics, 674:A2.

Elitzur, M. (2008). Agn uni cation: An updated paradigm. New Astronomy Reviews, 52(6-7):274 288.

Eyer, L., Audard, M., Holl, B., Rimoldini, L., Carnerero, M. I., Clementini, G., et al. (2023). Gaia data release 3: Summary of the variability processing and analysis of photometric and spectroscopic time series of 1.8 billion sources. Astronomy & Astrophysics, 674:A13.

Fabian, A. C. (2012). Observational evidence of active galactic nuclei feedback. Annual Review of Astronomy and Astrophysics, 50:455 489.

Fausnaugh, M. M. et al. (2017). Reverberation mapping of optical emission lines in ve active galaxies. Astrophysical Journal, 840:97.

Gaia Collaboration, Prusti, T., de Bruijne, J., et al. (2016). The gaia mission. Astronomy Astrophysics, 595:A1.

Graham, M. and Group, I. T. W. (2014). Ivoa table access protocol version 1.0. In International Virtual Observatory Alliance .

Greene, J. E. and Ho, L. C. (2005). Estimating black hole masses in active galaxies using the h emissionline: The Astrophysical Journal; 630(1) : 122 129.

Gunn, J. E., Carr, M. A., Rockosi, C. M., Sekiguchi, M., Berry, K., Elms, B., de Haas, E., Ivezi¢, , Knapp, G. R., Lupton, R. H., Pauls, G., Simcoe, R., Hirsch, R., Sanford, D., Wang, S.-i., York, D. G., Harris, F., Annis, J., Bartozek, L., Boroski, W. N., Bakken, J. A., Haldeman, M., Kent, S., Holm, S., Holmgren, D., Petravick, D., Prosapio, A., Rechenmacher, R., Doi, M., Fukugita, M., Shimasaku, K., Okada, N., Hull, C., Siegmund, W., Mannery, E., Blouke, M., Heidtman, D., Schneider, D. P., Lucinio, R., and Brinkman, J. (1998). The sloan digital sky survey photometric camera. The Astronomical Journal, 116(6):3040 3081.

Gunn, J. E., Siegmund, W. A., Mannery, E. J., Owen, R., Hull, C. L., Leger, F., Carey, L. N., Knapp, G. R., York, D. G., et al. (2006). The 2.5-m telescope of the sloan digital sky survey. Astronomical Journal, 131:2332 2359.

Heckman, T. M. (1980). An optical and radio survey of the nuclei of bright galaxies - activity in normal galactic nuclei. Astronomy and Astrophysics, 87:152 164.

Heckman, T. M. and Best, P. N. (2014). The coevolution of galaxies and supermassive black holes: Insights from surveys of the contemporary universe. Annual Review of Astronomy and Astrophysics, 52:589 660.

Ho, L. C. (2008). Nuclear activity in nearby galaxies. Annual Review of Astronomy and Astrophysics, 46:475 539.

Huang, J., Ji, J., Wu, X.-B., and et al. (2024). Comprehensive correction of gaia dr3 xp spectra: Improving spectral slope and color accuracy. Astronomy Astrophysics.

Ili¢, D., Raki¢, N., and Popovi¢, L. (2023). Fantasy: Fully automated agn spectral tool. Astrophysical Journal Supplement Series, 267(1):19.

Ivezi¢, š., Lupton, R. H., Schlegel, D., Boroski, B., Adelman-McCarthy, J., Yanny, B., Kent, S., Stoughton, C., Finkbeiner, D., Padmanabhan, N., et al. (2004). Sloan digital sky survey standard star catalog for stripe 82. Astronomical Journal, 128:502 512.

Ivezi¢, š., Smith, J. A., Miknaitis, G., Lin, H., Tucker, D., Lupton, R. H., Gunn, J. E., Knapp, G. R., Strauss, M. A., Sesar, B., et al. (2007). Sloan digital sky survey standard star catalog for stripe 82: The dawn of industrial 1% optical photometry. Astronomical Journal, 134(3):973 998.

Kaspi, S., Smith, P. S., Netzer, H., Maoz, D., Jannuzi, B. T., and Giveon, U. (2000). Reverberation measurements for 17 quasars and the size mass luminosity relations in active galactic nuclei. The Astrophysical Journal, 533(2):631 649.

Kau mann, G., Heckman, T. M., Tremonti, C., Brinchmann, J., Charlot, S., White, S. D. M., Ridgway, S. E., Brinkmann, J., Fukugita, M., Hall, P. B., Ivezi¢, š., Richards, G. T., and Schneider, D. P. (2003). The host galaxies of active galactic nuclei. Monthly Notices of the Royal Astronomical Society, 346(4):1055 1077.

Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., and Trevena, J. (2001). Theoretical modeling of starburst galaxies. The Astrophysical Journal, 556(1):121 140.

Kollatschny, W., Ulbrich, K., Zetzl, M., Kaspi, S., and Haas, M. (2014). Broad-line region structure and kinematics in the radio galaxy 3c 120. Astronomy and Astrophysics, 566:A2.

Komossa, S. (2008). Narrow-line seyfert 1 galaxies. In Revista Mexicana de Astronomía y Astrofísica Conference Series, volume 32, pages 86 92.

Kormendy, J. and Ho, L. C. (2013). Coevolution (or not) of supermassive black holes and host galaxies. Annual Review of Astronomy and Astrophysics, 51:511 653.

Kova£evi¢, J., Popovi¢, L. ƒ., and Dimitrijevi¢, M. S. (2010). Analysis of optical spectra of agn. i. fe ii emission in type 1 agn. Astrophysical Journal Supplement Series, 189:15 36.

Krolik, J. H. (1999). Active Galactic Nuclei: From the Central Black Hole to the Galactic Environment. Princeton University Press, Princeton, NJ.

Krolik, J. H. and Begelman, M. C. (1988). Molecular tori in seyfert galaxies: Feeding the monster and hiding it. The Astrophysical Journal, 329:702 711.

Kuhn, L. et al. (2020). Broad-line region geometry from multiple emission lines in a single-epoch spectrum. Monthly Notices of the Royal Astronomical Society, 498:3505.

Marrese, P. M., Marinoni, S., Fabrizio, M., and Altavilla, G. (2017). Cross-match between gaia and external catalogues: algorithm and statistics. Astronomy & Astrophysics, 607:A105.

Marrese, P. M., Marinoni, S., Fabrizio, M., and Altavilla, G. (2019). Gaia dr2 cross-match with external catalogues: algorithms and results. Astronomy & Astrophysics, 621:A144.

Montegri o, P., Andrae, R., De Angeli, F., Weiler, M., Evans, D. W., et al. (2023). Gaia data release 3: External calibration of bp/rp low-resolution spectrophotometric data. Astronomy & Astrophysics, 674:A3.

Netzer, H. (2013). The Physics and Evolution of Active Galactic Nuclei. Cambridge University Press.

Netzer, H. (2015). Revisiting the uni ed model of active galactic nuclei. Annual Review of Astronomy and Astrophysics, 53:365 408.

Osterbrock, D. E. and Ferland, G. J. (2006). Astrophysics of Gaseous Nebulae and Active Galactic Nuclei. University Science Books, Sausalito, CA, 2nd edition.

Osterbrock, D. E. and Pogge, R. W. (1985). The spectra of narrow-line seyfert 1 galaxies. The Astrophysical Journal, 297:166 176.

Padovani, P., Alexander, D. M., Assef, R. J., et al. (2017). Active galactic nuclei: what's in a name? Astronomy and Astrophysics Review, 25:2.

Peterson, B. M. (1993). Reverberation mapping of active galactic nuclei. Publications of the Astronomical Society of the Paci c, 105:247 268.

Peterson, B. M. (2001). Variability of active galactic nuclei. In Aretxaga, I., Kunth, D., and Mújica, R., editors, Advanced Lectures on the Starburst AGN Connection, volume 254 of ASP Conference Series, pages 3 67. Astronomical Society of the Paci c.

Peterson, B. M. (2004). Measuring the masses of supermassive black holes. Annual Review of Astronomy and Astrophysics, 42:553 601.

Planck Collaboration (2020). Planck 2018 results. vi. cosmological parameters. A&A , 641:A6.

Popovi¢, L. ƒ. (2012). Broad line region in active galactic nuclei: Physics and kinematics. New Astronomy Reviews, 56(2-3):74 91.

Popovi¢, L. (2008). The broad-line region of agn: Kinematics and physics. Serbian Astronomical Journal, 176:1 11.

Porqueres, N., Jasche, J., Enÿlin, T. A., and Lavaux, G. (2018). Imprints of the large-scale structure on agn formation and evolution. Astronomy & Astrophysics, 612:A33.

Raki¢, N. (2022). Broad-line region geometry and kinematics of type-1 agns with high-quality sdss spectra. Monthly Notices of the Royal Astronomical Society, 516(2):1624 1634.

Rakshit, S., Stalin, C. S., Chand, H., and Zhang, X. (2017). A catalog of quasar properties from sloan digital sky survey data release 12. The Astrophysical Journal Supplement Series, 229(2):39.

Rees, M. J. (1984). Black hole models for active galactic nuclei. Annual Review of Astronomy and Astrophysics, 22:471 506.

Richards, G. T., Fan, X., Newberg, H. J., Strauss, M. A., Vanden Berk, D. E., Schneider, D. P., Yanny, B., Boucher, A., Burles, S., Frieman, J. A., et al. (2002). Spectroscopic target selection in the sloan digital sky survey: The quasar sample. The Astronomical Journal, 123(5):2945 2975.

SDSS Collaboration, Pallathadka, G. A., et al. (2025). The Nineteenth Data Release of the Sloan Digital Sky Survey (SDSS DR19). arXiv e-prints. Accepted for publication, SDSS-V data release encompassing the rst full datasets from Black Hole Mapper, Milky Way Mapper, and Local Volume Mapper.

Sesar, B., Ivezi¢, š., Lupton, R. H., Juri¢, M., Gunn, J. E., Knapp, G. R., DeLee, N., Smith, J. A., Miknaitis, G., Lin, H., et al. (2007). Exploring the variable sky with the sloan digital sky survey. Astronomical Journal, 134(6):2236 2251.

Shapovalova, A. I., Popovi¢, L. ƒ., Burenkov, A. N., Chavushyan, V. H., Ili¢, D., Kova£evi¢, J., Bochkarev, N. G., and Leon-Tavares, J. (2012). Spectral optical monitoring of 3c 390.3 in 1995 2007. ii. variability of the broad emission lines. Astrophysical Journal Supplement Series, 202:10.

Shen, Y., Brandt, W. N., Dawson, K. S., Hall, P. B., McGreer, I. D., Schneider, D. P., Treu, T., Trump, J. R., Wood-Vasey, W. M., Zu, Y., and et al. (2015). The sloan digital sky survey reverberation mapping project: Technical overview. The Astrophysical Journal Supplement Series, 216(1):4.

Shen, Y., Richards, G. T., Strauss, M. A., Hall, P. B., Schneider, D. P., Snedden, S. A., Bizyaev, D., Brewington, H., Malanushenko, E., Malanushenko, V., et al. (2011). A catalog of quasar properties from sloan digital sky survey data release 7. The Astrophysical Journal Supplement Series, 194(2):45.

Smee, S. A. et al. (2013). The multi-object, ber-fed spectrographs for the sloan digital sky survey and the baryon oscillation spectroscopic survey. The Astronomical Journal, 146(2):32.

Storey-Fisher, K., Hogg, D. W., Rix, H.-W., Eilers, A.-C., Fabbian, G., Blanton, M. R., and Alonso, D. (2024). Quaia, the gaia-unwise quasar catalog: An all-sky spectroscopic quasar sample. The Astrophysical Journal, 964(1):69.

Urry, C. M. and Padovani, P. (1995). Uni ed schemes for radio-loud active galactic nuclei. Publications of the Astronomical Society of the Paci c, 107:803 845.

Vallenari, A., Brown, A. G. A., Prusti, T., de Bruijne, J. H. J., Arenou, F., Babusiaux, C., and et al. (2023). Gaia data release 3: Summary of the content and survey properties. Astronomy & Astrophysics, 674:A1.

Vanden Berk, D. E., Shen, J., Yip, C. W., Schneider, D. P., Connolly, A. J., Burton, R. E., and et al. (2006). Spectral variability of quasars in the sloan digital sky survey.Astronomical Journal, 131:84 99.

Véron-Cetty, M.-P., Joly, M., and Véron, P. (2004). A spectrophotometric atlas of narrow-line seyfert 1 galaxies.Astronomy & Astrophysics, 417:515 533.

Vestergaard, M. and Peterson, B. M. (2006). Determining central black hole masses in distant active galaxies and quasars. ii. improved optical and uv scaling relationships.The Astrophysical Journal, 641(2):689 709.

Yip, C. W., Connolly, A. J., Szalay, A. S., and et al. (2004a). Distributions of galaxy spectral types in the sloan digital sky survey. Astronomical Journal, 128:585 609.

Yip, C. W., Connolly, A. J., Vanden Berk, D. E., Ma, Z., Frieman, J. A., SubbaRao, M., and et al. (2004b). Spectral classi cation of quasars in the sloan digital sky survey: Eigenvectors. Astronomical Journal, 128:2603 2630.

York, D. G., Adelman, J., Anderson, J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Barkhouser, R., Bastian, S., Berman, E., et al. (2000). The sloan digital sky survey: Technical summary. Astronomical Journal, 120:1579 1587.

Appendix A: read_gaia Loader Function

The read_gaia function

The following Python class was developed to load and prepare Gaia XP spectra for direct use with the FANTASY fitting pipeline. It reads flux-calibrated spectra from CSV files produced after GaiaXPy calibration, attaches positional and redshift information from SDSS crossmatches, converts fluxes from SI units to the SDSS-style cgs scale, and defines the wavelength grid in vacuum Ångström before any optional conversion to air wavelengths in the fitting stage.

```
import pandas as pd
import numpy as np
import re
import ast

class read_gaia:
    def __init__(self, filename, source_id=None, source_index=None):
        """
        Reads Gaia XP spectra data for a specific source_id or row index.

        Parameters
        ----------
        filename : str
            Path to the CSV file containing Gaia XP spectra data.
        source_id : int, optional
            The Gaia source ID for which data should be loaded.
        source_index : int, optional
            The row index for the source to load (e.g., 5 for the 5th source).
        """
        self.filename = filename
        self.source_id = source_id
        self.source_index = source_index

        # Load the CSV file containing Gaia data
        df = pd.read_csv(self.filename)

        # Convert flux and flux_error columns to lists
        def safe_list_conversion(value):
            try:
                if isinstance(value, str):
                    value = re.sub(r'(?<=\d)\s+(?=\d)', ',', value.strip())
```

```
                return ast.literal_eval(value)
            return value
        except (SyntaxError, ValueError):
            return None

df["flux"] = df["flux"].apply(safe_list_conversion)
df["flux_error"] = df["flux_error"].apply(safe_list_conversion)

# Drop invalid entries
df = df.dropna()
if df.empty:
    raise ValueError("No valid spectra remaining after cleaning.")

# Load data by source_id or source_index
if self.source_id is not None:
    df_source = df[df["source_id"] == self.source_id]
    if df_source.empty:
        raise ValueError(f"Source ID {self.source_id} not found.")
    row = df_source.iloc[0]  # First matching row
elif self.source_index is not None:
    if self.source_index < 0 or self.source_index >= len(df):
        raise ValueError(f"Invalid source index {self.source_index}.")
    row = df.iloc[self.source_index]
else:
    row = df.iloc[0]

# Set attributes for the selected source
self.source_id = row["source_id"]
self.ra = row["ra"]
self.dec = row["dec"]
self.z = row["redshift"]

# Convert from W/m^2/nm to erg/s/cm^2/Å
self.flux = np.array(row["flux"]) * 100
self.err = np.array(row["flux_error"]) * 100

# Normalize to SDSS-style scale (10^-17 erg/s/cm^2/Å)
self.flux *= 1e17
self.err *= 1e17

# Gaia XP wavelength grid: 336 1020 nm     Å
self.wave = np.linspace(336, 1020, len(self.flux)) * 10
```

```
# Extract original_external_source_id as bestobjid
self.bestobjid = row["original_ext_source_id"]

# Name object
self.name = f"Gaia_{self.source_id}"

print(f"Data loaded for source ID {self.source_id}")
print(f"RA: {self.ra}, Dec: {self.dec}, Redshift: {self.z}")
print(f"Original External Source ID (bestobjid): {self.bestobjid}")
```