Tracing the Evolution of Super-Massive Black Holes through Cosmic Time with Luminous Active Galactic Nuclei

Julien Wolf

SIGILL.UNIVERSITAT.LUDOVIC.MAXIM.

München 2023

Tracing the Evolution of Super-Massive Black Holes through Cosmic Time with Luminous Active Galactic Nuclei

Julien Wolf

Dissertation
an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Julien Wolf
aus Colombes (Frankreich)

München, den 17.03.2023

Erstgutachter: Prof. Dr. Kirpal Nandra
Zweitgutachter: Prof. Dr. Jochen Weller
Tag der mündlichen Prüfung: 09.05.2023

Contents

List of Figures

List of Tables

List of Abbreviations

1(2)RXS	First (Second) ROSAT All-Sky Survey
4DE1	Four-Dimensional Eigenvector 1
ACIS	Advanced CCD Imaging Spectrometer
ADAF	Advection-Dominated Accretion Flow
AGN	Active Galactic Nuclei
ARF	Ancillary Response File
ART-XC	Mikhail Pavlinsky Astronomical Roentgen Telescope X-ray Concentrator
ASKAP	Australian Square Kilometre Array Pathfinder
BLR	Broad Line Region
BOSS	Baryon Oscillation Spectroscopic Survey
CCD	Charge-Coupled-Device
ChAMP	Chandra Multi-Wavelength Project
CFHQS	Canada–France High-z Quasar Survey
CFHQSIR	CFHQS in the Near Infrared
CMB	Cosmic Microwave Background
CXB	Cosmic X-ray Background
COSMOS	Cosmic Evolution Survey
DECam	Dark Energy Cam
DES	Dark Energy Survey
DESI	Dark Energy Spectroscopic Instrument
eBOSS	extended Baryon Oscillation Spectroscopic Survey
eFEDS	eROSITA Final Equatorial Depth Survey
EHT	Event Horizon Telescope
eROSITA	extended ROentgen Survey with an Imaging Telescope Array
eROSITA-DE(RU)	German (Russian) eROSITA consortium
EV1	Eigenvector 1
ExSeSS	Extragalactic Serendipitous Swift Survey
FDPL	Flexible Double Power-Law
FIRE	Folded Port Infrared Echelette
FWHM	Full-Width at Half-Maximum
GAMA	Galaxy And Mass Assembly
GLM	Generalised Linear Model
HEW	Half-Energy Wifth

HSC	Hyper Suprime Cam
iC-CMB	Inverse Compton-Scattering of CMB photons
IGM	Intergalactic Medium
IR	Infrared
JVLA	Karl G. Jansky Very Large Array
JWST	James Webb Space Telescope
KDE	Kernel Density Estimate
KiDS	Kilo Degree Survey
LAMOST	Large sky Area Multi-Object fiber Spectroscopic Telescope
LDDE	Luminosity-Dependent Density Evolution
LDSS-3	Low Dispersion Survey Spectrograph
LOFAR	LOw-Frequency Array
LS8	DESI Legacy Imaging Surveys DR8
MIR	Mid-Infrared
NIR	Near-Infrared
NLR	Narrow Line Region
NLS1	Narrow-Line Seyfert 1 Galaxy
Pan-STARR	Panoramic Survey Telescope
PCA	Principal Components Analysis
PDF	Probability Density Function
PSF	Point-Spread Function
PDR	Pure Density Evolution
RM	Reverberation Mapping
RMF	Response Matrix File
ROSAT	Röntgensatellit
SDSS	Sloan Digital Sky Survey
SED	Spectral Energy Distribution
SHELLQs	Subaru High-z Exploration of Low-Luminosity Quasars
S/N	Signal-To-Noise Ratio
SMBH	Super-Massive Black Hole
SPIDERS	SPectroscopic IDentifcation of ERosita Sources
SWAG-X	Survey With ASKAP of GAMA-09 + X-Ray
UKIRT	United Kingdom Infrared Telescope
UKIDSS	UKIRT Infrared Deep Sky Survey
UVOIR	Ultraviolet Optical Infrared
UV	Ultraviolet
VISTA	Visible and Infrared Survey Telescope for Astronomy
VHS	VISTA Hemisphere Survey
VIKING	VISTA Kilo-degree Infrared Galaxy
VLT(I)	Very Large Telescope (Interferometer)
WISE	Wide-field Infrared Survey Explorer
XLF	X-ray Luminosity Function
XMM-Newton	X-ray Multi-Mirror Mission

(2)4 XMM	XMM-Newton Serendipitous Survey
XMMSL1	XMM-Newton Slew 1 Survey

List of Constants, Units and Recurrent Symbols

Constants & Units

Symbol	Definition	Value (SI units)
c	Speed of light in vacuum	3.0×10^{8} m/s
G	Gravitational constant	6.7×10^{-11} $\mathrm{N} \cdot \mathrm{m}^2/\mathrm{kg}^2$
m_{e}	Electron mass	9.1×10^{-31} kg
m_{p}	Proton mass	1.7×10^{-27} kg
σ_{T}	Thomson cross-section for the electron	6.7×10^{-29} m^2
$M_{\odot}$	Solar mass	2.0×10^{30} kg
$L_{\odot}$	Solar luminosity	3.8×10^{26} J/s
$Z_{\odot}$	Solar metallicity	0.0134
$\mathring{A}$	Angstrom	10^{-10} m
pc	parsec	3.1×10^{16} m
yr	year	3.2×10^{7} s

Recurrent symbols

Symbol	Definition
r_s	Schwarzschild radius
L_{Edd}	Eddington luminosity
λ_{Edd}	Eddington ratio
M_{BH}	Black hole mass
L_{bol}	Bolometric luminosity
$\dot{M}$	Mass accretion rate
ϵ	Radiative efficiency

Zusammenfassung

Schwarze Löcher sind zentrale Bausteine der großskaligen baryonischen Struktur im Universum. Ihre Massen reichen von einigen wenigen bis hin zu einigen Milliarden Sonnenmassen. Supermassereiche schwarze Löcher (engl. supermassive black holes, kurz SMBHs), die größte Form kompakter Objekte, werden in den Kernen aller Galaxien vermutet. Korrelationen zwischen den Eigenschaften von Host-Galaxien und SMBHs deuten auf eine enge Koevolution zwischen Galaxien und ihrer zentralen schwarzen Löcher hin. Im tiefen Gravitationspotential akkretiert das SMBH Materie. Unter gewissen Bedingungen, setzt die visköse Strömung in der Akkretionsscheibe extreme Energiemengen in Form von Strahlung frei. Galaktische Zentren, die durch den Akkretionseffekt ihre Umgebung bestrahlen, werden aktive Galaxienkerne genannt (engl. active galactic nuclei, kurz AGN). In den letzten 20 Jahren wurden Quasare, die leuchtkräftigsten AGN, zu zunehmend höheren Distanzen entdeckt. Der derzeitig entfernteste Quasar befindet sich bei einer Rotverschiebung von $z \sim 7.6$. Unter der Annahme einer ΛCDM Standardkosmologie entspricht dies einer Distanz von umgerechnet 13 Milliarden Lichtjahren. Als diese Quasare ihr Licht entsannten, war das Universum etwa 700 Millionen Jahre alt. Einige Quasare in diesem frühen Universum beinhalten SMBHs deren Virialmassen Milliarden von Sonnenmassen überschreiten. Die Entdeckung von $> 10^9\,M_\odot$ SMBHs bei $z > 7$ stellt eine Herausforderung für die frühe Entstehungstheorie massiver schwarzer Löcher dar, da Stamm- und Entwicklungsmodelle das Wachstum dieser extremen Objekte über eine sehr kurze Zeitspanne gewährleisten müssen. Die Entstehung und Entwicklung der ersten SMBHs ist eine fundamentale, offene Frage der modernen Astronomie. Um die Wachstumshistorie von SMBH zu verstehen ist ein präziser Zensus und eine Charakterisierung entfernter Quasare (und AGN) als Proxypopulation essentiell. In der nachfolgenden Dissertation werden die zeitabhängigen Eigenschaften und die Demografie leuchtkräftiger AGN untersucht. Der Fokus liegt dabei auf zwei zentralen, beobachtbaren Größen: SMBH-Massen und die AGN-Raumdichte. Diese werden auf zwei distinkte Raumskalen untersucht: Der zentrale Parsec von Galaxien und mitbewegte kosmologische Volumen, die sich über mehrere Kubikmegaparsecs erstrecken.

Die Studien, die in dieser Arbeit präsentiert werden, basieren auf Röntgen- und Multiwellenlängenzusatzbeobachtungen. Um AGN in Röntgenhimmelsdurchmusterungen zu finden und ihre Rotverschiebung zu determinieren, müssen die Multiwellenlängengegenstücke der Röntgenquellen identifiziert werden. Als Teil dieses PhD-Projektes wurde eine Erweiterung basierend auf maschinellem Lernen für den bayesianischen Katalogskreuzassoziierungsalgorithmus `NWAY` entwickelt (Kapitel 2). Diese Methodologie wurde angewandt,

um die Quellenkataloge des Röntgenweltraumteleskops extended ROentgen Survey with an Imaging Telescope Array (eROSITA) mit Beobachtungen im optischen und infraroten Wellenlängenbereich zu assoziieren.

Zuerst wurde die Kinematik der sog. Broad Line Region (kurz BLR) über eine statistische Analyse der Formparameter breiter Emissionslinien in den optischen Spektren Röntgenleuchtkräftiger AGN untersucht (Kapitel 3). Die Charakterisierung verschobener, breiter Komponenten, die in AGN Spektra durch nicht-Keplercher Kinematik entstehen, ist wichtig, da dies einen signifikanten Effekt auf die Genauigkeit virialer Massenmessungen schwarzer Löcher hat. Die Spektren wurden im Kontext des Sloan Digital Sky Survey-IV/Spectral Identification of eROSITA Sources (SDSS-IV/SPIDERS) Programms erhoben. Die spektrale Varianz von AGN wurde über eine Hauptkomponentenanalyse exploriert. Im Kontext der sog. Quasarenhauptsequenz wird gezeigt, wie die Formen der breiten Emissionslinien mit globalen Parametern des schwarzen Lochs zusammenhängen. Die Asymmetrie der breiten Hβ Linie ist ein treibender Parameter der Sample-Varianz. Ein zentrales Ergebnis ist das blau-asymmetrische breite Balmer-Emissionslinien, die auströmendes BLR-Gas signalisieren, sich bevorzugt am schnell akkretierenden Ende der Quasaranhauptsequenz befinden. Des Weiteren werden Hinweise für eine geschichtete, selbstabschirmende Struktur der BLR aufgedeckt.

Um die Demografie entfernter AGN zu charakterisieren, wurde die eROSITA Final Equatorial Depth Survey (eFEDS) Himmelsdurchmusterung nach hochrotverschobenen Quasaren durchsucht. Es konnte die Röntgendetektion eines leuchtkräftigen $z = 5.81$ SDSS Quasaren bestätigt werden (Kapitel 4). Diese Detektion im derzeit größten, nahuniformen und zusammenhängenden Röntgenfeld ermöglichte die Herleitung der ersten spektroskopischen Beschränkung der AGN Röntgenleuchtkraftsfunktion (engl. X-ray Luminosity Function, kurz XLF) bei Rotverschiebung $z \sim 6$. Unsere Analyse unterstützt eine flache Steigung der $L_{\mathrm{X}} > 10^{45}\,\mathrm{erg\,s^{-1}}$ und $z \sim 6$ XLF. Dies steht im Kontrast zu Modellextrapolationen aus der Literatur, die eine niedrigere Anzahl leuchtkräftiger AGN im jungen Universum vorhersagen. Nach Senkung der Quellendetektionswahrscheinlichkeit wurde eine ähnliche Suche im Feld wiederholt. Dies ermöglichte die unsichere eROSITA Detektion eines weiteren Quasaren bei $z = 6.56$ und führte zu dessen subsequenten Bestätigung mit einer Chandra-Nachbeobachtung. Es ist derzeit die entfernteste blinde Röntgendetektion. Das Nahinfrarotspektrum dieser Quelle weist typische Signaturen eines stark-akkretierenden SMBH mit geringer Masse auf ($\sim 10^8 M_\odot$). Die Eigenschaften dieses Quasars ähneln denen von englinigen Seyfertgalaxien 1 (eng. Narrow-Line Seyfert 1, kurz NLS1), die üblicherweise mit jungen schwarzen Löchern in Verbindung gesetzt werden. Die Detektion der zwei distanten eFEDS Quasaren wurde genutzt, um die Akkretionsratendichte schwarzer Löcher bei $z \sim 6$ zu beschränken. Dieses Resultat kann dann mit Vorhersagen theoretischer Entstehungs- und Entwicklungsmodelle verglichen werden. Die eFEDS Daten sind konsistent mit Modellen, in denen SMBH-Wachstum nur in den massivsten Dunkle Materie Halos stattfindet ($> 5 \times 10^{11}\, M_\odot$). Das tiefe Gravitationspotential dieser Halos verhindert Gaserschöpfung durch Supernovaerückkoppelungseffekte.

Weiterhin werden erste Resultate einer Pilotstudie zur Suche unbekannter, seltener Quasare bei $z > 5.7$ in der eROSITA Ganzhimmelsdurchmusterung (engl. eROSITA All-

Sky Survey, kurz eRASS) präsentiert (Kapitel 6). Die Entdeckung solch extrem seltener, entfernter, leuchkräftiger Quasare demonstriert das Potential großflächiger Röntgenhimmelsdurchmusterungen, wie eRASS, die Akkretionshistorie schwarzer Löcher bis in das frühe Universum nachzuverfolgen.

Die Resultate der nachfolgenden Dissertation entstammen einem PhD-Projekt unter der Leitung von Prof. Dr. Kirpal Nandra und Dr. Mara Salvato am Max-Planck-Institut für extraterrestrische Physik.

Abstract

Black holes are an essential building block of the baryonic structure in the Universe. Their masses range from a few times to a few billion times the mass of our Sun. Super-massive black holes (SMBHs), the largest form of compact objects, are believed to reside at the centre of all galaxies. Scaling relations between host properties and SMBH masses hint at a tight co-evolution between galaxies and their central black holes. In the deep nuclear gravitational potential, matter is accreted onto the SMBH. Under certain conditions, the viscous flow in the accretion disk releases extreme amounts of energy in the form of radiation. The galactic cores irradiating their environment under the effect of accretion are called active galactic nuclei (AGN). Over the last 20 years, quasars, the most luminous sub-species of AGN, have been discovered at ever-increasing distances. The current quasar redshift frontier is at $z \sim 7.6$, corresponding to a distance of 13 billion light-years assuming a standard ΛCDM cosmological model. When these quasars emitted their light, the Universe was still in its infancy, about 700 million years old. Some of the quasars at these cosmic times were found to host SMBHs whose virial masses exceed billions of solar masses. The discovery $> 10^9\,M_\odot$ SMBHs at $z > 7$ challenges the early formation theory of massive black holes, as seed and growth models are required to support the assembly of these extreme objects on short timescales. The formation and evolution channels of the first SMBHs remain a fundamental open question in modern astronomy. In order to understand the SMBH growth history, it is essential to obtain an accurate census and characterisation of quasars (and AGN at large) as proxy population. This thesis investigates the properties and demographics of the most X-ray luminous AGN across cosmic time. The focus lies on two central observable quantities: SMBH masses and the space density of AGN. These are investigated on two distinct spatial scales: the central parsec of galaxies and comoving cosmological volumes extending over many cubic megaparsecs.

The studies presented in this work are based on X-ray and ancillary multi-wavelength observations. Finding AGN in X-ray surveys and associating them to a redshift requires the identification of multi-wavelength counterparts to X-ray sources. As part of this PhD project a machine learning-based extension for the Bayesian catalogue cross-matching framework `NWAY` was developed (Chapter 2). This methodology was adopted to obtain optical and near-infrared counterparts to the source catalogues of the extended ROentgen Survey with an Imaging Telescope Array (eROSITA).

We first investigated the kinematics of the Broad Line Region (BLR) via a statistical analysis of the shape parameters of emission lines in the optical spectra of X-ray luminous

AGN (Chapter 3). It is critical to characterise broad and shifted spectral components imprinted on AGN spectra by non-Keplerian BLR kinematics as they have a significant effect on the accuracy of black hole virial mass estimates. The spectra were collected in the context of the Sloan Digital Sky Survey-IV/Spectral Identification of eROSITA Sources (SDSS-IV/SPIDERS) programme. The spectral variance of AGN is explored by performing a principal component analysis. We show how the broad line asymmetries are related to the global black hole parameters in the context of the quasar main sequence. The broad Hβ line asymmetry are a driving parameter of the total sample variance; a marker of diversity in Type 1 AGN. A core result of this work is that blue-asymmetric broad Balmer emission line profiles, a signature of outflowing BLR gas, preferentially populate the rapidly accreting end of the quasar main sequence. We further provide evidence for a layered, self-shielding structure of the BLR.

In order to characterise the demographics of distant AGN, we searched for high-redshift quasars in the eROSITA Final Equatorial Depth Survey (eFEDS) and reported the X-ray detection of a luminous $z = 5.81$ SDSS quasar (Chapter 4). This detection in the largest contiguous, near-uniform X-ray survey to date enables the formulation of the first spectroscopic-grade constraints on the bright end of the X-ray luminosity function (XLF) of AGN at $z \sim 6$. Our analysis supports a shallow slope on the $L_{\rm X} > 10^{45}\,{\rm erg\,s^{-1}}$ and $z \sim 6$ XLF, in contrast to models from literature extrapolated from lower redshift ranges. After decreasing the detection likelihood threshold, a similar search was performed in the field. A tentative eROSITA detection of a second quasar at $z = 6.56$ and its confirmation via a Chandra follow-up observation was subsequently reported (Chapter 5). It is, to date, the most distant serendipitous X-ray detection. Its near-infrared spectrum shows typical signatures of a low-mass ($\sim 10^8 M_\odot$), rapidly accreting black hole. The quasar's properties are remarkably similar to that of narrow-line Seyfert 1 galaxies, which are usually associated with SMBHs caught in an early evolutionary stage. The detection of the two distant eFEDS quasars is used to derive constraints on the black hole accretion rate density at $z \sim 6$, which can be compared to theoretical predictions from various seed and growth models. The eFEDS data is consistent with models that limit SMBH growth to the most massive dark matter halos ($> 5 \times 10^{11}\, M_\odot$), in which the gravitational potential suppresses gas depletion through supernova feedback.

Finally, early results of a pilot study designed to uncover yet unknown, rare quasars at $z > 5.7$ in the eROSITA All Sky-Survey (eRASS) are presented (Chapter 6). The discovery of extremely rare, luminous high-redshift quasars demonstrates the potential of wide X-ray surveys such as eRASS to track the black hole accretion history at cosmic dawn.

The following thesis reports the results of a PhD project supervised at the Max-Planck-Institute for Extraterrestrial Physics by Prof. Dr. Kirpal Nandra and Dr. Mara Salvato.

Chapter 1

Active Galactic Nuclei from the local Universe to the first billion years after the Big Bang

1.1 From spacetime singularities to astrophysical black holes

Black holes shine an unsettling light on the fabric of spacetime itself. In our collective imagination, they are star-devouring behemoths lurking in the confines of the Universe. Over the last century, theorists and observers have gradually uncovered a natural phenomenology that is more nuanced, complex and sometimes counter-intuitive. A good starting point and simple thought experiment is to consider objects that can entrap light in Newtonian mechanics. With conservation of energy, the escape velocity from the surface of a spherically symmetric body of mass M exceeds the speed of light c if its radius is larger than:

$$R = \frac{2GM}{c^2}, \tag{1.1}$$

where G is the gravitational constant. Black holes have been a prediction from theory long before entering their astrophysical era. Rather than celestial objects, black holes should be understood as regions of spacetime where gravitation is such that it secludes contained information of any form from the Universe. Solving the field equations of the theory of general relativity (Einstein 1916) for a spherically symmetric mass and using the weak-field approximation, one obtains a spacetime metric:

$$ds^2 = -\left(1 - \frac{2G\,M}{c^2\,r}\right) c^2 dt^2 + \frac{dr^2}{1 - 2G\,M/c^2 r} + r^2 d\theta^2 + r^2 sin^2\theta d\phi^2 \tag{1.2}$$

Here the metric components are along the spherical coordinates (r, ϕ, θ) and the time dimension t. This exact solution to Einstein's field equations is called the Schwarzschild metric (Schwarzschild 1916). Setting the radial coordinate to:

$$r_s = \frac{2G\,M}{c^2} \tag{1.3}$$

leads to a coordinate singularity and sets the radius of the event horizon in the case of a non-rotating black hole. Coincidentally, r_s is equal to the radius of a Newtonian, light-confining body. The Schwarzschild metric becomes singular at $r = 0$, causing a real spacetime singularity. The physical interpretation of singularities in general relativity has been intensively debated. Formally, the $r = 0$ singularity cannot be circumvented by coordinate transformations (e.g., Stoica 2018, and references therein). It is widely believed that complementing general relativity with quantum theory would close this loophole in the theory of black holes (e.g., Bonanno & Reuter 2000; Ashtekar & Bojowald 2005; Frolov 2014; Gambini & Pullin 2014; Haggard & Rovelli 2015).

This geometry leads to a complex causal structure and a collection of somewhat surprising physical effects, which can qualitatively be explored in a Penrose-Carter diagram (Penrose 1963; Carter 1966, see Fig. 1.1). The conformal diagram for a spherically symmetric geometry shown here describes causal relations in our Universe, within a black hole and between these two distinct regions. In this diagram, light emitted in the very distant Universe reaches us in the very distant future on trajectories maintaining a 45° angle in every point: the light cone is preserved (see description in caption of Fig. 1.1). The lines defining constant times and positions are exchanged in the black hole. As light cones are preserved, any particle crossing the event horizon cannot return in our Universe and will inexorably follow a trajectory to the $r = 0$ singularity. Indeed the event horizon has been interpreted as a *unidirectional* membrane restricting the permitted causal interactions to the inward direction (Finkelstein 1958). Events triggered within the black hole are entirely secluded from the Universe. To the observer in our Universe, they never happened and will never happen.

Suppose we observe a particle falling onto a black hole. In that case, any light emitted by the particle at the exact position of the horizon can only reach our Universe at infinite times, i.e. the upper corner of the region I. This signal is lost to the Universe. Approaching the horizon, the light emitted from the particle would have to travel along increasingly stretched patches of spacetime (represented in the diagram by the compacting grid structure). To the observer, it would appear to progressively slow down until halting completely at the horizon.

The derivation of the Schwarzschild metric was followed by solutions for static and electrically charged black holes (Reissner 1916; Nordström 1918), rotating and uncharged black holes (Kerr 1963), as well as for rotating and charged black holes (Newman et al. 1965). From the investigation of these solutions, it emerged that only three parameters govern black holes: their mass, their spin angular momentum and their charge (no-hair theorem Israel 1967; Carter 1971; Robinson 1975, for a recent review see Bambi 2020). Astrophysical black holes are believed to be born from stellar collapse or even direct gravitational instabilities. The Kerr solution is, therefore, of particular interest from an astrophysical point of view as black holes born from the collapse of rotating stars are expected to have non-vanishing angular momentum. Additionally, any residual charge is expected to be

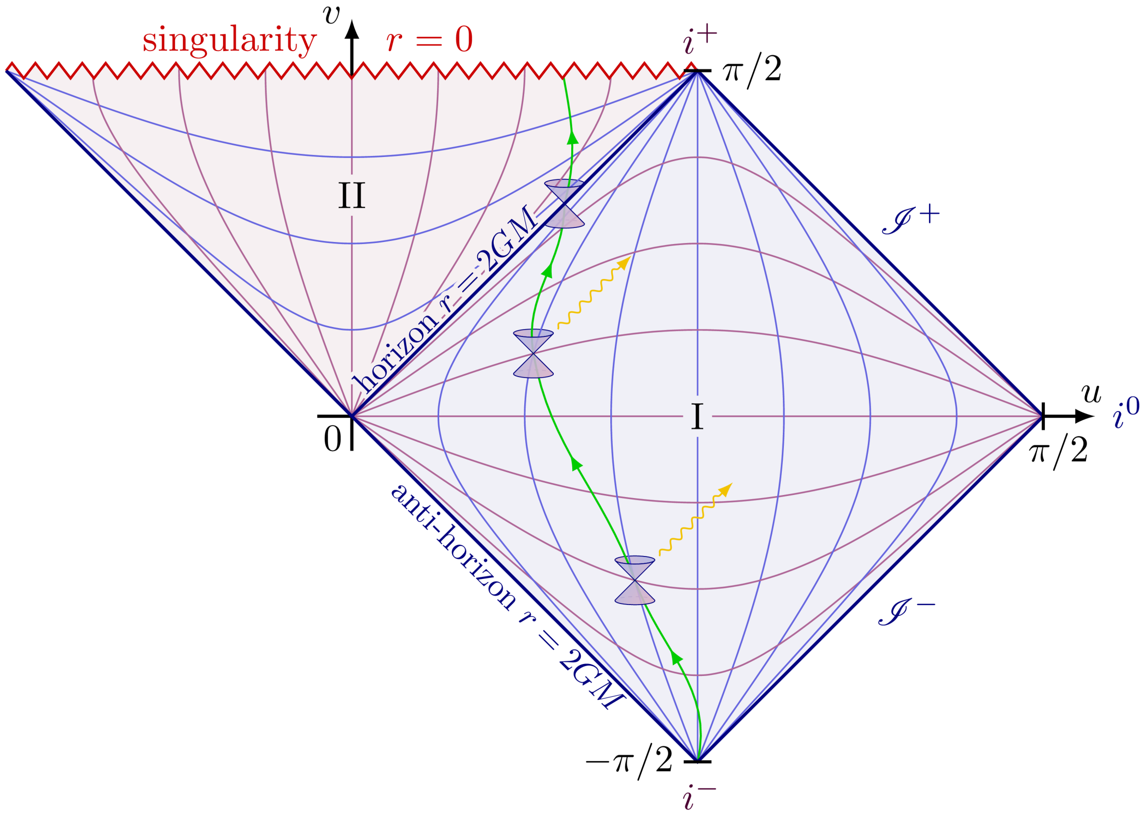


Figure 1.1: Penrose diagram for a Schwarzschild black hole. In this conformal diagram, two regions are defined: our Universe (I) and the inside of a black hole (II). Corners in I define time-like infinities (i^+ and i^-) and distance-like infinity (i^0). The black hole is at the origin of the coordinate system. Red and blue curves denote fixed moments in time and locations in space respectively. The event horizon delimits I and II. The spacetime singularity at $r = 0$ is shown as spacelike boundary. The trajectory of a particle through our Universe and crossing the event horizon is shown in green with individual light cones. This representation is based on a conformal transformation: i.e. light rays always propagate at an angle of +45° (the shape and orientation of lightcones are preserved at any point of the spacetime map). Figure reproduced using the TiKZ package (`https://github.com/pgf-tikz/pgf`)

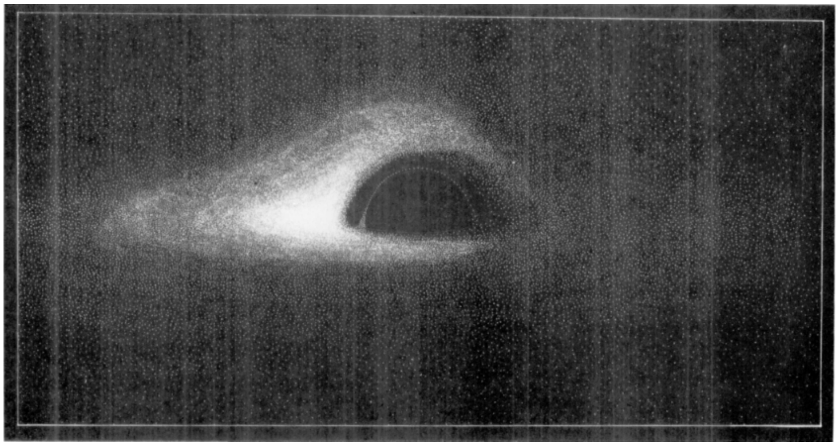


Figure 1.2: The appearance of an idealised black hole by Luminet (1979).

neutralised in the radiation field near the surface.

Astronomers have searched for the signatures of extreme physical conditions expected to be caused by black holes. Cosmic X-ray emission is associated with highly energetic astrophysical phenomena. Cygnus X-1, one of the brightest (hard) X-ray sources in the sky (Bowyer et al. 1965), was identified as a spectroscopic binary system by Webster & Murdin (1972) and Bolton (1972). Measuring the motion of the stellar companion revealed a compact object in the binary system to be a stellar-mass black hole. Some of the most substantial evidence for black holes came directly from our cosmic neighbourhood: measuring stellar proper motions in infrared wavebands of the Sgr A* nuclear star cluster of our galaxy, two teams of astronomers could simultaneously demonstrate the presence of a *dark mass* of $2.6 \times 10^6\,\mathrm{M}\odot$ at its core (Genzel et al. 1997; Ghez et al. 1998). This result and subsequent long monitoring campaigns by both collaborations were recently rewarded with the 2020 Nobel prize in physics for Reinhard Genzel and Andrea Ghez. The consortia of the gravitational wave detectors LIGO and VIRGO achieved another breakthrough when they measured a signal related to a binary black hole merger (Abbott et al. 2016). While these results are unmistakably associated with black hole activity, astronomers had yet to deliver a direct visualisation of a black hole. In 1979, Luminet presented a rendering of the distorted light of an accretion disk warping around a black hole. He had traced light trajectories in the gravitational field of an idealised black hole (see Fig. 1.2). Combining radio observatories around the globe into a single, giant interferometer, the *Event Horizon Collaboration* (EHT) produced the first images of the shadow cast by the event horizon of the super-massive black hole (SMBH) in the core of the galaxy M87 (Event Horizon Telescope Collaboration et al. 2019a, see Fig. 1.3, left panel). In 2022, a second set of images was published. This time the EHT collaboration targeted the black hole in the centre of the Milky Way, finding the shadow of Sgr A* (Event Horizon Telescope Collaboration et al. 2022, see Fig. 1.3, right panel).

Active galactic nuclei (AGN) are titanic manifestations of black hole activity in the Universe. They are luminous structures located at the centre of galaxies. AGN sign-post

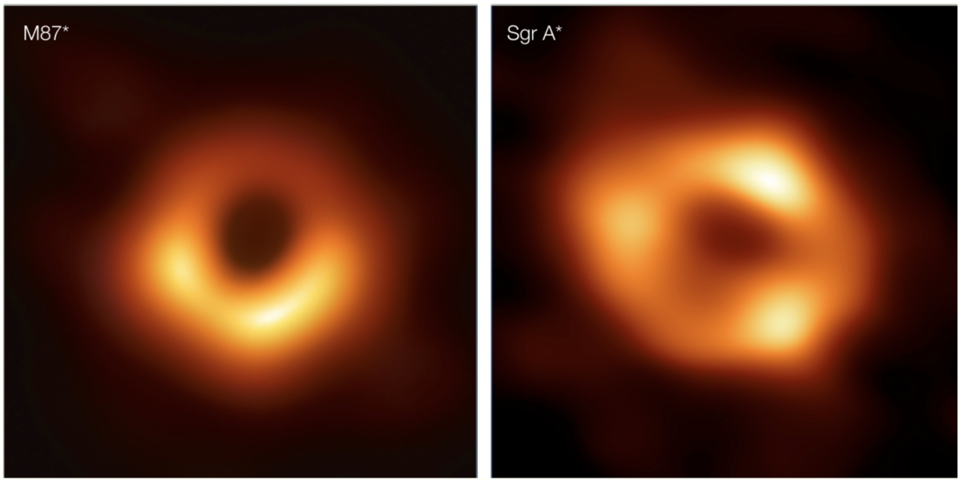


Figure 1.3: The EHT renderings of the black holes in the centres of M87 (left) and the Milky Way (right). Adapted from Event Horizon Telescope Collaboration et al. (2019a, 2022).

the population of accreting SMBHs and their evolution throughout cosmic time. The black-hole accretion mechanism drives their extreme luminosities and often causes the cores to outshine their host galaxies across the electromagnetic spectrum. While researchers are now converging towards a standard model of AGN, these objects have been classified in a notoriously complex, layered system of sub-types. The initial taxonomy of AGN was driven by the iterative discovery of peculiar electromagnetic signatures associated with flavours of the AGN phenomenon. While studying this semantic labyrinth and various classification schemes is of great historical value, here I will primarily focus on the key discoveries that established our modern model of AGN in Section 1.2. In Section 1.3, I will briefly review the modern AGN standard model and its key components. A central question driving the work presented in this thesis is: *How and when did the first SMBHs form ?* Measurements of the black hole mass and the space density of AGN are essential to tackle this issue. I present an overview of black hole mass estimates from broad emission lines and X-ray AGN population studies in Sections 1.4 and 1.5. With these tools in hand, I will move back in cosmic time to the first gigayear after the Big Bang in Section 1.6, and will discuss how the population of quasars, the brightest class of AGN, can be used to trace accretion onto $> 10^8\, M_\odot$ SMBH in the early Universe.

1.2 The AGN Standard Model I: Milestone discoveries

About a century ago, Hubble (1926) reported his observations of *planetary spectra* in the supposedly stellar nuclei of certain extragalactic spiral nebulae, inadvertently documenting an early AGN observation. By then, Curtis (1918) had already observed a jet emanating from the nucleus of M87. In his PhD thesis, Seyfert (1943) interpreted the broad emission lines observed in the spectra of bright, *nearly stellar nuclei* as the signature of strongly Doppler-shifted gas distributions. The term *Seyfert galaxy* now refers to objects with a detectable host galaxy and a nucleus emitting high-ionisation lines. These initial milestone results regained interest in the 1960s after astronomers had started to identify discrete celestial radio sources (e.g., Ryle et al. 1950; Smith 1952; Edge et al. 1959; Bennett 1962). The astrometric accuracy of the identified radio sources enabled the identification of optical counterparts (e.g., Baade & Minkowski 1954). Some counterparts to these radio sources had the morphological appearance of stars and were called *quasi-stellar radio sources.* Investigating a spectrum of the optical counterpart to one of the components of the source 3C 273, Schmidt (1968) identified a sequence of emission lines as the redshifted Balmer series, setting the source at $z = 0.16$. The first *quasar* had been discovered (Schmidt 1968; Oke 1963; Hazard et al. 1963). Soon after, the cosmological interpretation of the redshift of 3C 273 and similar objects was adopted (Greenstein & Schmidt 1964).

The large distances to these newly discovered bright sources implied they were extremely luminous[1]. Which physical mechanism could be responsible for the observed energy release? In 1964, Zel'dovich and Salpeter proposed black hole accretion as the central powering mechanism of quasars, an idea that further coalesced into a disk of accreted matter responsible for photoionizing radiation (Lynden-Bell 1969). Shakura & Sunyaev (1973) proposed the thin accretion disk model to describe the inward accretion flow.

With the quasar redshifts established as cosmological, Gunn & Peterson (1965) postulated that the neutral hydrogen distribution along our line of sight towards these quasars should imprint Lyman α absorption features in the quasar spectrum bluewards of the Lyman α emission line at the redshift of the absorbing system. Extended distributions of neutral hydrogen would cause large absorption troughs. This finding established quasars as cosmological tools to probe foreground structures and neutral hydrogen fractions (for recent results, see McGreer et al. 2015; Bañados et al. 2018). From an observational point of view, the *Gunn-Peterson* troughs allow the selection of high-redshift quasars well into the epoch of reionization. They are, therefore, crucial to this thesis (see Section 1.6.1).

One key observable for AGN is the width of certain emission lines. It gave rise to a classification in *Seyfert 1* vs *Seyfert 2*, historically denoting *Seyfert* galaxies with or without broad emission lines. An early observation was that only permitted lines can be broad, while forbidden emission lines remain narrow in AGN spectra (e.g., Khachikian & Weedman 1974). This dichotomy progressively led to the postulation of two distinct emission regions (e.g., Woltjer 1959; Soffrin 1969; Shields 1974). The idea of a Broad Line Region (BLR) of gravitationally bound, rapidly orbiting gas in the inner region of the nucleus emerged. At larger distances from the black hole, a distinct low-density gaseous region called the Narrow Line Region (NLR) emits the narrow lines. With the energetics derived from the emission lines, photoionization was quickly identified as the line emission

---

[1] In this thesis, I will use the term *quasar* to refer to luminous AGN in which the nuclear emission dominates the host.

mechanism (Souffrin 1969).

How could it be explained that certain AGN had broad permitted lines and others not? Performing spectropolarimetric observations of NGC 1068, (Antonucci & Miller 1985) found broad Balmer lines in the polarised flux of the source, resembling those observed in *Seyfert 1* galaxies. Extending the analysis to imaging polarimetry, Miller et al. (1991) could corroborate this result. The observations established that the BLR could be embedded and obscured by a dusty toroidal structure. In such a model, the polarised light observed in *Seyfert 2* arises from nuclear emission made visible to the observer by polar scattering close to the axis of the dusty torus.

In 1962 Riccardo Giacconi et al. developed and launched a rocket experiment that measured soft X-ray emission from a source outside the solar system. The detection of Scorpius X-1, the first-known X-ray binary system, paved the way for modern X-ray astronomy, a field that rapidly caught up with the successive AGN discoveries made in the optical and radio wavebands. Indeed, X-ray sources coincident with quasars and Seyfert galaxies were soon found (Friedman & Byram 1967; Bowyer et al. 1970; Gursky et al. 1971). Elvis et al. (1978) identified X-ray emission as a ubiquitous characteristic of Seyfert 1 galaxies and located its origin at the very core of AGN, close to the black hole. They were the first to measure an X-ray luminosity function (XLF) for a sample with luminosities $42 < \log\left(L_{\mathrm{X}}/(\mathrm{erg\,s^{-1}})\right) < 45$.

One distinctive property of AGN caught the eye of astronomers early on: the substantial variability of their emission across the entire electromagnetic spectrum. In an *information bulletin*, Sharov & Efremov (1963) reported long-term variability of the primordial quasar, 3C 273, measured from 44 photometric plates taken between 1896 and 1960. Variability on the scale of days was reported in the optical wavebands by Matthews & Sandage (1963) for the quasar 3C 48. In 1965 Dent also measured radio variability in three quasars. In X-rays, extreme variability on scales of days and years was observed in the cores of active galaxies by Winkler & White (1975); Ives et al. (1976) and Marshall et al. (1981). In addition to broad-band variability tracing changes in the continuum, AGN display spectral variability: changes in the continuum are followed by variations of the broad lines with a slight delay (of a couple of weeks, Cherepashchuk & Lyutyi 1973; Boksenberg et al. 1978). This delay between continuum and emission line variations was interpreted as response delay, i.e. the time the light from the continuum takes to reach the line-emitting clouds in the BLR. Finding response delays in the BLR is called *reverberation mapping* (RM). This technique reveals the structure, scale and kinematics of the BLR (Bochkarev & Antokhin 1982; Blandford & McKee 1982; Peterson 1993).

Early very-long baseline interferometric observations of quasars led to the discovery of apparent superluminal motion in the radio cores of quasars. In 1971, consecutive radio observations of 3C 279 revealed an expanding "double" source on the scale of 10^{-3} arcsec. Given the redshift of the source, this expansion appeared to progress at a velocity of $10c$ (Knight et al. 1971; Whitney et al. 1971; Cohen et al. 1971). This apparent superluminal motion is now interpreted as a special relativistic effect: the interval at which light is emitted from a distant object moving close to the speed of light towards the observer is gradually reduced, giving rise to the effect. This idea was formalised by Blandford &

Königl (1979), who brought forward the idea of relativistic jets. An additional classification arose from the presence of this extended radio emission. Fanaroff & Riley (1974) classified radio galaxies according to the separation of the bright lobes on each side of the galaxies, effectively separating core-dominated (*FR 1*) from lobe-dominated galaxies (*FR 2*).

While the black hole is usually dwarfed by its host galaxy (e.g., the analogy of *a coin in comparison to the earth*, Fabian 2012), the large amounts of energies generated by the AGN can potentially have a significant effect on the gas reservoirs of their host galaxies. After the initial landmark findings discussed above gradually provided a standard picture of AGN, the discovery of two scaling relations particularly motivated AGN feedback and black hole-galaxy co-evolution studies: the black hole mass to the stellar velocity dispersion in the bulge and to the bulge stellar mass (Magorrian et al. 1998; Gebhardt et al. 2000a; Ferrarese & Merritt 2000; Marconi & Hunt 2003; Häring & Rix 2004). While these scaling relations suggest a physical coupling between the growth of SMBHs and host galaxies, Jahnke & Macciò (2011) argued that they could also arise from hierarchical assembly through mergers.

Two modes of feedback have been identified: the radiative mode and the kinetic mode. The radiative mode is associated with black holes accreting at the Eddington limit, the theoretical maximal luminosity reachable in hydrostatic equilibrium:

$$L_{Edd} = \frac{4\pi G m_{\mathrm{p}}}{\sigma_{\mathrm{T}}} M_{\mathrm{BH}} \tag{1.4}$$

σ_{T} is the Thomson scattering cross-section, m_{p} the mass of the proton and M_{BH} the black hole mass. This feedback mechanism is related to winds or radiation pressure and is expected to be dominant at cosmic noon. The kinetic mode is related to shocks induced by the relativistic jet and is expected to become important at lower redshifts. These feedback mechanisms heat the gas and suppress star formation (e.g., Springel et al. 2005).

A significant development was the attempt to unify the observational signatures of quasars and, more generally, AGN into one coherent model based on the orientation of these sources (Antonucci 1993; Urry & Padovani 1995). The key components of this model are the black hole and accretion disk, the dusty torus, the BLR, the hot electron corona, the NLR and, in some instances, the jet. Each of these components imprints characteristic features on the SED of AGN. A schematic view of the AGN SED by Harrison (2014) is presented in Fig. 1.4.

In the following Section, I will briefly review critical components of the contemporaneous standard model of AGN and discuss their typical observational signatures: the accretion disk, the torus, the BLR and the hot electron corona. A schematic representation of this model is shown in Fig. 1.5. Various lines of sight explain the different classes of AGN that can be observed.

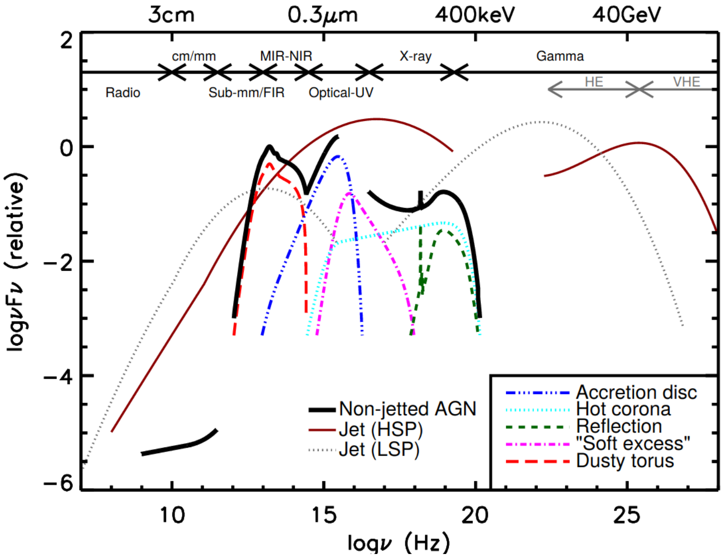


Figure 1.4: The AGN SED and its individual components (listed in the legend). AGN emit over the entire electromagnetic spectrum. A non-jetted composite model is shown in solid black. The contribution from low- and high-synchrotron peaked jet-dominated sources (LSP and HSP) are also shown. From Padovani et al. (2017), initially adapted from Harrison (2014). Image credit: C. M. Harrison.

1.3 The AGN Standard Model II: A contemporary view of the AGN topology

1.3.1 Black Hole Accretion: Disks as central engines

As we have seen in the previous section, AGN are powered by a central SMBH. The gravitational energy of infalling gas is converted efficiently to radiation due to the viscosity of the flow (Salpeter 1964; Zel'dovich 1964; Zel'dovich & Novikov 1967; Lynden-Bell 1969).

In a viscous disk, the differential rotation of the disk causes friction, which causes the accreted material to lose angular momentum and ultimately fall onto the event horizon. As angular momentum is conserved, other components of the disk must gain angular momentum in the accretion process.

The basic model is the $\alpha-$disk, proposed in 1973 by Shakura & Sunyaev. Accounting for orientation and relativistic effects Sun & Malkan (1989) fitted disk models to quasar continua finding that for both Schwarzschild and Kerr geometries, extreme luminosities were driven by accretion close to the Eddington limit. The typical quasar $\alpha-$disk reaches temperatures of 10^5 K and accretes matter in a thin, optically thick disk configuration. This is the source of the principal continuum emission via black-body radiation.

Other accretion modes supporting advective energy transport were later explored. Most prominently, an optically thin, geometrically thick (and radiatively inefficient) advection

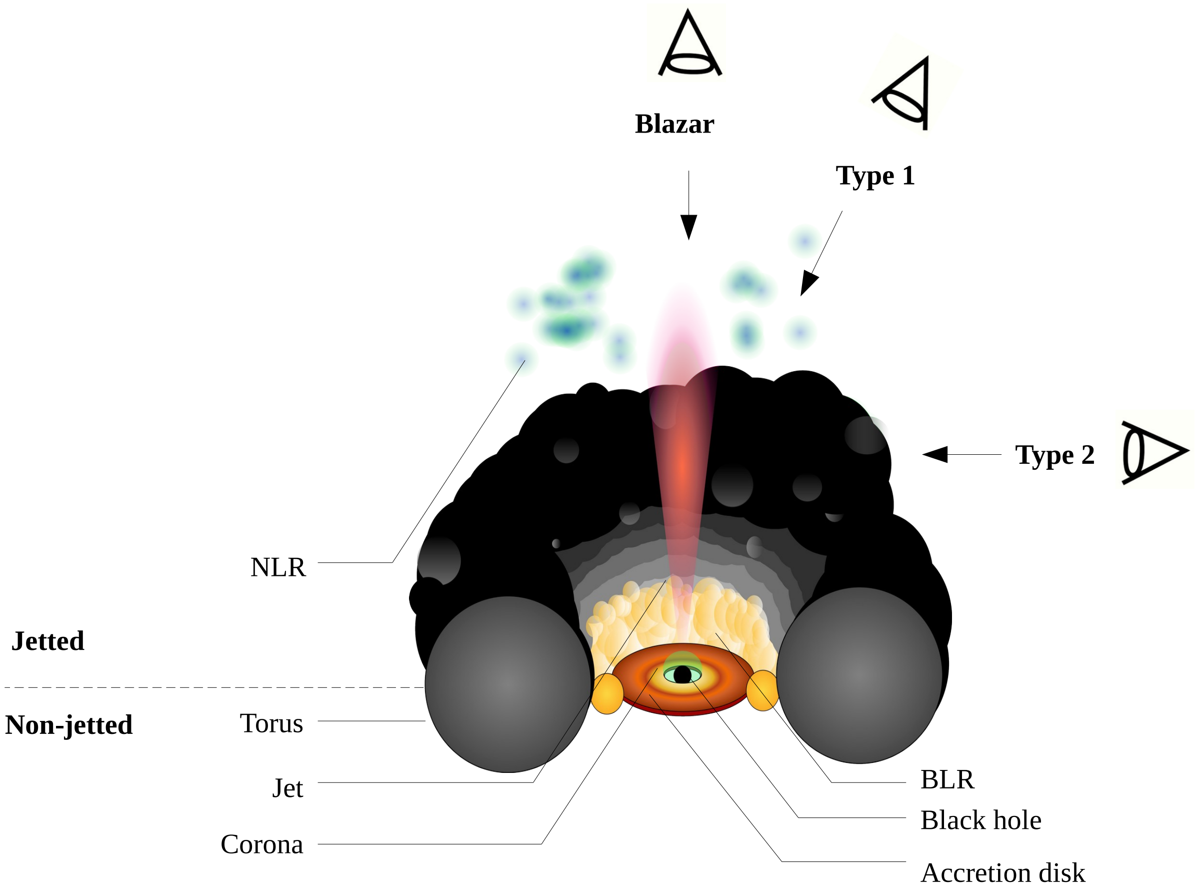


Figure 1.5: The AGN basic unified model. The main components are the black hole, the accretion disk, the BLR, the NLR, the corona, the jet and the obscuring torus. Different angles of view explain the different observable AGN SEDs and the original AGN taxonomy. For instance, blazars are jetted AGN seen pole on.

dominated accretion flow (ADAF) model has been developed to account for cases where black holes radiate at a few per cent of the Eddington luminosity (Ichimaru 1977; Rees et al. 1982; Narayan & Yi 1994, 1995a,b; Abramowicz et al. 1995). Fitting models of the spectral energy distribution (SED) of AGN accounting for a composite accretion mode (ADAF in the inner region and disk in the outer), it could indeed be shown that the disk is truncated at ~ 100 times the gravitational radius $R_g = GM/c^2$ (e.g., Nemmen et al. 2014) for low-luminosity AGN.

1.3.2 The dusty torus

The key component to explain the anisotropic emission from AGN is the torus, a dusty molecular structure surrounding the nuclear black hole, disk and BLR complex. "Surrounding" is a deliberately vague term, as the exact geometry of the dust structure is still intensively debated.

The torus was initially introduced to explain nuclear obscuration in optical bands, separating *Seyferts* in two types, "1" (unobscured) and "2" (obscured) (Antonucci & Miller 1985). At equatorial viewing angles, the torus obscures the BLR, justifying the classification as Seyfert 2, or in more modern terminology, a *Type 2* AGN. Henceforth, I will adopt the following convention for the obscuration-based dichotomy of AGN (e.g., Ramos Almeida & Ricci 2017):

- Type 1: AGN that show broad emission lines in their spectra, i.e. with full-width at half-maximum (FWHM) $> 2000\,\mathrm{km\,s^{-1}}$

- Type 2: AGN that show only narrow emission lines in their spectra, i.e. they have FWHM $< 2000\,\mathrm{km\,s^{-1}}$

The near-infrared (NIR, $\sim 1-7\mu$m) and mid-infrared (MIR, $\sim 7-30\mu$m) SED of AGN has been studied with high-resolution ground and space-based imaging campaigns (e.g., Alonso-Herrero et al. 2003; Ramos Almeida et al. 2009; Prieto et al. 2010). For reviews on the restframe IR emission of AGN see Ramos Almeida & Ricci (2017) and Hönig (2019). Regardless of their type, all AGN appear to share a common feature in the infrared part of their SED: a bump peaking in the MIR range. It is now assumed that the mechanism responsible for this bump is the isotropic reprocessing of nuclear continuum emission by dust grains in the torus. A ubiquitous feature in the NIR SED of AGN is a turnover at $\sim 1\,\mu$m. This break can be explained by considering dust sublimation. Each dust species has a set temperature above which it will evaporate, e.g.,~ 1400 K for silicate grains and ~ 1800 K for graphite grains (Barvainis 1987; Netzer 2015, and references therein). The inner radius of the torus corresponds to the dust sublimation radius, which has been resolved with NIR observations using the Very Large Telescope Interferometer (VLTI, Weigelt et al. 2012; GRAVITY Collaboration et al. 2020) and the Keck telescope (e.g., Swain et al. 2003; Pott et al. 2010; Kishimoto et al. 2011; Koshida et al. 2014). It has also been investigated in RM studies (e.g., Suganuma et al. 2006).

These observations established that the dust sublimation radius scales with the continuum luminosity as $r_s \sim L^{1/2}$.

High-resolution MIR imaging campaigns have allowed estimates of the scale of the torus (0.1-10 pc, e.g., Packham et al. 2005; Radomski et al. 2008). These scales are also supported by MIR interferometric observations with the VLTI, which resolves cold dust structures at larger radii than the dust sublimation radius r_s (Jaffe et al. 2004; Tristram et al. 2007; Meisenheimer et al. 2007; Raban et al. 2009).

An important tracer of anisotropy in the MIR band is the silicate feature at 9.7 μm. The strength of the feature correlates with the absorbing HI column densities in X-rays, which are due to obscuring material along the line of sight (see 1.3.4). While absorption at 9.7 μm was found to be coincident with high HI column densities, emission at 9.7 μm corresponds to low HI column densities (e.g., Shi et al. 2006). 9.7 μm absorption has, however, been observed in Type 1 AGN (Roche et al. 1991). These findings seriously challenged the torus geometry (and, by extension, the basic unification scheme): How could the shallow absorption in the silicate feature be explained in the absence of obscuring dust along the line of sight? Instead of a smooth gas distribution, the possibility of a clumpy material distribution was explored with radiative transfer models by Nenkova et al. (2002). The clumpy model was further formalised and investigated by Nenkova et al. (2008a,b). The clumpiness could naturally account for silicate absorption in Type 1 AGN. In a clumpy torus, individual clumps at different radii can be illuminated simultaneously, as there is no uniform shielding from nuclear radiation.

MIR-interferometry also revealed that the geometry of the dusty obscurer might be more complex than the canonical torus (Raban et al. 2009; Hönig et al. 2012, 2013; Tristram et al. 2014; López-Gonzaga et al. 2016). These studies have led to the development of an alternative, two-component model of the torus consisting of a thin extension of the BLR and a polar component formed by dusty winds (e.g., Hönig & Kishimoto 2017). It is conceptually close to the Failed Radiatively Accelerated Dusty Outflow model for the BLR proposed by Czerny & Hryniewicz (2010). The polar dust has, however, been detected on scales of parsecs (e.g., Asmus et al. 2016), geometrically coinciding with the NLR. Whether the polar dust contributes to the obscuration of the inner parsec remains to be confirmed. A schematic representation of this modern view of the torus is shown in Fig. 1.6.

The dusty gas can also be probed at sub-mm wavelengths with the Atacama Large sub-Millimeter Array (ALMA), by tracing molecular lines. This led to the discovery of extended, circumnuclear disk-like molecular tori on scales of a few parsecs (Combes et al. 2019). The molecular torus is seen edge-on for most Type 2 AGN in this survey.

1.3.3 The Broad Line Region and nuclear kinematics

Type 1 AGN restframe optical/UV spectra show remarkable similarities, including e.g., the relative strength of emission lines. This implies that the photoionisation equilibrium is attained at roughly the same temperature of $T \sim 10^4$K (e.g., Peterson 2006). AGN typically display broad and narrow emission lines which arise from the different gaseous regions of different densities (Khachikian & Weedman 1974). The BLR is the source of the

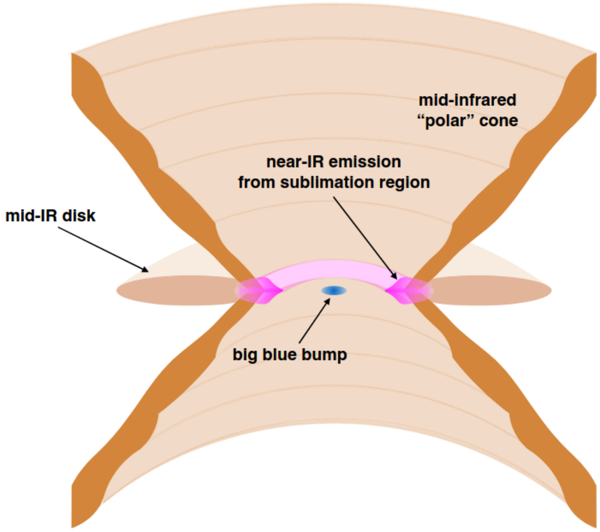


Figure 1.6: Modern schematic view of the torus. NIR interferometry supports a two-component model with dust polar and equatorial dust distributions. The MIR emission arises from the outer regions of the equatorial dust (isotropically) and the polar dust, while the NIR emission arises where the dust is warmer, close to the sublimation radius. From Hönig & Kishimoto (2017).

broad emission lines in Type 1 AGN and is located close to the core. The dust sublimation radius of the torus delimits it. Photoionised by the central continuum emission, the BLR emits transition lines, which are expected to be Doppler-broadened by the motion of the gas (e.g., Baldwin et al. 1995; King 2016). The emitting gas itself is assumed of high electron density ($n_e > 10^9 cm^{-3}$), given the absence of broad components in forbidden lines. The broad lines have FHWM of 1000-25000 $\mathrm{km\,s^{-1}}$ (Peterson 2006).

Taking advantage of flux and spectral variability in AGN, the structure of the BLR can be probed with the RM technique (Blandford & McKee 1982; Peterson 1993). As introduced in Section 1.2, this method cross-correlates variations of the continuum flux and variations in the flux of broad emission lines. The time delay τ between continuum and line variations is measured. Under the assumption of a Keplerian velocity field, the time delay corresponds to the light travel time between the disk and the BLR, yielding the radius of the BLR as $R_{\mathrm{BLR}} = \tau/c$. The BLR typically extends over 10-100 light-days (e.g.,Wandel et al. 1999). In 1.4, I will discuss how RM can further be used to measure the mass of the central SMBH in AGN.

The debate over the actual BLR kinematics and geometry has flourished through the past decades; the variety of the often asymmetric (sometimes double-peaked) broad lines and their variability has led to an abundant diversity of competing models. The basic view is that the BLR is composed of virialised clouds (Peterson & Wandel 1999). A flattened BLR geometry was initially supported by the relation between the width of Hβ and the

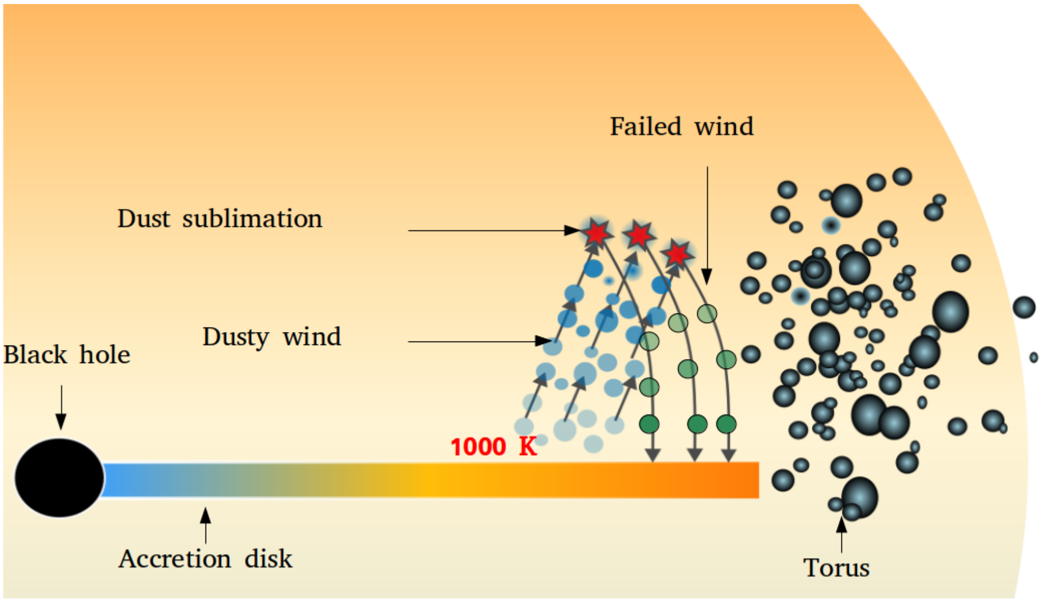


Figure 1.7: BLR kinematics from failed radiative driven outflows, a possible model to explain the asymmetric shape of broad emission lines. Dust above the outer edges of the disk ($T \sim 1000\,\mathrm{K}$) is driven out vertically by radiative pressure. The dusty gas is exposed to the inner, hotter part of the disk. As the dust evaporates, the plasma falls back on the disk. Adapted from Czerny et al. (2018).

radio core-to-lobe ratio. (Wills & Browne 1986; Jarvis & McLure 2006; Brotherton et al. 2015) and by the double shoulders and peaks observed in certain broad emission lines that favour a disk configuration (Oke 1987; Chen et al. 1989; Chen & Halpern 1989; Eracleous & Halpern 1994, 2003; Lewis et al. 2010; Storchi-Bergmann et al. 2016). The double-peaked profile variability has been shown not to correlate with continuum changes. Thus, the peak shifts are not a reverberation effect Wanders & Peterson (1996).

Another, possibly complementary, explanation for the observed broad line shapes is the presence of (rotational) disk winds (Emmering et al. 1992; Murray et al. 1995; Elvis 2000; Everett 2005; Keating et al. 2012). Czerny & Hryniewicz (2010) have developed a physically motivated model in which outflowing dust grains play a role in the BLR kinematics. This model describes a failed dusty wind. Dust can exist above the outer edge of the accretion disk, where the atmosphere temperature is below the dust sublimation temperature. Under radiative pressure, a dusty wind is driven out vertically. Once it reaches a sufficient height above the disk, the plasma is not shielded anymore and is irradiated by the inner parts of the disk. The dust evaporates, causing the outward radiative force to vanish, the wind fails and dust-free gas falls back on the disk. This model naturally reproduces blueward and redward asymmetries in broad Balmer lines. A schematic representation of the model is shown in Fig. 1.7,

It was not until recently that the kinematics of this region could be resolved for the first time. Measuring the spatial offset between the red and blue shifted Paschen-α line in

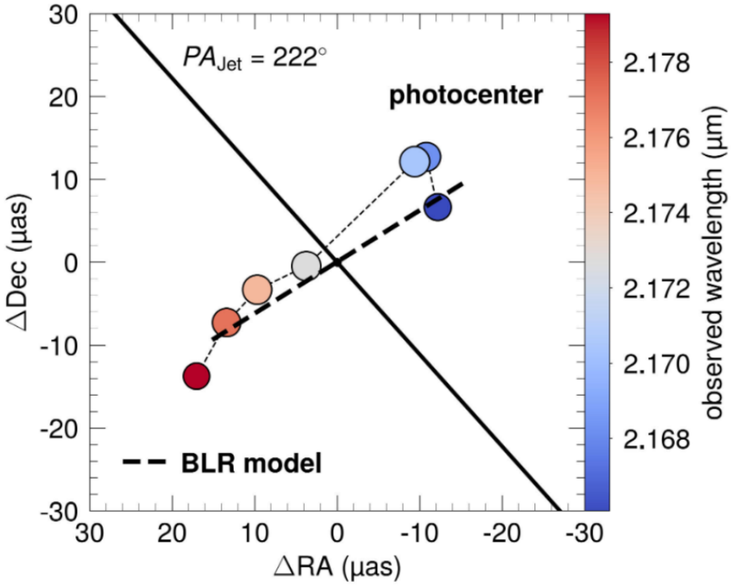


Figure 1.8: Spatial offsets of the broad Paschen-α line in 3C 273. The points are colour-coded according to their wavelength. The black line represents the jet axis of the AGN. From Sturm et al. (2018).

3C 273, Sturm et al. (2018) revealed a velocity gradient perpendicular to the jet axis. The photo-centre offsets are shown in Fig. 1.8. The data favours a thick disk model for the BLR as an orbiting extension of the accretion disk, with an extent of 150 light days.

1.3.4 The hot electron corona and the jet: X-ray emission mechanisms

For this thesis, observations in the soft X-ray band have been central to tracking the AGN population through cosmic time. Here we briefly discuss the origin of the ubiquitous high-energy emission in AGN. In the nuclear region, thermal radiation emerging from the accretion disk is reprocessed by a hot ($\sim 10^9$ K) and optically-thin plasma: the corona. The optical/UV disk photons are inverse-Compton scattered to high energy by the relativistic coronal electrons (Haardt & Maraschi 1991). This is the primary X-ray emission mechanism in AGN and the multiple inverse Compton scatterings produce a power-law continuum (see e.g., Rybicki & Lightman 1979):

$$P\,[\text{photons}\,\text{s}^{-1}\,\text{keV}^{-1}] \sim E^{-\Gamma} \times e^{-E/E_{\text{cut-off}}} \tag{1.5}$$

P is the photon flux, E is the energy, $E_{\text{cut-off}}$ is the cut-off energy and Γ is the photon index. For typical Type 1 AGN, the photon index is $\Gamma = 1.8{-}2.0$ (Nandra et al. 1997), while populations exist with softer indices (Boller et al. 1996, see also Chapter 5 of this thesis). The coronal continuum of AGN displays a rollover at high-energies ($> 100\,\text{keV}$ e.g., Perola et al. 2002; Malizia et al. 2014; Ricci et al. 2018). In Eq. 1.5, this is phenomenologically accounted for with the parameter $E_{\mathrm{cut-off}}$. $E_{\mathrm{cut-off}}$ is related to the corona temperature (e.g., Middei et al. 2019).

The corona irradiates the nuclear region with X-rays. The irradiation of the accretion disk, the BLR and the torus give rise to reflection components in the AGN spectra (e.g., Miller 2007). The most prominent are the Fe Kα 6.4 keV line and the Compton reflection hump at $\sim$ 30keV (George & Fabian 1991; Matt et al. 1991). The relativistic, broad Fe Kα line is expected to be emitted from the reflection at the accretion disk. It is, therefore, a probe of the conditions deep in the gravitational well of the central black hole. Its asymmetric profile was predicted for Schwarzschild and Kerr black holes by Fabian et al. (1989) and Laor (1991). It was since confirmed in systematic surveys (e.g., Nandra et al. 1997). The innermost stable circular orbit of an infalling particle depends on the spin of the black hole. The shape and width of a line emitted in the inner accretion disk can thus, in principle, provide constraints on the black hole spin (see, e.g., Miller 2007; Reynolds & Fabian 2008, and references therein)

The corona is confined to a small region of the AGN, within 2-20 gravitational radii R_G from the black hole. Similarly to RM for the BLR, its geometrical structure can be constrained by cross-correlating variability in the coronal emission and reflection features. In practice, this is done by measuring lags between various X-ray bands in Fourier-frequency space (e.g., Zoghbi et al. 2010; De Marco et al. 2013; Kara et al. 2016, for a review, see Uttley et al. 2014). Further constraints on the size of the primary X-ray emission region is provided by X-ray photometry from gravitationally lensed quasars and the comparison to modelled microlensing lightcurves (e.g., Guerras et al. 2017).

X-ray photons, at $\sim 0.1-10$ keV, can be heavily affected by intervening matter because of the photoelectric effect. This effect is less critical for hard X-rays, at $>$ 10 keV, as the probability of photoelectric absorption is inversely proportional to the cube of the energy of the incident photons. For AGN, this absorption is expressed in terms of the line-of-sight hydrogen column density N_{H}. Soft X-ray absorption is preferably found in Type 2 AGN, further supporting the AGN unification model (e.g., Matt et al. 1997).

At <1 keV, an emission in excess of the extrapolated hard X-ray continuum of Type 1 AGN has been discovered by Singh et al. (1985) and Arnaud et al. (1985). It has since been confirmed in most unobscured AGN (e.g., Turner & Pounds 1989; Bianchi et al. 2009; Scott et al. 2012). The physical origin of this *soft excess* is still debated. Possible explanations are inverse Compton scattering of disk photons in a warm corona, i.e. a plasma distinct from the continuum-emitting hot corona (e.g., Magdziarz et al. 1998), or the relativistic blurring of emission lines emitted from the inner accretion disk by the SMBH (Crummy et al. 2006). Simulated Type 1 AGN X-ray spectra accounting for the above components are shown in Fig. 1.9.

In addition to direct and reprocessed coronal emission, jetted and non-blazar AGN often display a second hard X-ray power-law which becomes dominant at $>$ 2 keV (e.g., Grandi & Palumbo 2004; Madsen et al. 2015, see jet X-ray contributions in Fig. 1.4). This power-law is thought to arise from inverse Compton scattering and synchrotron self-Compton emission at the base of the jet (e.g., Potter & Cotter 2012). Jets can also reprocess ambient photons

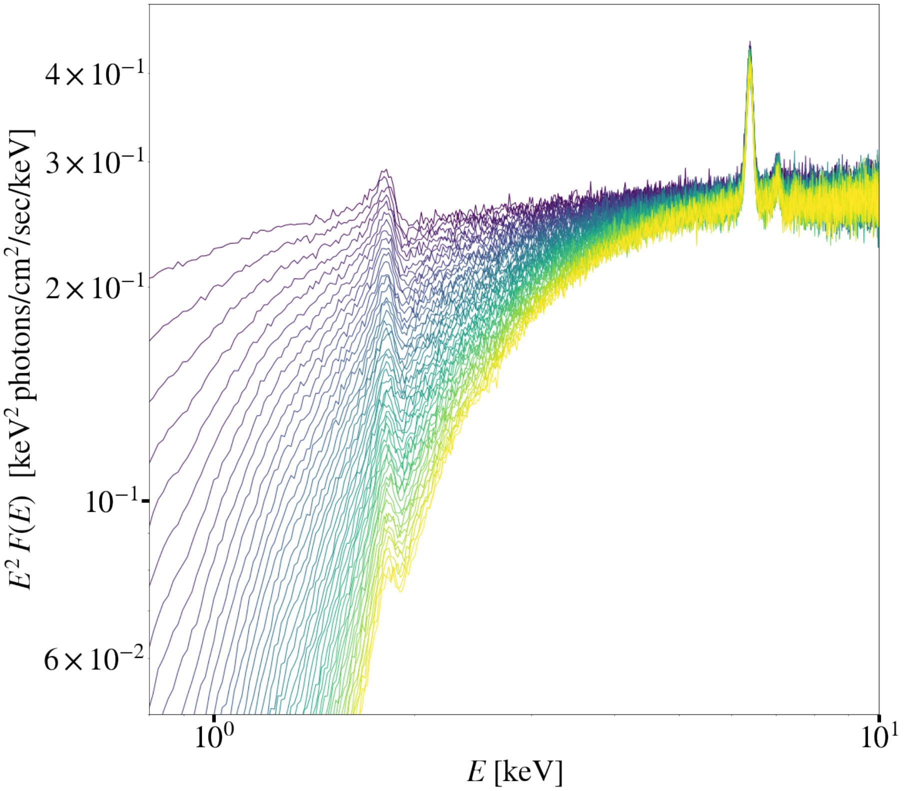


Figure 1.9: Simulated AGN X-ray spectra with increasing absorbing column density N_H from the host. Using the X-ray spectral fitting tool `XSPEC` (Arnaud 1996), I have modelled the X-ray emission from a typical Type 1 AGN with a simple power-law with a photon index $\Gamma = 1.95$ (Nandra & Pounds 1994) and an exponential cut-off at 200 keV for the primary coronal emission (`cutoffpl`), a second power-law weighted by a factor of 0.03 (`constant`) for the intrinsic soft-scattered component (e.g. Gilli et al. 2007) and a Compton reflection component (`pexmon`, Nandra et al. 2007). I account for host obscuration with the Tübingen-Boulder absorption model `tbabs` (Wilms et al. 2000) and let the absorbing column vary between $N_\mathrm{H} = 5.0 \times 10^{20} - 1.5 \times 10^{23}\,\mathrm{cm}^{-2}$. The model was convolved with the Chandra ACIS-S instrument response matrices: the Response Matrix File (RMF) and Ancillary Response File (ARF)(`https://cxc.cfa.harvard.edu/caldb/prop_plan/imaging/`). The total model is: `tbabs` $*$ (`cutoffpl` $+$ `constant` $*$ `cutoffpl` $+$ `pexmon`). The increasing absorbing column density (from purple to yellow) results in increasing depletion of the total soft X-ray flux.

to boost the X-ray emission of AGN. The high X-ray detection rate of radio-loud/jetted quasars has motivated the development of a model in which cosmic microwave background (CMB) photons are inverse Compton-scattered by relativistic electrons in the jet (iC-CMB, Tavecchio et al. 2000; Celotti et al. 2001). As the energy density of the CMB is proportional to $(1+z)^4$, this effect is expected to enhance the X-ray emission of radio-loud AGN at higher redshifts significantly (Schwartz 2002). When this thesis was written, X-ray jets have been discovered up to $z = 6.1$ (Ighina et al. 2022).

1.4 SMBH population studies I: Black hole masses from broad AGN emission lines

In this thesis, I characterise the evolution of black holes, as determined by their comoving space density at different times. However, the growth channels of black holes also play a critical role in our understanding of the overall population (see 1.6.2). The determination of SMBH masses M_{BH} is therefore essential. The deep gravitational potential of the SMBH strongly affects stellar and gas kinematics in their host galaxies, motivating the development of multiple and complementary techniques for M_{BH} measurements (for a review see Peterson 2010).

Nuclear stellar dynamics in the Milky Way have been used to measure the mass of its central black hole (Ghez et al. 2000; Genzel et al. 2000). This technique is, however, limited to our Galaxy as it is the only case for which the proper motion of nuclear stars can be resolved. Fitting general relativistic magnetohydrodynamic models to the EHT images of M87, the Event Horizon Telescope Collaboration et al. (2019b) derived an angular gravitational radius for the gas distribution at the horizon and, thus, obtained a direct estimate of M_{BH}. Dynamical modelling of the spatially resolved gas and stellar distributions close to the nucleus of nearby galaxies can also be applied to constrain central mass distributions (e.g., Macchetto et al. 1997; Gebhardt et al. 2000b, 2003). To measure M_{BH} for AGN out to more considerable distances, where central gas kinematics are unresolved and outshone by the disk continuum, the indirect techniques of RM and the related single-epoch method are used. These techniques require the presence of broad emission lines in the AGN spectra to trace intrinsic BLR kinematics, i.e. they are limited to the Type 1 population. Recently, Baron & Ménard (2019) reported a correlation between the ionisation state of the NLR and the BLR kinematics, enabling, in principle, the estimate of M_{BH} in obscured AGN.

In RM, the time lags τ between variations in the disk continuum emission and the broad emission lines provide an estimate for the size of the BLR as $r_{\mathrm{BLR}} = \tau / c$ (Blandford & McKee 1982; Peterson 1993). The broadening of the emission lines can, under certain geometric assumptions, be used to estimate M_{BH}. The FWHM of broad emission lines serves as a proxy for the radial velocity of the BLR gas, which in combination with the distance to the black hole, yields an estimate for its virial mass:

$$M_{\mathrm{BH}} = \frac{r_{\mathrm{BLR}} f_{\mathrm{BLR}} \mathrm{FWHM_{BLR}}^2}{G} \tag{1.6}$$

$f_{\rm BLR}$ denotes the geometric form factor of the BLR. The Keplerian velocity $v_{\rm Kepl}$ of the gas is obtained by correcting for projection effects encoded in the geometric form factor: ${v_{\rm Kepl}}^2 = f_{\rm BLR}{\rm FWHM_{BLR}}^2$. $f_{\rm BLR}$, which takes into account the source orientation θ and the shape of the velocity field, described by the motion of the gas in the BLR, is calibrated using the $M_{\rm BH}$-stellar dispersion relation (Onken et al. 2004; Negrete et al. 2018) and accounting for the host bulge type (Ho & Kim 2014). The RM technique has been successfully applied in large AGN monitoring campaigns (e.g. Wang et al. 2014; Bentz & Katz 2015; Grier et al. 2017).

A by-product of RM studies was the confirmation of a robust correlation between the accretion disk continuum luminosity and the BLR size $r_{\rm BLR}$ (e.g. Kaspi et al. 2000), which does not rely on long-term monitoring of AGN and uses only continuum luminosities and line widths to determine virial masses. This method is referred to as *single-epoch* mass estimation (e.g. Woo & Urry 2002; Greene & Ho 2005; Vestergaard & Osmer 2009; Shen et al. 2011). An example of its application is presented in 3.6.

Broad, low-ionisation lines, such as Hβ (restframe wavelength at 4862$\AA$) and the MgIIλ2800 doublet (MgII), are generally thought to yield more reliable virial broadening estimators than, for instance, the high ionisation transition CIVλ1549 (CIV, e.g., Trakhtenbrot & Netzer 2012; Mejia-Restrepo et al. 2016; Mejía-Restrepo et al. 2018). This is in part due to the presence of prominent blueshifted components in the CIV line profiles of rapidly accreting sources, which may not be related to the orbital motion of the gas and would affect the inferred velocities (e.g., Richards et al. 2002; Zamanov et al. 2002; Marziani et al. 2017; Sulentic et al. 2017). At higher redshifts ($z > 6$), the Balmer line Hβ is redshifted out of optical and NIR spectrographs. To constrain the $M_{\rm BH}$ in the early Universe, authors have primarily made use of broad restframe UV lines observed in NIR spectra such as MgII (e.g., Kurk et al. 2007; Jiang et al. 2007; Willott et al. 2010; Mortlock et al. 2011; Onoue et al. 2019; Yang et al. 2021).

However, Hβ lines possess profiles which are difficult to reconcile with simple Keplerian motion around the central black hole and this could significantly affect the $M_{\rm BH}$ estimates (Negrete et al. 2018). Accurate black hole mass measurements are essential for the development of galaxy evolution models. It is thus critical to understand the mechanisms responsible for the observed broad line shapes. Aspects of the kinematics and structure of the BLR are explored in detail in Chapter 3.

1.5 SMBH population studies II: AGN demographics with X-ray surveys

In the previous section, I discussed how to capture SMBH growth via the estimation of individual black hole masses. To investigate the general evolution of accretion, we need to produce a census of the population of SMBH at different cosmic epochs. The light emitted from the accretion process is a beacon to pinpoint massive black holes and provides a measure of accretion power. By constraining the demographics of AGN, one can directly access the underlying population of strongly accreting and massive black holes over a large dynamical range of redshifts.

Hard X-rays photons (> 2 keV) are a clear signature of AGN activity and can reveal the population of obscured AGN and, to some extent, Compton-Thick AGN ($N_{\rm H} > 10^{24} cm^{-2}$). X-ray surveys of various depth and area configurations have allowed access to the AGN population. Missions such as the *Chandra X-ray Observatory* (Weisskopf et al. 2002), the *X-ray Multi-Mirror Mission* (XMM-Newton, Jansen et al. 2001), the *Neil Gehrels Swift Observatory* (Barthelmy et al. 2005) and ROSAT (Truemper 1982), have contributed milestone surveys for AGN demographics. The X-ray coverage of the Cosmic Evolution Survey (COSMOS, Scoville et al. 2007; Hasinger et al. 2007; Elvis et al. 2009), the AEGIS field (Nandra et al. 2015), as well as the *Chandra* Deep Field Surveys (Brandt et al. 2001; Xue et al. 2011; Ranalli et al. 2013) provided deep pencil beam surveys that sample AGN down to a limiting flux of $\sim 6.4 \times 10^{-18}\,{\rm erg\,s^{-1}\,cm^{-2}}$ in the $2 - 10$ keV band (Luo et al. 2017). The sensitivity of these surveys comes at the cost of smaller surveyed areas, making them complementary to wider X-ray surveyed fields such as the Sloan Digital Sky Survey (SDSS) Stripe 82 (LaMassa et al. 2013), XMM-XXL (Pierre et al. 2016) or field concatenations such as the *Chandra* Multi-Wavelength Project (ChaMP, Kim et al. 2004), the XMM-Newton serendipitous survey (4XMM, Watson et al. 2009; Webb et al. 2020) or the Extragalactic Serendipitous Swift Survey (ExSeSS, Delaney et al. 2022). The 2019-launched X-ray observatory extended ROentgen Survey with an Imaging Telescope Array (eROSITA, Merloni et al. 2012; Predehl et al. 2021) is performing an all-sky survey in the soft X-ray range and providing the largest AGN sample to date. More details are presented in Chapter 2. The sensitivity of some of the above X-ray surveys as a function of surveyed area are compared in Fig. 1.10. The limiting flux of soft X-ray surveys as a function area is also presented in Fig. 2.1.

A large fraction of this thesis focuses on quasars at $z > 5.7$ which have emitted the light we can observe in the first billion years after the Big Bang. Hard X-ray photons at > 10 keV are little affected by relatively large column densities ($\log N_{\rm H} < 23$, see Fig. 1.9). The emitted wavelength of a photon from a source at redshift z is stretched as it travels to the observer ($z = (\lambda_{\rm observed} - \lambda_{\rm restframe})/\lambda_{\rm restframe}$, equivalently, the restframe energy of the photons is reduced). For a source at $z > 5.7$, the hard X-ray emission at 10 keV will thus be collectable in the soft X-ray band as $E_{\rm observed} = E_{\rm restframe}/(1+z) \approx 1.5$ keV. Soft X-ray surveys provide a way to sample the high-redshift AGN population that is less prone to absorption biases than optical surveys.

The XLF quantifies the comoving space density of AGN in X-rays as a function of luminosity $L_{\rm X}$ and redshift z. $L_{\rm X}$ and z have to be derived from samples with well defined selection function. The redshift is obtained from the optical, NIR and UV counterparts of the X-ray sources. In Chapter 2, I present the core concepts of the identification procedure of multi-wavelength counterparts to X-ray sources. Ideally, the redshifts are spectroscopic, i.e. derived from the shift of emission and absorption lines in the restframe optical/UV spectra of the AGN. However, depending on the multi-wavelength coverage of the X-ray field and the availability of large-scale and sufficiently deep spectroscopic data (e.g., Comparat et al. 2020), photometric redshifts may be required (Salvato et al. 2019a). The

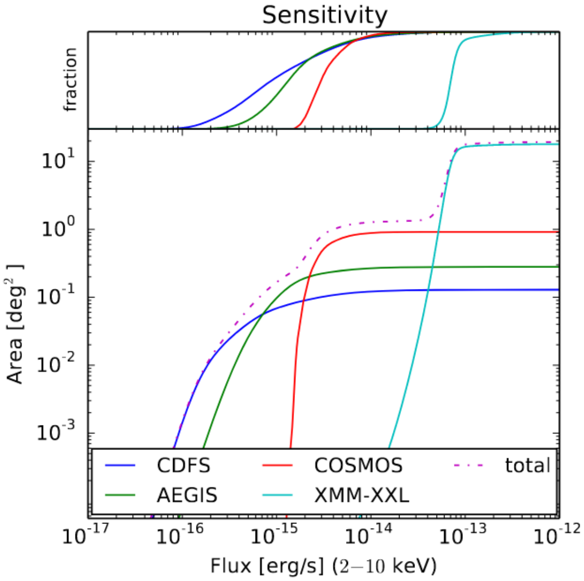


Figure 1.10: Area sensitivity curves of various X-ray surveys. From Buchner et al. (2015).

luminosities L_X of the AGN are obtained via the analysis of their X-ray spectra, which requires a model accounting for the various components of the AGN emission (see 1.3.4) and the redshift of the source. For instance, (Buchner et al. 2015) use the following `XSPEC` model for their AGN sample: `torus+pexmon+scattering`. This model accounts for the intrinsic coronal power-law and its Compton scattering off the torus, the Compton reflection and a soft scattering component. Such models account for the intrinsic obscuration of the AGN. It is essential to confirm the X-ray sources as AGN via the spectra of their counterparts directly, their broad-band photometric colours (e.g., Salvato et al. 2022) and their X-ray properties.

The XLF is measured by estimating the space density of AGN in bins of luminosity and redshift (e.g., Ueda et al. 2003; Hasinger et al. 2005; Silverman et al. 2005; Ueda et al. 2014; Miyaji et al. 2015; Buchner et al. 2015; Georgakakis et al. 2015; Aird et al. 2015; Fotopoulou et al. 2016; Ananna et al. 2019). Formally the XLF ϕ expresses the number N of objects per unit comoving volume V and X-ray luminosity L_X:

$$\phi = \frac{\mathrm{d}^2 N}{\mathrm{d}V\,\mathrm{d}\log L_\mathrm{X}} = \frac{\mathrm{d}\Phi}{\mathrm{d}\log L_\mathrm{X}}(z,\,\log L_\mathrm{X}) \tag{1.7}$$

Following Page & Carrera (2000), under the assumption that the XLF changes little in a given bin of redshift and luminosity $(\Delta z, \Delta L_\mathrm{X})$ one can estimate the XLF from the number N of detected sources in this bin as:

$$\phi_{est} = \frac{N}{\int\int A(\log L_\mathrm{X}, z)\frac{\mathrm{d}V}{\mathrm{d}z}\mathrm{d}\,z\,\mathrm{d}\log L_\mathrm{X}} \tag{1.8}$$

where $A(\log L_{\rm X}, z)$ is the sky area of the survey sensitive to $L_{\rm X}$ at redshift z and $\mathrm{d}V/\mathrm{d}z$ is the differential comoving volume. The statistical uncertainty in ϕ_{est} is given by:

$$\sigma\phi_{est} = \frac{\sigma_N}{\int\int A(\log L_{\rm X}, z)\frac{\mathrm{d}V}{\mathrm{d}z}\mathrm{d}\,z\,\mathrm{d}\log L_{\rm X}} \tag{1.9}$$

where σ_N is the 1σ Poisson uncertainty for N counts (e.g., Gehrels 1986). The XLF can be parametrized as a double power-law:

$$\phi_m = \frac{K}{(L_{\rm X}/L_*)^{\gamma_1} + (L_{\rm X}/L_*)^{\gamma_2}} \tag{1.10}$$

where K is the normalization, γ_1 and γ_2 the slopes of the power-law components and L_* the break luminosity. This double power-law is modified by a redshift-evolution term which can either be applied to the normalization or the break luminosity (or both).

Ueda et al. (2014), Vito et al. (2014), Buchner et al. (2015), Miyaji et al. (2015), Aird et al. (2015) and Georgakakis et al. (2015) have measured the XLF in the redshift bin $3 < z < 5$.

Vito et al. (2014) and Georgakakis et al. (2015) reported that a pure-density evolution (PDE) model best fits their data. This model assumes an evolutionary term, parameterised as a multiplicative factor to the normalisation K:

$$e(z) = \left(\frac{1+z}{1+z_0}\right)^q \tag{1.11}$$

where q controls the density evolution and $z_0 = 3$. Ueda et al. (2014) assumed a luminosity-dependent density evolution (LDDE) parametrisation. Based on previous observations of a decline in the comoving number of $\log\left(L_{2-10\,\rm keV}/(\rm erg/s)\right) > 44$ AGN at higher redshifts (Brusa et al. 2009; Civano et al. 2011; Hiroi et al. 2012), two cut-off redshifts were introduced in the parameterisation of the evolutionary term (Ueda et al. 2014, for further details). Miyaji et al. (2015) also used LDDE, accounting, in addition to absorption effects, for the probability distribution of photometric redshifts. Aird et al. (2015) introduced a flexible double power-law (FDPL) as parameterisation of the XLF. It allows any parameter in Eq. 1.10 to evolve with redshift. This redshift dependence is modelled by polynomials of $\log(1+z)$. Respectively using Bayesian modelling and neural networks, Buchner et al. (2015) and Ananna et al. (2019) derived non-parametric forms of the XLF. The above models describe an exponential decline of the AGN population with redshift in all luminosity bins. The space density of the most X-ray luminous quasars at $z > 5.7$ is the object of the work presented in Chapters 4 and 5 of this thesis.

An additional ingredient is required to quantify the entire parent population of active SMBHs: a measure of the effect of obscuration on AGN space density (e.g., Ueda et al. 2003, 2014; Aird et al. 2015; Buchner et al. 2015; Ananna et al. 2019). The fraction of obscured Compton-thin AGN and Comtpon-thick AGN shows dependencies on both $L_{\rm X}$ and z. Modelling and fitting the absorption function yields an estimate of the fraction of obscured Compton-thin and Compton-thick AGN in the Universe at different epochs. The

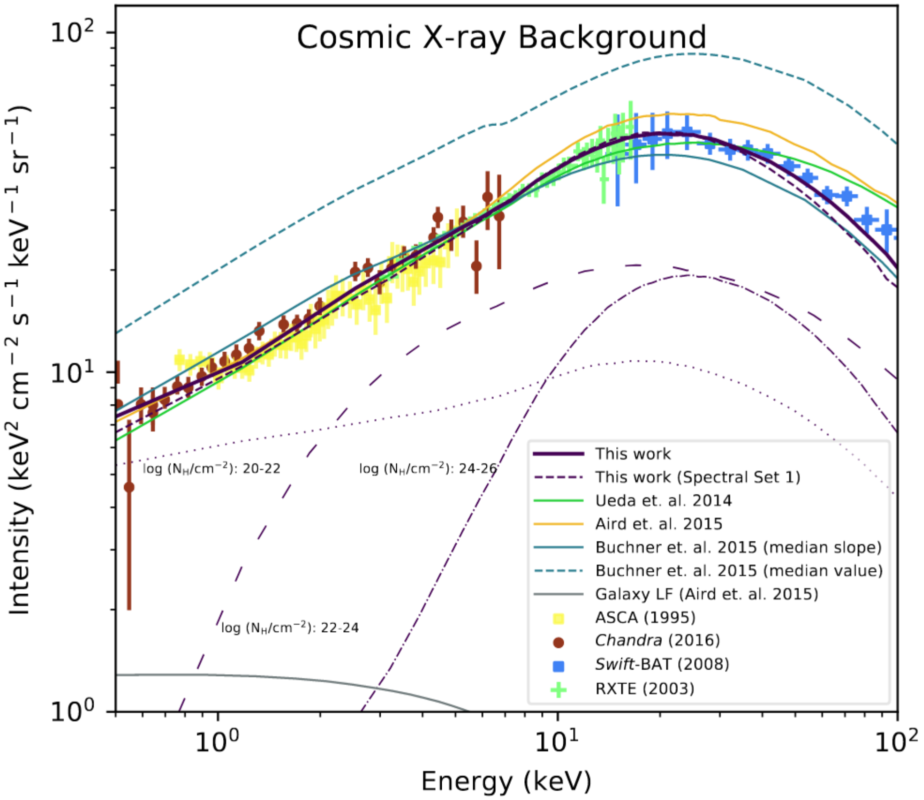


Figure 1.11: The empirical CXB from X-ray surveys compared to integrated X-ray AGN population synthesis models from Ueda et al. (2014), Buchner et al. (2015),Aird et al. (2015) and Ananna et al. (2019). Figure from Ananna et al. (2019).

results of Buchner et al. (2015) are consistent with a constant Compton-thick fraction of $\sim 35\%$. Ananna et al. (2019) find a significantly higher fraction of $\sim 56\%$ at $z = 1$.

The AGN X-ray spectral model, the XLF and the absorption function can be converted to an *AGN population synthesis model*. Integrating the emission of all AGN described by such a model should yield the Cosmic X-ray Background (CXB, e.g., Comastri et al. 1995; Gilli et al. 1999; Ueda et al. 2003; Gilli et al. 2007; Ueda et al. 2014). The CXB results from the cumulated emission of all X-ray sources and is expected to be dominated by AGN. It is mainly resolved into discrete sources at < 10 keV and provides strong constraints on the population of accreting SMBHs traced via X-ray surveys, as it is measured over large areas and across experiments. A recent figure by Ananna et al. (2019) showing the empirical CXB spectrum derived from various X-ray surveys and results integrated from the population models by Ueda et al. (2014), Aird et al. (2015), Buchner et al. (2015) and Ananna et al. (2019) is presented in 1.11.

With well-constrained XLFs, one can study the evolution of SMBHs via the black hole accretion rate density and the converted SMBH mass density (see Chapter 5 and Eq. 1.13 respectively). The bolometric luminosity of an AGN L_{bol} is a measure of its output over the entire electromagnetic spectrum. A general conversion from the X-ray luminosity L_X to the bolometric luminosity, the *bolometric correction*, can be derived for various X-ray bands from multi-wavelength AGN surveys (e.g., Hopkins et al. 2006; Lusso et al. 2012; Duras et al. 2020): $K_\mathrm{bol} = L_\mathrm{bol}/L_\mathrm{X}$. A bolometric luminosity function ϕ_bol can thus be derived from the XLF (e.g., Ueda et al. 2014). If the total bolometric output of AGN is dominated by the accretion process, the emitted radiation is related to the mass accretion rate onto the SMBH as:

$$L_\mathrm{bol} = \epsilon\,\dot{M}\,c^2, \tag{1.12}$$

where $\dot{M}$, the mass accretion rate and ϵ the radiative efficiency. The black hole spin a_* constrains the radiative efficiency *geometrically* as it determines the innermost circular stable circular orbit, which is the limit from which material falls onto the black hole without loss of energy. For a standard accretion disk, the maximum spin is $a_* = 0.998$, which translates to $\epsilon = 0.31$ (Thorne 1974). Fitting accretion disk models to AGN SEDs to determine $\dot{M}$ and simultaneously measuring their bolometric luminosities L_bol typically yields $\epsilon \sim 0.1$ (e.g., Bian & Zhao 2003; Davis & Laor 2011).

Soltan (1982) argued that the total accreted mass density of AGN throughout cosmic time and the local SMBH mass density $\rho(z = 0)$ should be commensurate if the central black holes in inactive galaxies are *dead quasars*. Generalizing Soltan's argument, the SMBH mass density at any given redshift is (e.g., Ueda et al. 2014):

$$\rho(z) = \rho(z_\mathrm{init}) + \frac{1-\epsilon}{\epsilon c^2}\int_z^{z_\mathrm{init}} dz\frac{dt}{dz}\int_{\mathrm{L_{bol,min}}}^{L_\mathrm{bol,max}} L_\mathrm{bol}\frac{d\phi_\mathrm{bol}(z, L_\mathrm{bol})}{d\log L_\mathrm{bol}}d\log L_\mathrm{bol} \tag{1.13}$$

ϵ is set to an average value. z_init indicates the redshift upper bin edge. $\rho(z_\mathrm{init})$ is the initial SMBH mass density at this starting redshift and can directly be estimated by from the space density of AGN at high redshifts (i.e. all SMBHs are assumed to live in AGN at z_init). $L_\mathrm{bol,min}$ and $L_\mathrm{bol,max}$ delimit the considered bolometric luminosity interval. The average radiative efficiency and the Eddington ratio $\lambda_\mathrm{Edd} = L_\mathrm{bol}/L_\mathrm{Edd}$ of the AGN population can be further constrained by comparing the SMBH mass density function from AGN population studies (Eq. 1.13) and the mass density function from inactive galaxies (Tamura et al. 2006; Cao & Li 2008; Li et al. 2012).

1.6 SMBH growth in the early Universe

In Sections 1.4 and 1.5 I introduced observational methods and the formalism to investigate the evolution of SMBHs. In the early Universe, their growth is signposted by powerful quasars. Here, I present this population of ultra-rare and luminous AGN, which were active when the Universe was less than a gigayear old. I will further discuss the stringent observational constraints they provide on black hole formation and evolution.

1.6.1 Quasars in the first Gyr

Over the last 20 years, quasars have been discovered at ever-increasing redshifts and well into the epoch of reionization (e.g., Fan et al. 2001; Willott et al. 2009; Mortlock et al. 2012; Wu et al. 2015; Jiang et al. 2016; Matsuoka et al. 2018a; Bañados et al. 2016, 2018; Wang et al. 2021; Banados et al. 2022). At these cosmic times ($z > 5.7$), quasars are embedded in a neutral intergalactic medium (IGM, e.g., Wise 2019, and references therein). Intervening clouds of neutral hydrogen along the line of sight to quasars imprint Lyman absorption features on the blue side of the Lyman α emission line of quasars ($\lambda_{\rm rest} \sim 1216$Å, $n = 1 \rightarrow 2$ transition). The optical depth τ is defined as the probability of scattering a Lyman α photon in a proper length element $c\,\mathrm{d}t$ and it is determined by:

$$\mathrm{d}\tau = n_{\rm HI}\sigma_{1\rightarrow 2}(\nu, \nu_\alpha)c\,\mathrm{d}t \tag{1.14}$$

$\sigma_{1\rightarrow 2}$ is the radiative cross-section for a transition from the ground state to the first excited state and $n(t)$ is the neutral hydrogen density at a cosmic time t. It is a function of the frequency of incoming photons ν in the restframe of the scattering atoms and peaks at the Lyman α frequency ν_α:

$$\sigma_{1\rightarrow 2}(\nu, \nu_\alpha) = \frac{\pi e^2}{m_{\rm e} c} f_{1\rightarrow 2}\phi(\nu - \nu_\alpha) \tag{1.15}$$

$m_{\rm e}$ is the mass of the electron, e the electric charge, $f_{1\rightarrow 2}$ the oscillator strength for the Lyman α transition and ϕ is a Voigt profile centred at ν_α.

Following Gunn & Peterson (1965):

$$\tau(z) = n_{\rm HI}\frac{\pi e^2}{m_e \nu_\alpha} f_{1\rightarrow 2} H^{-1}(z), \tag{1.16}$$

where H is the Hubble constant at a redshift z. In a standard ΛCDM, Universe this becomes (e.g., Fan et al. 2006a):

$$\tau(z) = 4.9 \times 10^5 \left(\frac{\Omega_m h^2}{0.13}\right)^{-1/2} \left(\frac{\Omega_b h^2}{0.02}\right) \left(\frac{1+z}{7}\right)^{3/2} \left(\frac{n_{\rm HI}}{n_{\rm H}}\right) \tag{1.17}$$

Here Ω_m and Ω_b are the cosmological density parameters of matter and baryons (e.g., Planck Collaboration et al. 2020). h is the dimensionless Hubble constant derived from H_0 (the Hubble constant today, $h = H_0/(100\,\mathrm{km\,s^{-1}\,Mpc^{-1}})$). $n_{\rm H}$ is the number density of hydrogen. Thus, at high redshifts, a small neutral fraction $\frac{n_{\rm HI}}{n_{\rm H}} \sim 10^{-4}$ already warrants optical thickness to Lyman α photons. Each absorber at a redshift $z_{\rm abs}$ along the line of sight generates a Lyα absorption line in the quasar spectrum at $(1 + z_{\rm abs}) \times 1216$Å. Close to the redshift of distant quasars, the increasing density of neutral hydrogen (e.g., Songaila 2004) clouds generate a dense succession of absorption features appearing as troughs in the quasar spectrum (see Fig. 1.12).

The presence of large absorption troughs has motivated the development of specific methods to sample the high-redshift quasar population. The drop in flux on the blue side

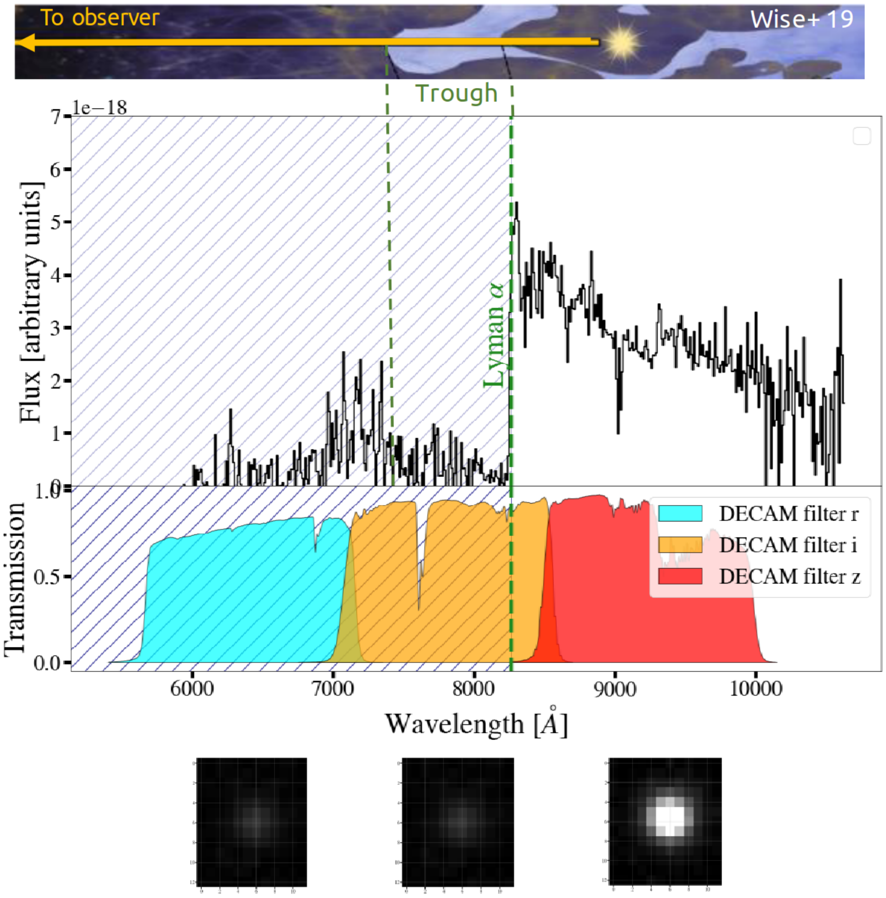


Figure 1.12: An illustration of a high-redshift absorbed quasar spectrum. The location of the Lyman α emission line is indicated by the vertical, dashed, green line on the right. The absorption trough (or sometimes *Gunn-Peterson trough*) caused by the dense neutral hydrogen at reionization is clearly visible between the dashed, green lines. The lower panel shows the optical r, i and z filters of the Dark Energy Cam (DECam, `https://noirlab.edu/science/programs/ctio/filters/Dark-Energy-Camera`). The appearance of the source in these filters is shown in the image cutouts. The upper schematic representation of the quasar is adapted from a figure in Wise (2019).

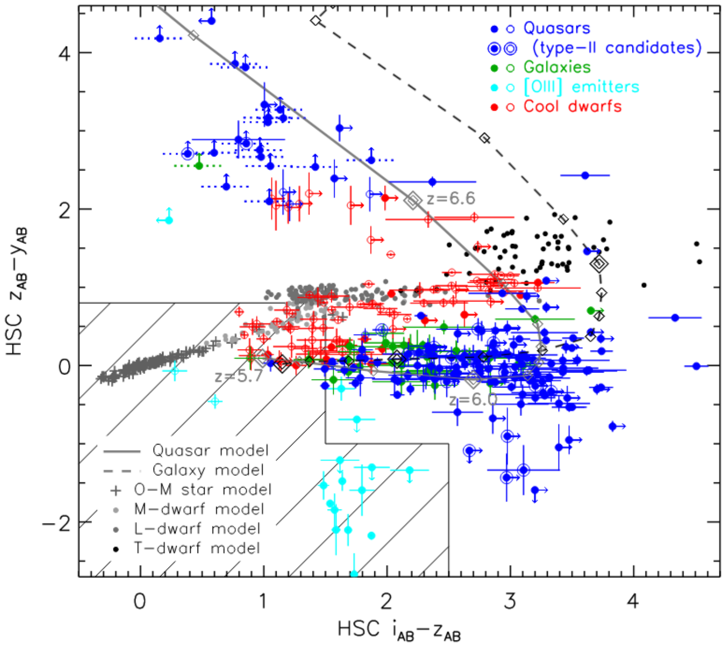


Figure 1.13: A typical optical colour selection plane ($i-z$ vs $z-y$). The selection method presented here was developed in the context of the SHELLQs survey with HSC photometry (e.g., Matsuoka et al. 2016; Matsuoka et al. 2019; Matsuoka et al. 2022). Spectroscopically confirmed quasars in this survey are shown in blue. Contaminants of the search, such as [OIII] emitters and cool dwarfs, are also marked on this figure. The black dashed and solid lines show the theoretical colour-redshift tracks of a fiducial galaxy and quasar model, respectively. Figure from Matsuoka et al. (2022) .

of the Lyman α line, at $\lambda < 1216\mathring{A}$ in restframe, causes the redshifted SEDs to *drop out* of optical, UV and sometimes NIR photometric bands. Dedicated high-redshift quasar imaging surveys, therefore, collect samples of very red dropouts over vast areas of the sky. The search for quasars at cosmic dawn is a challenging task due to their low surface density on the sky: at a fixed luminosity, the XLF shows an exponential decline with redshift (see Section 1.5). Measuring the UV high-redshift quasar luminosity function, Wang et al. (2019) find a space density at $z \sim 6.7$ of $\rho(M_{1450} < -26) = 0.39 \pm 0.11 \mathrm{Gpc}^{-3}$. In addition, the red optical and IR SEDs of high-redshift quasars resemble those of Galactic L, M and T-dwarfs, the most numerous stars in the Galaxy (e.g., Magaudda et al. 2022). The surface density of these contaminants can exceed the $z > 5.7$ quasar surface density by a factor of $10^2 - 10^3$ (e.g., Euclid Collaboration et al. 2019). Once selected, high-redshift quasar candidates are confirmed via optical and NIR spectroscopy. An optical and NIR colour selection plane by Matsuoka et al. (2022) is shown in Fig. 1.13. Spectroscopically confirmed quasars and contaminants are marked.

Classical colour selection methods have been recently extended with Bayesian model comparison (Mortlock et al. 2012) and supervised machine learning models such as random forest (e.g., Wenzl et al. 2021 and Chapter 6 of this thesis). With these methods, the redshift frontier for quasars has been continuously pushed. Dedicated optical surveys with SDSS (York et al. 2000, high-z survey: e.g., Jiang et al. 2016), the Panoramic Survey Telescope & Rapid Response System (Pan-STARR; Kaiser et al. 2010, high-z survey: e.g., Banados et al. 2022), the Dark Energy Survey (DES; Dark Energy Survey Collaboration et al. 2016, high-z survey: e.g., Reed et al. 2015), the Dark Energy Spectroscopic Instrument (DESI) Legacy Surveys (Dey et al. 2019, high-z survey: e.g., Wang et al. 2019), the Hyper Suprime Cam (HSC; Miyazaki et al. 2018, high-z survey: Subaru High-z Exploration of Low-Luminosity Quasars [SHELLQs], e.g., Matsuoka et al. 2022), the Canada–France High-z Quasar Survey (CFHQS; e.g., Willott et al. 2010) have uncovered the bulk of the currently known $z > 5.7$ quasar population. NIR infrared surveys such as the United Kingdom Infrared Telescope (UKIRT) Infrared Deep Sky Survey (UKIDSS; Lawrence et al. 2007, high-z survey: e.g., Mortlock et al. 2011), the Visible and Infrared Survey Telescope for Astronomy (VISTA) Kilo-degree Infrared Galaxy (VIKING; Edge et al. 2013, high-z survey: Venemans et al. 2013), the VISTA Hemisphere Survey (VHS; McMahon et al. 2013 high-z survey: Reed et al. 2019) and the CFHQS in the Near Infrared (CFHQSIR, Pipien et al. 2018) have increased the redshift limit of the distant quasar searches. The NIR photometry combined with MIR data from the Wide-field Infrared Survey Explorer (WISE, Wright et al. 2010) gives access to a colour space that is less affected by galactic contamination (e.g., Wang et al. 2019). More than 400 quasars at $z > 5.7$ have been discovered in optical/IR surveys to date. A compilation of all spectroscopically confirmed high-redshift quasars has been recently published by Fan et al. (2022). These discoveries are summarised in Fig. 1.14. The most distant quasars to date are J1342+0928 ($z = 7.54$, Bañados et al. 2018), J1007+2115 ($z = 7.52$, Yang et al. 2020) and J0313-1806 ($z = 7.64$, Wang et al. 2021). While X-rays are a ubiquitous signature of AGN activity, only ~ 50 of the currently known $z > 5.7$ quasars have been detected at higher energies (e.g. Nanni et al. 2017; Vito et al. 2019a; Pons et al. 2020; Wolf et al. 2021, 2023). This sample and its properties are further discussed in Chapters 4 and 5. It is, however, important to note at this stage that X-rays can, in principle, reveal heavily obscured quasars at the epoch of reionization. A candidate has been reported by Vito et al. (2019b). Vito et al. (2019a) also performed a joint X-ray spectral analysis of $z > 6$ quasars and obtained a photon index $\Gamma \sim 2.3$, a value that is steeper than in local Type 1 AGN (e.g., Nandra & Pounds 1994). Notwithstanding selection effects, this indicates that quasars at $z > 6$ generally accrete at high Eddington rates. The relative optical and X-ray emission of high-redshift quasars is extensively covered in Chapters 4 and 5.

Quasars at $z > 5.7$ have SEDs that are remarkably similar to their less distant counterparts (Shen et al. 2019; Yang et al. 2021). The chemical abundances in the BLR traced from the broad emission lines ratios in their optical/NIR spectra also show no evolution (Barth et al. 2003; Fan et al. 2004; Iwamuro et al. 2004; Jiang et al. 2007; Juarez et al. 2009; De Rosa et al. 2014; Lai et al. 2022), i.e. the BLR reaches its final metallicity early in the quasar lifetime.

As for lower redshift Type 1 AGN, the broad emission lines in quasar at $z > 5.7$ can be used in the single-epoch method to estimate the M_{BH}. In Section 1.4, we have already

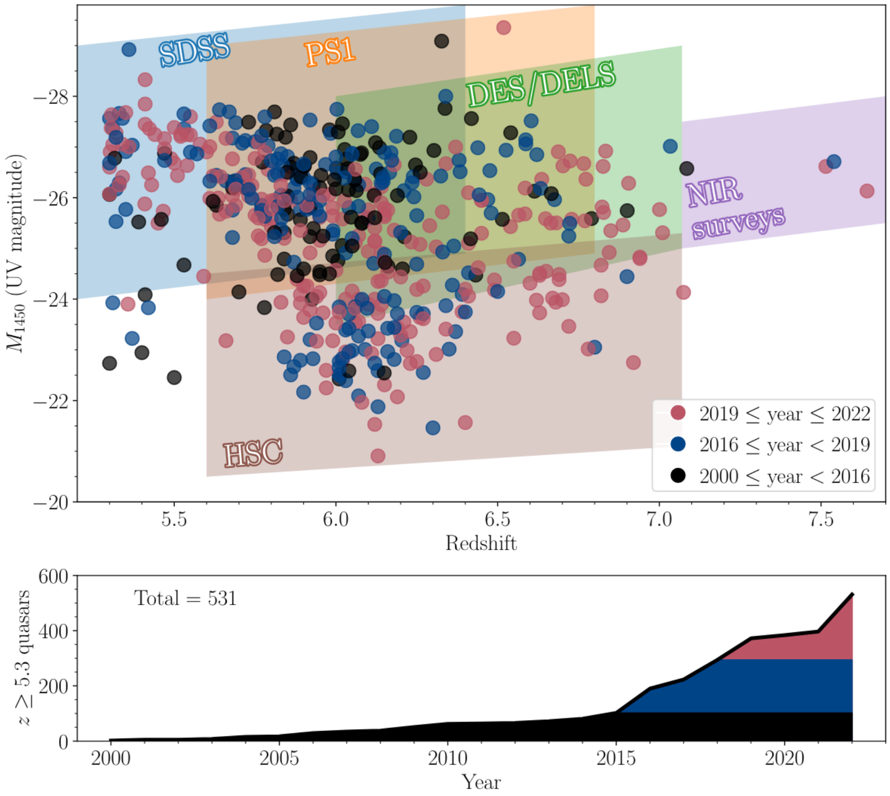


Figure 1.14: The magnitude redshift distribution of all confirmed high-redshift quasars known to date. The quasars are colour-coded by the period of discovery. Transparent polygons represent the sensitivity of various surveys. The lower panel shows the cumulative number of discoveries as a function of years. Figure from Fan et al. (2022).

seen that MgII is the most reliable line for virial mass estimation at $z > 5.7$, as the Hβ line is shifted in the MIR and the CIV line profile is affected by broad blueshifted, outflow components. MgII is shifted to the K-band at $6.0 < z < 7.6$ and $M_{\rm BH}$ measurements from this line, therefore, require NIR spectroscopic observations (e.g., Shen et al. 2019; Onoue et al. 2019; Yang et al. 2021). For MgII mass derivations, the most frequently used scaling relation was introduced by Vestergaard & Osmer (2009).

There are currently 119 quasars at $z > 5.9$ with reliable MgII-based single-epoch $M_{\rm BH}$. The median black hole mass is $1.3 \times 10^9 M_\odot$, with the most massive quasar reaching $> 10^{10} M_\odot$ (at $z = 6.33$, Wu et al. 2015, mass demographics summarised in the review by Fan et al. 2022). While the sample of high-redshift quasars with mass estimates is biased towards the very luminous end of the high-redshift quasar population, the mere existence of such high black hole masses in the first gigayear after the Big Bang challenges our understanding of SMBH seeding and evolution.

1.6.2 Early black hole assembly: Seeding and accretion modes

SMBHs discovered at $z > 5.7$ with $M_{\rm BH} \sim 10^9 M_\odot$ are but the tip of the iceberg of the high-z black hole general population: they are massive and power the most luminous quasars by selection. However, they represent a serious conundrum in terms of black hole accretion timescales. The central question is:

How did the most distant SMBHs grow to such masses in less than a gigayear?

In the standard picture, the accreted mass evolves as:

$$M_{\rm BH} = M_{\rm seed} \times \exp\left[(1-\epsilon)\lambda_{\rm Edd} f_{\rm duty} \frac{t}{t_{\rm Salp}}\right] \tag{1.18}$$

Here $M_{\rm seed}$ is the mass of the black hole seed. The duty cycle of quasars $f_{\rm duty}$ is defined by the fraction of active galaxies. The Salpeter timescale of black hole growth $t_{\rm Salp}$ is given by:

$$t_{\rm Salp} = \frac{\epsilon\, \sigma_T\, c}{4\pi G m_{\rm p}}, \tag{1.19}$$

The remnants of the first generation of stellar objects, Population III stars (hereafter Pop III), were initially proposed as natural progenitors of the SMBHs powering high-redshift quasars. Pop III stars are believed to form in dark matter *mini-haloes* ($\sim 10^6\, M_\odot$) at $z \sim 30$ (e.g., Haiman et al. 1996; Madau & Rees 2001; Tegmark et al. 1997; Yoshida et al. 2003) with initial masses $M_* \lesssim 100\, M_\odot$ (e.g., Hirano et al. 2014). The exact shape of the initial mass function of Pop III stars at the high-mass end is of interest to study the progenitors of $M_{\rm BH} \sim 10^9 M_\odot$ SMBHs at $z > 6$ and is still being intensively investigated (e.g., Toyouchi et al. 2023). How many Pop III stars are born with masses $\sim 100\, M_\odot$? In the case of Eddington-limited accretion ($\lambda_{\rm Edd} = 1$) and assuming a duty cycle $f_{\rm duty} = 1$ (e.g., Chen & Gnedin 2018) and a radiative efficiency $\epsilon = 0.1$ (see Section 1.5) Eq. 1.18 yields an e-folding of black hole growth of ~ 50 Myr. Even if we consider a massive Pop III star with a mass $M_* =\sim 100\,M_\odot$, it would need to grow ~ 805 Myr under these conditions to reach an SMBH mass of $10^9\,M_\odot$. The seeds would have formed immediately after the Big Bang. Furthermore, this scenario requires a sustained accretion with $f_{\rm duty} = 1$ and $\lambda_{\rm Edd}$ over the entire lifetime of the quasar, which is unlikely due to feedback processes in the mini-halo environments (e.g., Alvarez et al. 2009). SMBH seeding and growth models attempt to circumvent this timescale tension by modifying the seeds or the growth conditions (for in-depth reviews see Volonteri 2010; Haiman 2013; Johnson & Haardt 2016; Latif & Ferrara 2016; Inayoshi et al. 2020; Volonteri et al. 2021). In the following, I summarise some of the main theoretical growth channels:

Heavy seeds ($> \mathbf{10^5 M_\odot}$): One possible solution to the timescale issue is to increase the seed masses to $M_{\rm seed} \sim 10^5\,M_\odot$. Such massive seeds can be formed through the direct and isothermal collapse of pristine gas clouds to massive protostars and subsequently $\sim 10^5\,M_\odot$ black holes (Loeb & Rasio 1994; Begelman et al. 2006; Lodato & Natarajan 2006; Shang et al. 2010; Choi et al. 2015; Yue et al. 2017). This model requires large clouds of gas contained in primordial dark matter halos to collapse without fragmenting. The conditions in the halos need to sustain large Jeans masses for the contained gas. The so-called atomic cooling halos in which such conditions are given have the following properties:

- Relatively low metallicities ("pristine" gas): $Z/Z_\odot \leq 10^{-5}$ (Latif et al. 2016)
- Large virial temperatures. Following Barkana & Loeb (2001) the minimum required virial temperature for collapse is $T_{\rm vir} > 2.9 \times 10^3[(1+z)]^2$ K which is equivalent to a minimum halo mass:

  $$M_{\rm ACH,min} = 5 \times 10^3 \left(\frac{\Omega_m h^2}{0.15}\right)^{-1/2} \left(\frac{\Omega_b h^2}{0.022}\right)^{-3/5} \left(\frac{1+z}{10}\right)^{3/2} M_\odot \tag{1.20}$$

  At $z > 10$, $T_{vir} \sim 4 \times 10^4$ K.
- A mechanism suppressing efficient gas cooling and fragmentation. Possible candidates for this are photodissociative Lyman-Werner radiation (Haiman et al. 1997; Omukai 2001; Dijkstra et al. 2008; Latif et al. 2013; Regan et al. 2014; Dunn et al. 2018), heating via accretion and mergers (Yoshida et al. 2003) or remnant streaming velocities imprinted on the baryons at recombination (e.g., Tanaka & Li 2014; Hirano et al. 2017).

Light seeds ($\mathbf{10^{1-2} M_\odot}$): The sparse population of massive SMBHs at $z > 5.7$ could also have emerged from a small fraction of light seeds that were able to sustain Eddington-limited or super-Eddington accretion over 6 or 7 orders of magnitude on the mass scale (e.g., Inayoshi et al. 2020, for further references). This could be the case for Pop III stars located in over-massive primordial dark matter halos ($\sim 10^8\,M_\odot$, Tanaka & Li 2014). Such halos would have sufficiently deep potentials to suppress gas depletion via feedback effects. It is important to note that super-Eddington accretion has been observed for certain SMBHs (see Chapter 5), i.e., the high radiation pressure in these sources does not disrupt the accretion flow. Super-Eddington accretion can be explained by the physical effect of *photon trapping*, in which the infalling matter damps the emerging radiation (e.g., Begelman 1979; Ohsuga et al. 2005; Takahashi & Ohsuga 2015). Such an effect requires the radial gas inflow velocity to be superior to the photon diffusion velocity. In addition to high accretion rates, Pop III remnant seeds can further be grown via mergers with other compact objects (e.g., Micic et al. 2007).

Further scenarios for SMBH seeding have been explored, such as collisions in dense star clusters in the first galaxies (e.g., Begelman & Rees 1978; Omukai et al. 2008; Devecchi & Volonteri 2009). Better observational constraints will be required to understand the growth of the first black holes. The spectroscopic identification of fainter, low mass high-redshift quasars with the *James Webb Space Telescope* (*JWST*, Gardner et al. 2006) will extend the probed mass scale and allow constraints on the formation channels of the bulk of the SMBH population (Habouzit et al. 2022). A complementary search for the most massive black holes in wide surveys has the potential to exclude the light seed, Pop III formation channel (Pacucci & Loeb 2022).

This thesis

In this thesis, my aim is to shed light on the multi-wavelength properties and the demographics of the most luminous quasars, which trace the most massive and/or rapidly accreting SMBHs. The published work presented in Chapters 3, 4 and 5 is based on X-ray data from the eROSITA space-based observatory, as well as ROSAT, XMM-Newton and Chandra. These datasets were complemented with ancillary spectroscopic and imaging surveys in the optical/IR. The structure of this work is as follows (see also flowchart connecting various sections in Fig. 1.15):

- **Chapter 2: Finding AGN in the eROSITA All-Sky X-ray Survey.** I present the X-ray mission eROSITA and its main AGN surveys, as well as the optical, IR and UV counterpart identification procedure for eROSITA sources.

  *Partially based on: "The eROSITA Final Equatorial-Depth Survey (eFEDS): Identification and characterization of the counterparts to the point-like sources", Salvato, M., **Wolf, J.**, Dwelly, T., et al. (2022), Astronomy and Astrophysics, 661, A3.*

- **Chapter 3: Broad Line Region kinematics and Active Galactic Nuclei diversity.** This is a statistical analysis of spectral properties for a sample of X-ray luminous AGN at $z = 0.01-0.80$ detected with XMM-Newton and ROSAT. The spectroscopic data exploited here was taken in the context of the SDSS-IV/SPectroscopic IDentifcation of ERosita Sources (SPIDERS) programme prior to the launch of eROSITA. The BLR kinematics and structure are explored in the context of the quasar main sequence.

  *Based on: "Exploring the diversity of Type 1 active galactic nuclei identified in SDSS-IV/SPIDERS" **Wolf, J.**, Salvato, M., Coffey, D., et al. (2020), Monthly Notices of the Royal Astronomical Society, 492, 3580*

- **Chapter 4: Constraining the AGN X-ray luminosity function at $z \sim 6$.** I present the eROSITA detection of the X-ray luminous quasar J0836+0054 ($z = 5.81$). From this detection in an X-ray survey with a well-defined selection function, first spectroscopic constraints on the demographics of the X-ray sampled AGN population at $z \sim 6$ could be derived.

  *Based on: "First constraints on the AGN X-ray luminosity function at $z \sim 6$ from an eROSITA-detected quasar" **Wolf, J.**, Nandra, K., Salvato, M., et al. (2021), Astronomy and Astrophysics, 647, A5.*

- In **Chapter 5: An X-ray loud quasar at $z > 6$.** I present the eROSITA and Chandra detection of the quasar J0921+0007. It is powered by a young, rapidly accreting black hole at $z = 6.56$. The strong X-ray emission of this source stands in contrast to its relatively low optical luminosity. The consequences of this detection for SMBH population studies are investigated.

Based on: "X-ray emission from a rapidly accreting narrow-line Seyfert 1 galaxy at z = 6.56" ***Wolf J.****, Nandra K., Salvato M., Buchner J. et al., (2022), Astronomy and Astrophysics, 669, A127*

- **Chapter 6: Conclusions and outlook**. I summarise the findings in this work. As an outlook, I present a newly developed selection pipeline designed to discover new quasars in the eROSITA All-Sky Survey (eRASS).

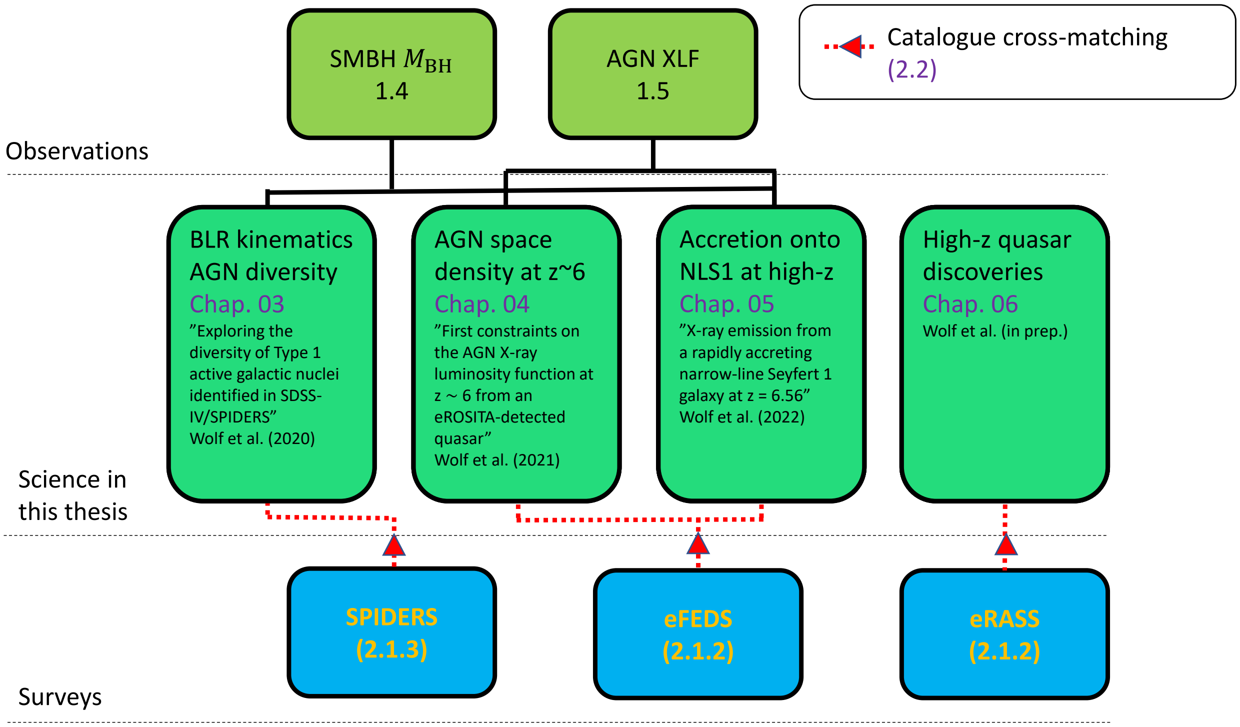


Figure 1.15: A flowchart linking the various sections of this thesis.

Chapter 2

Finding AGN in the eROSITA All-Sky X-ray Survey

The data collected with the eROSITA X-ray instrument provide the backbone of this thesis. In this Chapter, I will present the telescope and its main surveys. I will then present the basics of catalogue cross-matching, an essential procedure to identify AGN in X-ray surveys.

2.1 eROSITA: a new era in X-ray astronomy

eROSITA is a soft X-ray telescope ($0.2 - 8$ keV) that is based on technology developed at the Max-Planck-Institute for Extraterrestrial Physics (MPE, Merloni et al. 2012; Predehl et al. 2021). It was launched aboard the Spectrum-Roentgen-Gamma observatory (SRG, Sunyaev et al. 2021) from the Baikonur cosmodrome in July 2019. SRG also carries the Mikhail Pavlinsky Astronomical Roentgen Telescope X-ray Concentrator (ART-XC, Pavlinsky et al. 2021), which is complementary to eROSITA as it is sensitive to hard X-rays ($4 - 30$ keV).

As a successor to ROSAT (Truemper 1982; Voges et al. 1999; Boller et al. 2016), eROSITA has been developed to be a flagship X-ray survey instrument. Its design answers the need for a sensitive all-sky X-ray survey. The main scientific drivers for eROSITA are cluster cosmology and AGN demographics. Clusters, the most massive gravitationally bound structures in the Universe, entrap gas which emits X-rays via thermal bremsstrahlung. As clusters are located in density peaks of the large-scale matter distribution, sampling them via their X-ray emission and determining their evolution provides constraints on cosmological parameters such as the matter density parameter Ω_m (for reviews see Allen et al. 2011; Clerc & Finoguenov 2022). With a factor 20 deeper all-sky survey than its predecessor ROSAT (e.g., Boller et al. 1996), eROSITA also enables a census of the AGN population over large dynamical ranges in redshift, luminosity and obscuration: millions of AGN are expected to be detected by the end of the survey (Comparat et al. 2019; Seppi et al. 2022). With its unique combination of area and depth eROSITA also samples rare

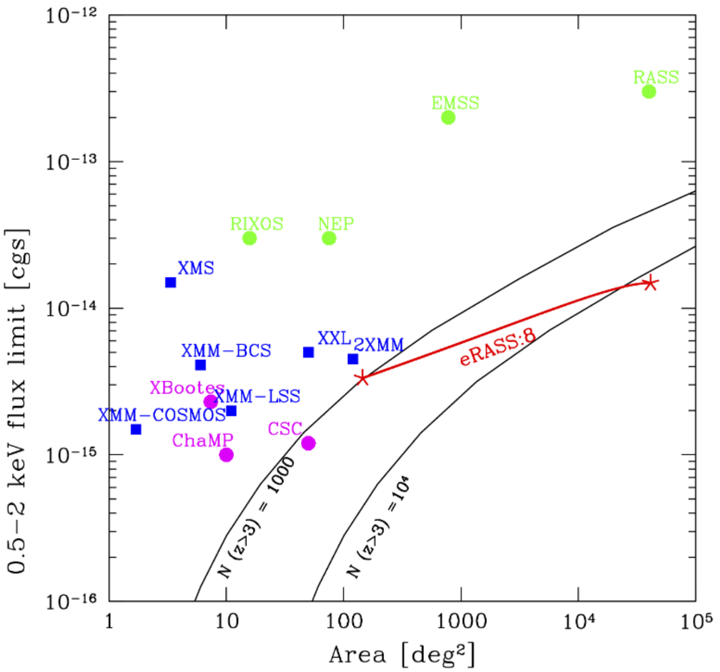


Figure 2.1: Area-sensitivity of previous X-ray surveys compared to eROSITA. Surveys performed with Chandra (purple), XMM-Newton (blue) and ROSAT (green) are shown. The prediction for the final all-sky survey (eRASS:8) is shown as a red line. The solid black lines mark the loci of number count predictions of $z > 3$ quasars extrapolated from lower-redshift XLFs. eROSITA is expected to produce a large sample of high-redshift sources. From Merloni et al. (2012).

populations such as the main objects of interest for this thesis: high-redshift quasars (see Fig. 2.1). The detection of a large population of quasars at $z > 6$ was indeed initially formulated as one of the driving scientific cases for the mission (Merloni et al. 2012).

In the following sections, I will present key technological aspects of the telescope (2.1.1) and its main surveys (2.1.2). I will also briefly present the large complementary spectroscopic programme SPIDERS (2.1.3) on which the work in Chapter 4 is based.

2.1.1 Instrument

The technical details summarised here can be found in more detail in the instrument paper by Predehl et al. (2021). eROSITA carries seven co-aligned telescope modules. Each telescope module comprises a mirror assembly associated with an individual camera assembly. The mirrors are arranged in a Wolter (type I) geometry. Wolter telescopes make use of grazing incidence optics to focus X-rays. They are composed of paraboloid and hyperboloid mirrors. Each telescope module of eROSITA contains 54 mirror shells that

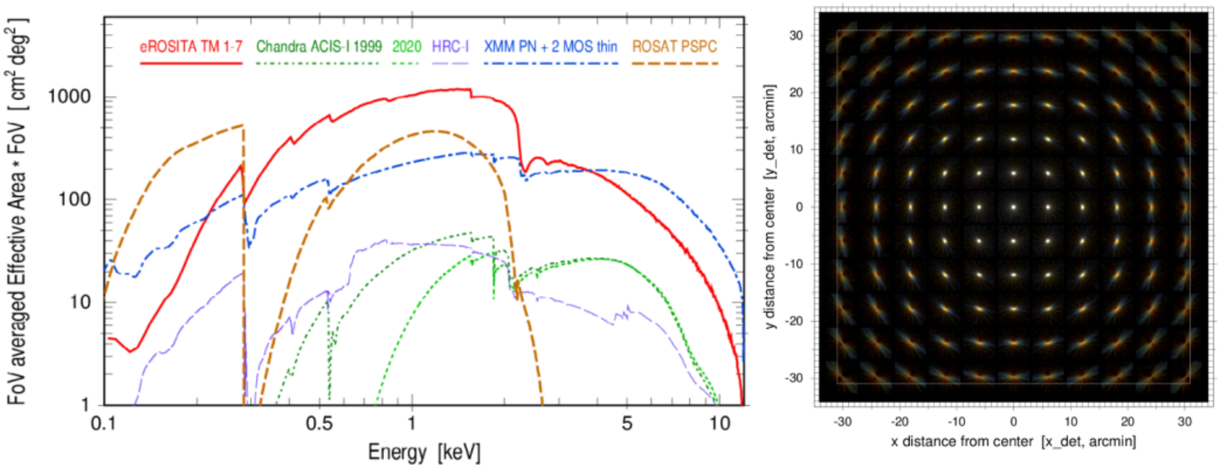


Figure 2.2: *Left panel*: Grasp of the eROSITA telescope modules (red curve) compared to other instruments. eROSITA has the largest grasp in the energy range $\sim 0.2 - 2.3$ keV. *Right panel*: Image of the PSF mapping of an eROSITA telescope module. A gradual blurring occurs off-axis. Both figures from Predehl et al. (2021).

are gold-coated on the reflecting side.

An important parameter to quantify the mirror performance is the effective area, which measures the ability of the telescope to collect X-rays over a range of energies. Alternatively, this sensitivity can be weighted by the surveyed area by computing the grasp of the telescope: the product of its field-of-view averaged effective area and the field of view. The grasp of eROSITA is compared to telescope modules of Chandra, XMM-Newton and ROSAT in Fig. 2.2 (left panel). It is unmatched in the energy range $\sim 0.2 - 2.3$ keV. At 1.5 keV the combination of the seven mirror assemblies reaches an on-axis resolution of $\sim 16.1''$ half-energy width (HEW) of the point-spread function (PSF), which corresponds to the angular diameter at which half of the reflected photons are collected. The gradual off-axis degradation of the PSF is shown in Fig. 2.2 (right panel). At an off-axis angle of $30'$, the HEW of the PSF reaches $\sim 26''$.

The mirror assemblies focus the reflected X-rays on a charge-coupled-device (CCD). These are arrays of coupled capacitors. Incoming photons are absorbed in a semiconductor substrate. The photoelectric absorption generates a charge cloud which is read out by applying an electric field. The eROSITA CCDs process X-ray photons on 384×384 pixels arranged on a 8.3 cm^2 surface. To reduce the effect of read-out streaks, the eROSITA CCDs store image frames in a time-shifted area. For details of the instrument calibration, please refer to Predehl et al. (2021, and references within).

After its launch on July 13, 2019, the SRG spacecraft was manoeuvred into an orbit around the Lagrange point L2 in the Sun-Earth system (see Fig. 2.3). From this point, it can observe in three modes: survey mode, pointing mode and field scan mode. The main survey mode consists of a rotational scanning pattern, while the pointing and field mode

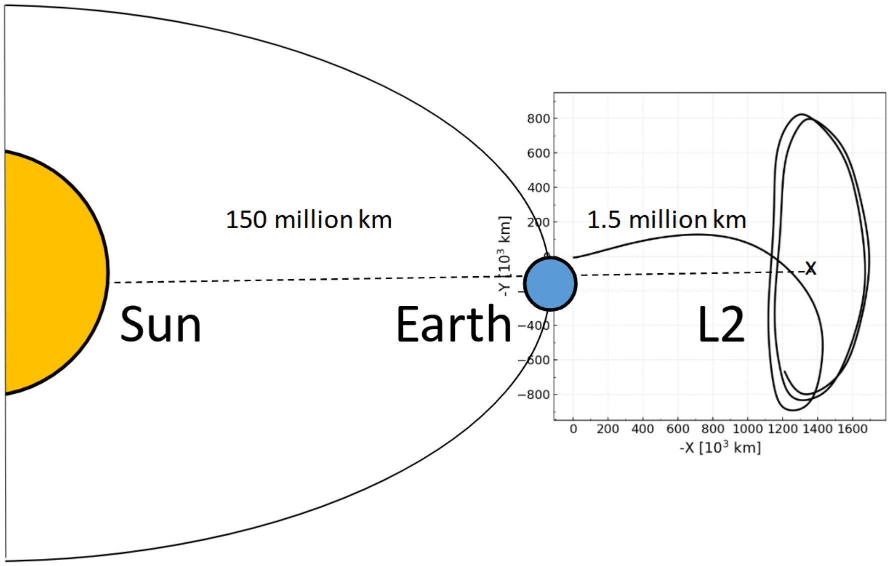


Figure 2.3: The location and orbit of the SRG observatory at L2. The black track is adapted from Predehl et al. (2021) and is a 2D projection of the trajectory of eROSITA in the first 450 days after its launch.

can sustain longer observations of positions or sky patches (up to $\sim 160\mathrm{deg}^2$), respectively.

2.1.2 Surveys: eFEDS and eRASS

eROSITA is an X-ray instrument with unique survey capabilities. The combination of its wide field of view and sensitivity makes it particularly competitive for population studies over a large area of the sky. For the work presented in this thesis, two surveys were of particular importance: the eROSITA Final Equatorial Depth Survey (eFEDS) and the successive iterations of the all-sky survey, eRASS.

eFEDS

eFEDS is a mini-survey, designed to demonstrate and test the instrument capabilities during the mission's performance verification phase. It was defined as a contiguous field of $\sim 140\,\mathrm{deg}^2$ arranged in 4 tiles of $\sim 35\,\mathrm{deg}^2$. Fig. 2.4 shows the final exposure map in the 0.2–2.3 keV band. It was observed in field scan mode in November 2019 for a total exposure depth of $\sim$ 2.2 ks, or 1.2 ks vignetting-corrected exposure. This corresponds to a soft flux limit of $F_{0.2-2.3\,\mathrm{keV}} \sim 10^{-14}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$. The source detection is performed in the 0.2-2.3 keV band using a sliding box algorithm from the eROSITA Science Analysis Software System (eSASS) and the resulting X-ray source catalogue is presented in Brunner et al. (2022). The location of eFEDS in the sky was selected due to the extensive coverage by ancillary imaging and spectroscopic surveys. eFEDS benefits from the coverage of optical/NIR photometry from HSC, the DESI Legacy Imaging surveys, the Kilo-Degree

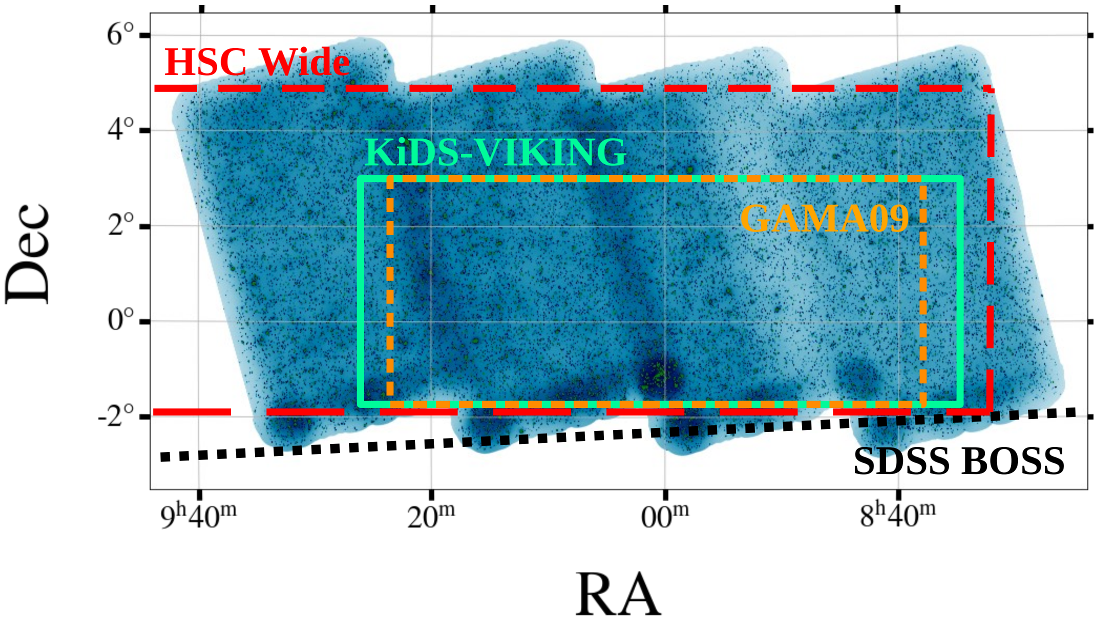


Figure 2.4: The 0.2-2.3 keV source map of the eFEDS field. The footprints of ancillary surveys are marked (see text for more details). Reproduced from Brunner et al. (2022).

Survey-VIKING survey (KiDS-VIKING, Wright et al. 2019). In addition spectroscopic data from SDSS (see further details in 2.1.3), Galaxy And Mass Assembly (GAMA, Driver et al. 2009), the Large sky Area Multi-Object fibre Spectroscopic Telescope (LAMOST, Zhao et al. 2012) and WiggleZ (Drinkwater et al. 2010). The eFEDS X-ray catalogue is used in Chapters 4 and 5 of this thesis.

eRASS

The most important data products of eROSITA will be collected over a succession of scans of the entire sky: eRASS. The instrument's rotation, its orbit around L2 and the Earth's revolution define the survey strategy. As the instrument rotates on its axis and is carried along with L2, projected vertical bands (or great circles) of the celestial sphere are observed. Following Predehl et al. (2021), the rotation of the spacecraft defines the scan rate, which is the advancement along one vertical band of the sky ($0.025\,\mathrm{deg\,s^{-1}}$). Through a combination of the Earth's orbit around the Sun, the orbit of the instrument around L2 and the instrument's revolution, the observed projected circles on the celestial sphere are progressively shifted until the full sky is surveyed every six months. Following the above survey strategy, two polar regions of deeper exposure are generated: the southern and northern ecliptic poles. An average exposure of ~ 275 s is reached for each eRASS while at the ecliptic poles it can reach > 1 ks. The resulting typical point-source sensitivity in the most sensitive $0.2 - 2.3$ keV band per single eRASS is $\sim 4.5 \times 10^{-14}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$. Each full scan of the sky, or eRASS, is performed within ~ 6 months, with 8 consecutive eRASS planned. The total all-sky average exposure after 8 scans will be ~ 2.2 ks.

The X-ray source catalogues are generated with eSASS for each scans individually and for the cumulative survey. Generating successive, single eRASS catalogues allows, for instance, the monitoring of long-term variability, while the deeper cumulative catalogue increases the survey sensitivity. Each survey is split into two halves owned and analysed by the eROSITA-DE and RU consortia, respectively. The final catalogue is expected to contain about 3 million sources. A simulation of X-ray detections in the first eRASS (eRASS1[1]) was presented by Seppi et al. (2022).

The all-sky survey was initiated in December 2019. As of January 2023, four out of eight eRASS have been completed.

2.1.3 SPIDERS

A critical piece of information for the characterisation of AGN in catalogues of X-ray sources is a measure of their redshift. The most accurate redshifts can be obtained by spectroscopy. Currently, the spectroscopic completeness of X-ray-selected AGN samples is low. For example a few thousand AGN have been spectroscopically identified in the northern XMM-XXL field (Menzel et al. 2016), one of the largest spectroscopic samples published so far, but the completeness is only 32 %.

Determining the exact astrometric position of the X-ray emission is challenging, as the positional uncertainties of X-ray detections are usually on the order of a few arcseconds. X-ray sources must therefore be matched to optical/IR counterparts with sub-arcseconds positional uncertainties. Knowledge of the restframe UV, optical and IR (UVOIR) SED of AGN is furthermore critical to redshift estimation. State-of-the-art techniques of catalogue cross-matching are presented in Section 2.2.

SPIDERS is a large SDSS (Gunn et al. 2006; Smee et al. 2013) spectroscopic follow-up campaign for X-ray sources embedded in the SDSS-IV (Blanton et al. 2017) and SDSS-V (Kollmeier et al. 2017) projects. It is divided into two sub-projects: one for AGN (Dwelly et al. 2017; Coffey et al. 2019; Comparat et al. 2020) and another for galaxy clusters (Clerc et al. 2016). The present summary is limited to the selection and observation of AGN targets.

SPIDERS was started as part of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey programme (eBOSS, Dawson et al. 2016), before the launch of eROSITA. During this early period, the scope of the SPIDERS programme was demonstrated by providing spectroscopy for X-ray sources detected in XMM-XXL (Menzel et al. 2016), the first and second ROSAT all-sky surveys (1RXS and 2RXS, Voges et al. 1999; Boller et al. 2016) and the XMM-Slew 1 and 2 surveys (XMMSL1,XMMSL2). These are *Tier-0* surveys. The target selection of 1RXS and XMMSL1 sources is presented by Dwelly

---

[1] Throughout this work, I will adopt the standard eROSITA nomenclature for the all-sky surveys: eRASSN for the N-th individual all-sky survey and eRASS:N for the cumulative surveys from 1 to N.

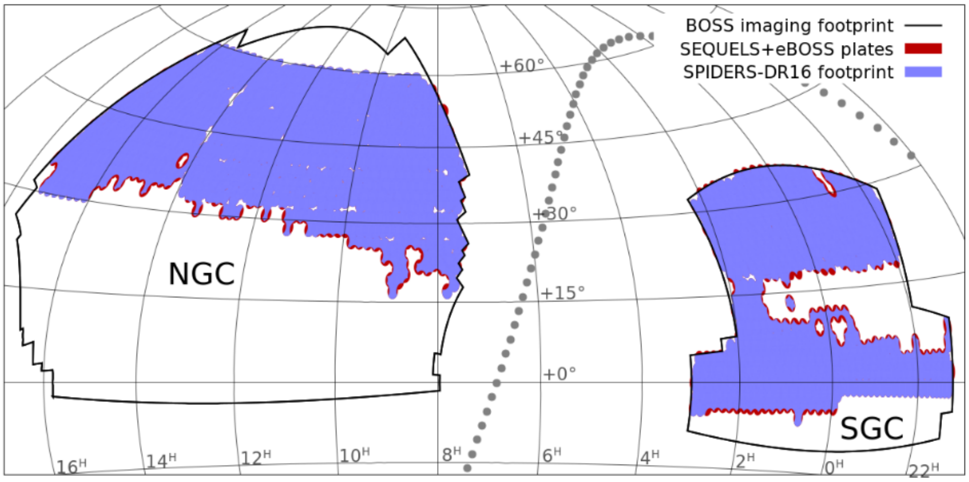


Figure 2.5: The footprint of the SPIDERS survey in SDSS DR16 (purple region). The solitd black line delimits the total BOSS imaging footprint. Additionally, the underlying red region shows the combined Sloan Extended QUasar, ELG and LRG Survey (SEQUELS) and eBOSS plates. From (Comparat et al. 2020)

et al. (2017). The 2RXS and XMMSL2 targeting is detailed by Comparat et al. (2020). For the determination of the target, these authors have made use of the ubiquitous and strong MIR emission of AGN (e.g., Elvis et al. 1994; Assef et al. 2013). X-ray sources were cross-matched to the AllWISE release (Cutri et al. 2013) of WISE data using the Bayesian cross-matching algorithm NWAY (Salvato et al. 2018). Dwelly et al. (2017) tested the reliability of their target selection using a sample of a well-characterised, bright sources selected in the 3XMM catalogue (Rosen et al. 2016). After the release of a value-added SPIDERS AGN catalogue in the fourteenth SDSS data release (SDSS DR14 Abolfathi et al. 2018) by Coffey et al. (2019), the SPIDERS data were released as part of the SDSS DR16 (Ahumada et al. 2020). All the targets were selected within a 5128.9deg^2 sub-region in the Baryon Oscillation Spectroscopic (BOSS) footprint (see Fig. 2.5). A list of 11,092 spectra was observed in total Comparat et al. (2020). The redshifts were derived automatically and subsequently visually inspected following Dwelly et al. (2017). After the launch of eROSITA in 2019, SDSS-IV/SPIDERS spectra were collected for eRASS1 sources and will be presented by Merloni et al. (in prep.). The collection of SPIDERS spectra continues within SDSS-V under the umbrella of the Black Hole Mapper project (for more details see Kollmeier et al. 2017).

2.2 Multi-wavelength counterpart identification of eROSITA sources

Some of the results and concepts presented in this section were published in "The eROSITA Final Equatorial-Depth Survey (eFEDS): Identification and characterization of the counterparts to the point-like sources" Salvato, M., ***Wolf, J.****, Dwelly, T., et al. (2022), Astronomy & Astrophysics, 661, A3, to which I have contributed the text for section 3.1. A slightly adapted version of this contribution is included here.*

Distance information via the redshift is essential to the physical characterisation of an astronomical source. The redshift of X-ray sources can be obtained via three methods, whose applicability depends on the intrinsic obscuration of the source and the presence of ancillary UVOIR data. In the case of heavily obscured sources, the source may not appear in multi-wavelength surveys at longer wavelengths, but the redshift can be estimated from the X-ray spectrum (Simmonds et al. 2018; Sicilian et al. 2022). This technique requires sufficiently good photon counting statistics and is therefore inapplicable for faint sources such as those at higher redshifts. The two main techniques for non-heavily obscured sources are indirect, as they require a cross-match to a counterpart in an ancillary survey. If only imaging data are available, a photometric redshift z_{phot} can be determined using broad or narrow band photometric measurements of the UVOIR counterparts (for a review see Salvato et al. 2019b). Pre-existing or follow-up spectroscopy, as provided in the SPIDERS programme, can yield more reliable spectroscopic redshifts. At any rate, the need for a distance measurement has motivated the development of sophisticated cross-matching techniques, which take into account astrometric positions, positional uncertainties and independent source properties (e.g., photometry, morphology ...). Cross-matches can be performed between two or more catalogues and require a probabilistic treatment (for a review, see Pineau et al. 2017).

eROSITA has PSF HEW of $\sim 16''$, which is large with respect to typical positional uncertainties for point sources in optical imaging survey ($< 1''$). Multiple possible counterparts can therefore exist within the error radius around an eROSITA centroid. In some cases, the X-ray emission from multiple neighbouring sources can be blended. Fig. 2.6 illustrates the challenge of cross-matching sources with relatively large positional uncertainties. An eFEDS source from the X-ray point source catalogue (ID 6931, Brunner et al. 2022; Salvato et al. 2022) is shown. Within a 3σ positional uncertainty ($\sim 12''$) there are seven potential counterparts from the DESI Legacy Survey Data Release 8 (LS8). It is a non-trivial task to determine the optical counterpart of the X-ray source. For the eROSITA catalogues, one of the techniques applied to determine UVOIR counterparts is called `NWAY` (Salvato et al. 2018). I will present the statistical concept of this cross-matching technique and how I have contributed to extending it with machine learning models to determine counterparts for the X-ray point source catalogue eFEDS (Salvato et al. 2022).

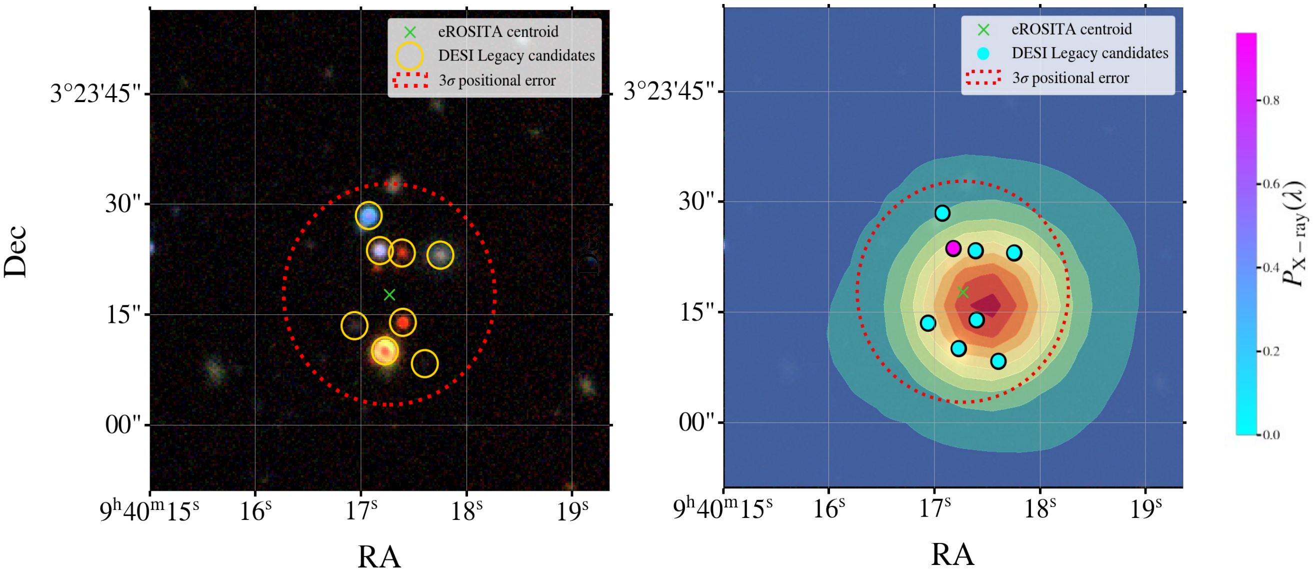


Figure 2.6: *Left:* The centroid and 3σ positional uncertainty of the source ID 6931 from the eFEDS catalogue (Brunner et al. 2022) are shown in green. The RGB image is extracted from the LS8 survey. The blue squares mark detected optical sources in LS8 located within the astrometric error circle of the source. A red square marks the actual candidate. It is a non-trivial task to determine this counterpart. *Right:* The same cutout is shown with contours from the eFEDS image. The potential candidates are marked according to their random forest-assigned probability of emitting in the X-ray (see more details on its calculation in 2.2.2).

2.2.1 Bayesian catalogue cross-matching

Cross-matching two catalogues is the simplest and most frequently encountered case in practice for finding multi-wavelength counterparts. One approach to this task is the Likelihood Ratio technique (Sutherland & Saunders 1992), in which a primary and a secondary catalogue are set. For a match candidate in the secondary catalogue, the ratio of the likelihood of being a true counterpart over the likelihood of being a chance alignment is computed. This calculation accounts for the astrometry, positional uncertainties, magnitudes and source densities (for an example of application to X-ray data, see Brusa et al. 2009)

The two-way cross-match can be put into a Bayesian framework in order to account for more complex priors. To this effect, each possible association is considered between the two catalogues to match. Each of these potential associations is characterised by the distance between the centroids of the two sources, normalised by the total positional uncertainties D. Following Pineau et al. (2017), one then computes the histogram of the number of all astrometric matches considering both true associations (T) and chance alignments (F). The probabilities of matches being real given an error-normalised distance D, $p(D \mid T)$, are distributed as a Rayleigh distribution. The probability of false association, $p(D \mid F)$, is linear in D. The histogram of matches $h(D)$ is then given by:

$$h(D) = (N_{\mathrm{tot}} - N_{\mathrm{F}}) \times p(D \mid T) + N_{\mathrm{F}} \times p(D \mid F) \tag{2.1}$$

Here N_{tot} is the total number of matches and N_F the total number of chance alignments. Let $N_T = N_{\mathrm{tot}} - N_F$ be the total number of true matches. With the knowledge of $p(D \mid T)$, $p(D \mid F)$ and N_{tot}, N_F and thus N_T can be derived by fitting $h(D)$.

The probability of a match being true is:

$$p(T \mid D) = \frac{N_T \times p(D \mid T)}{h(D)} = \frac{N_T/N_{\mathrm{tot}} \times p(D \mid T)}{h(D)/N_{\mathrm{tot}}} = \frac{p(T) \times p(D \mid T)}{p(D)} \tag{2.2}$$

where the following priors have been defined: $p(T) = N_T/N_{\mathrm{tot}}$ and $p(D) = h(D)/N_{\mathrm{tot}}$. The probability of a true match given the normalised distance is thus expressed in terms of Bayes' theorem. The Bayesian counterpart identification framework has been extended to multiple catalogue matching, i.e. when more than two catalogues are involved (Budavári & Szalay 2008; Pineau et al. 2017). With `NWAY`, Salvato et al. (2018) provided an algorithmic structure based on the Budavári & Szalay (2008) formalism, extending it to allow one or more catalogues not to be considered in certain matching instances. This feature is particularly handy in the case of catalogues of different photometric depths, which is nearly always the case in practice.

The following description of the `NWAY` formalism is based on Appendix B5 of Salvato et al. (2018).

Bayes' theorem yields the posterior probability of an association given some data θ as:

$$P(H \mid \theta) \propto P(H) \times P(\theta \mid H), \tag{2.3}$$

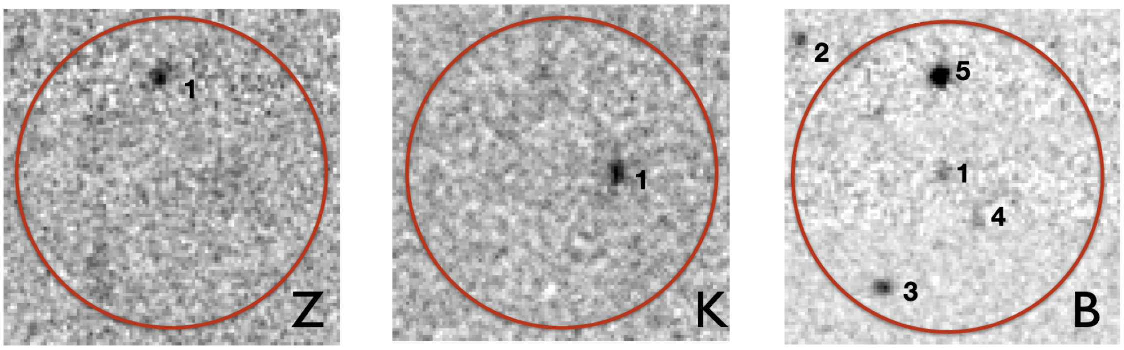


Figure 2.7: Cross-matching an X-ray source to multiple ancillary catalogues. The counterpart candidates are not detected in each survey. For instance, counterpart 1 in the survey B, is not detected in K and Z. Example from Brusa et al. (2010). Courtesy of Dr. Mara Salvato.

where $P(H)$ is the completeness-corrected prior for astrometric co-incidence and $P(\theta \mid H)$ the likelihood. The computation of the likelihood in Salvato et al. (2018) accounts for possible missing counterpart candidates in some catalogues. In practice, this is the case when catalogues of different depths are matched. Fig. 2.7 presents a typical case in which an X-ray source from the XMM-Newton COSMOS survey (Brusa et al. 2010) is matched to three ancillary catalogues in different bands. The images centered on the X-ray centroid display potential counterpart candidates in each band (or survey). Certain candidates are only visible and detected in a subset of bands. The details of the calculation of the posterior probabilities and the Bayes factor can be found in Budavári & Szalay (2008) and Salvato et al. (2018).

The Bayesian framework allows the likelihood to incorporate information on non-astrometric features such as photometry. If for the parameter set $\theta = (D, \phi)$, D denotes the error-normalised distance, and ϕ is an independent feature (e.g. a broad-band magnitude), the likelihood for the hypothesis H can be split as:

$$P(\theta \mid H) = P(D \mid H) \times P(\phi \mid H) = P(D \mid H) \times \frac{p_{\mathrm{true}}(\phi)}{p_{\mathrm{field}}(\phi)}, \tag{2.4}$$

where $p_{\mathrm{true}}(\phi)$ is the probability of a true counterpart having feature ϕ and $p_{\mathrm{field}}(\phi)$ the probability of the general field population having feature ϕ. In Salvato et al. (2018), AllWISE counterparts to 2RXS and XMMSL2 sources are determined with `NWAY`. The likelihood incorporates photometric information in the form of a colour-magnitude criterion $W1-W2$ vs $W2$ (colour and magnitude in first and second broad WISE bands). The ratio of p_{true} and p_{field} is derived from a training sample extracted from 3XMM. The ratio is obtained by computing density ratios in bins of the $W1 - W2$ vs $W2$ plane (see Fig. 2.8). Accounting for photometric information can break astrometric degeneracies.

Each `NWAY` cross-match solution to a source of the primary catalogue is summarised by the two main statistics: `p_any` and `p_i`. `p_any` is the probability for a primary source to

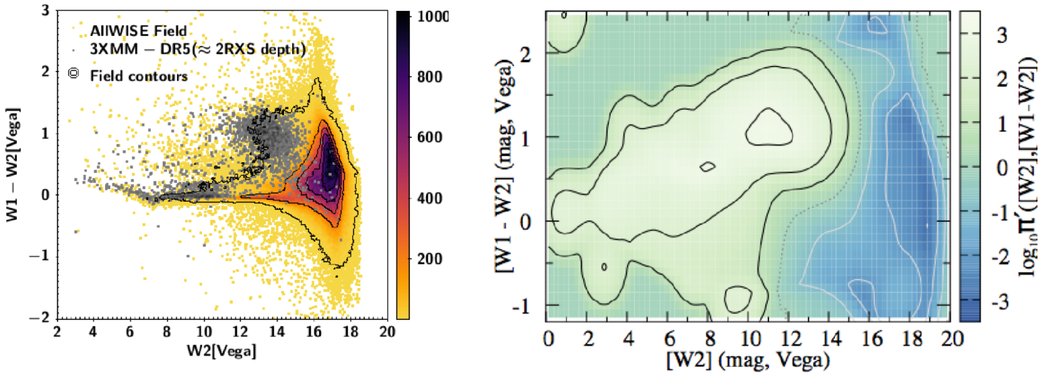


Figure 2.8: The AllWISE colour-magnitude prior from Salvato et al. (2018). *Left panel*: The distribution of the 3XMM training sample and field sources in the AllWISE colour magnitude plane. The field sources populate a distinct region of the parameter space. *Right panel*: A prior is constructed from the ratio of the density 3XMM sources to field sources in the binned colour-magnitude plane. Figures from Salvato et al. (2018).

have any counterparts. If H_0 is the hypothesis that the primary source has no counterparts in any auxiliary catalogue and H_i the hypothesis that the primary source has a counterpart in catalogue i, then Salvato et al. (2018) define:

$$\mathtt{p_{any}} = 1 - \frac{P(H_0 \mid \theta)}{\sum_i P(H_i \mid \theta)} \tag{2.5}$$

For each entry in the primary catalogue, all the posterior probabilities $P(H_i \mid \theta)$ are normalised to unity. For any match instance `p_i` is the relative posterior probability:

$$\mathtt{p_i} = \frac{P(H_i \mid \theta)}{\sum_{i>0} P(H_i \mid \theta)} \tag{2.6}$$

To select high-confidence counterparts, thresholds on `p_any` and `p_i` can be derived from simulations (e.g., Salvato et al. 2022).

2.2.2 Random forest-trained priors for `NWAY`

The idea of a colour-magnitude "prior"[2] as presented in Salvato et al. (2018) can be generalised to higher dimensional parameter space. For instance, using machine learning models, it is possible to map from N-dimensional feature space to a probability of being an X-ray source. The models are trained on well-defined samples of X-ray sources and non-X-ray field sources in the survey of interest. It is worth stressing that such a model can provide

---

[2] We call models providing a probability of being X-ray emitters "priors" to `NWAY`, as they enter as a priori information in the ultimate matching process.

the probability of X-ray emission for the entirety of the ancillary catalogue. In the `NWAY` context, this probability can be related to p_{field} and p_{true} in Eq. 2.4.

In Salvato et al. (2022)[3] we have identified LS8 counterparts for eFEDS point-like sources. We have made use of Eq. 2.4 to incorporate non-astrometric information in the computation of posterior matching probabilities. We have trained a random forest classifier (`sklearn` implementation, Pedregosa et al. 2011) on a large number of features to reliably map the available LS8 information to real X-ray sources and real field objects. The trained classifier is then used to predict the probability of all counterpart candidates being X-ray emitting. This probability is directly used to compute $P(\theta \mid H)$.

We have used a 3XMM/LS8 training sample containing confirmed 3XMM X-ray sources and their secure LS8 counterparts. In addition, this sample contains non-X-ray emitting LS8 field sources (sample selection detailed in the appendix of Salvato et al. 2022). The training features extracted from this sample are:

- `flux_*/mw_transmission_*`: deredenned fluxes in LS8 bands: $g, r, z, W1, W2$.
- `gaia_phot_*_mean_mag`: Gaia photometry in the bands $G, G_{\text{bp}}, G_{\text{rp}}$ as listed in LS8
- `snr_*`: signal-to-noise (S/N) ratio for $g, r, z, W1, W2, G, G_{\text{bp}}, G_{\text{rp}}$ photometry
- $\sqrt{\texttt{pmra}^2 + \texttt{pmdec}^2}$: Gaia proper motion as listed in LS8
- `parallax`: Gaia parallax as listed in LS8
- $\texttt{g} - \texttt{r}$, $\texttt{r} - \texttt{z}$, $\texttt{z} - \texttt{W1}$, $\texttt{r} - \texttt{W2}$: deredened colours

In this training sample, X-ray sources are flagged as target class "1" and field objects as target class "0". About 15% of the 3XMM training samples (61821 sources) are randomly extracted for testing purposes and not further considered in the training procedure. The baseline random forest model is composed of 200 decision trees, allowing decision split points if at least 8 samples are left in each branch. All of the 22 features can be used for the decision-tree building, which makes use of bootstrap samples of the training set. The training sample is highly imbalanced by construction since the field objects strongly outnumber the X-ray sources. We, therefore, opted for a weighting scheme, automatically adjusting weights of training examples for the class imbalance. The trained model is evaluated on the test set and a value p_{X-ray} is predicted for each testing entry. A confusion matrix summarising the test run results is presented in Fig. 2.9.

We note that the cut in the class prediction for the presented confusion matrix is made at $p_{X-ray} = 0.50$, where p_{X-ray} is the predicted probability that a counterpart candidate is X-ray emitting. Since `NWAY` uses the continuous predicted probability as modifying factor for the likelihood $P(\theta \mid H)$, real counterparts with rare or untypical photometric features, i.e with $p_{X-ray} \ll 0.5$, may still be selected by the algorithm if the astrometric configuration favours them. We obtain a high recall fraction of $2585/(2585 + 457) = 85\%$, while the

---

[3] This section is a modified version of the text I wrote for the **section 3.1** of Salvato et al. (2022).

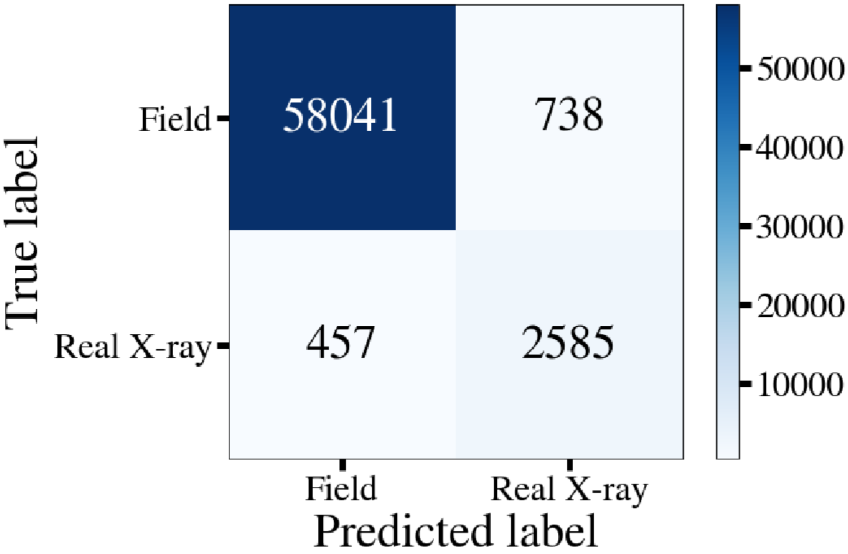


Figure 2.9: Performance of the random forest prior for `NWAY` matches between eFEDS and LS8. The figure shows the confusion matrix resulting from evaluating the trained random forest on an independent test set. It provides the number of real field and X-ray sources the model correctly or wrongly classified. The random forest model is then used to compute the prior ratio in Eq. 2.4. Figure from Salvato et al. (2022).

fractional leakage of contaminating field objects remains low: $738/(738 + 58041) = 1\%$. Using the trained model, we predict p_{X-ray} for all LS8 sources in the eFEDS field. We then run the `NWAY` matching procedure using the ratio $p_{X-ray}/(1-p_{X-ray})$ for $P(\theta \mid H)$. This is done by adding p_{X-ray} as a column to the LS8 catalogue and activating it as a prior column in `NWAY` with the `-mag` option. To calibrate the `p_any` thresholds, this cross-matching setup was applied to a simulated sample of eFEDS sources with associated LS8 counterparts (please refer to Salvato et al. 2022 for the construction of this simulated sample). The output of `NWAY` can be directly compared to the truth value, i.e. the true counterpart. Defining *purity* as the fraction of sources with a correct association and *completeness* as the fraction of counterparts identified given a `p_any` threshold, we obtained the purity and completeness curves presented in Fig. 2.10. Our machine Learning boosted Bayesian counterpart identification method reaches a purity and completeness of 95% at a low `p_any` threshold of ~ 0.035.

After the cross-matching procedure, we can proceed to identify AGN among the set of secure counterparts to the X-ray catalogue. This can be done in three different ways:

- Spectroscopic identification if spectroscopic data are available
- Colour-magnitude selection (Salvato et al. 2018, 2022)
- SED template fitting Salvato et al. (2019b)

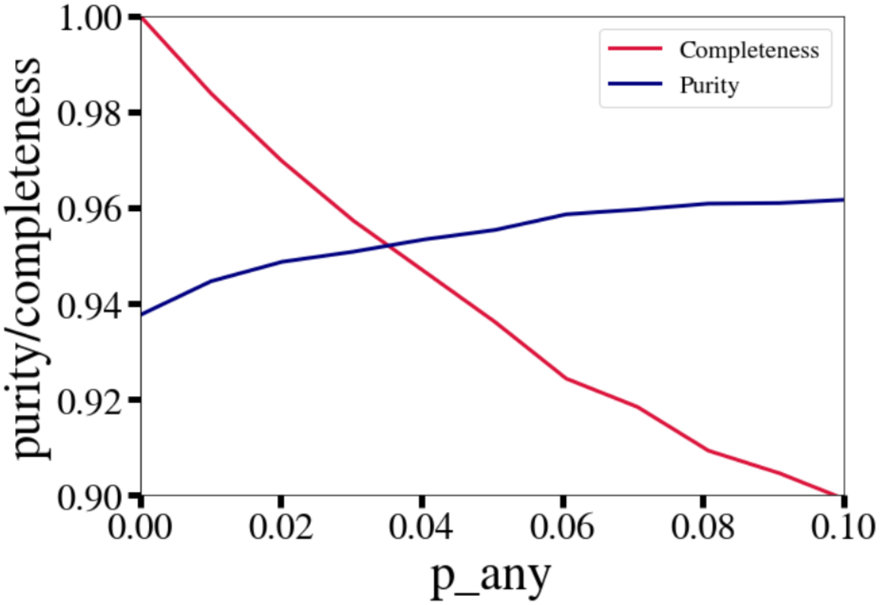


Figure 2.10: Evaluation of the `NWAY` setup on a simulated test sample. Purity is defined as the fraction of correctly associated counterparts and completeness the ratio of recovered counterparts. Both of these metrics are evaluated at threshold values `p_any` (i.e.,$\geq$ `p_any`). The completeness and purity are at 95% at `p_any`$= 0.035$. Figures from Salvato et al. (2022).

In summary, the technique presented in this Chapter extends the Bayesian catalogue cross-matching framework `NWAY` with high-dimensional machine learning trained priors. Reliable counterparts are selected based on the astrometric configuration, and the probability of emitting X-ray radiation derived from non-astrometric properties. The use of machine learning models such as random forests allow to take advantage of the wealth of information in the auxiliary catalogues. Tests on simulations suggest that this method yields highly pure and complete counterpart catalogues. `NWAY` has been used in combination with random forest-trained priors in Chapters 4 and 5 to identify high-redshift quasars among the counterparts of the eFEDS X-ray source catalogue.

Chapter 3

Broad Line Region kinematics and Active Galactic Nuclei diversity

The results presented in this Chapter were published in "Exploring the diversity of Type 1 active galactic nuclei identified in SDSS-IV/SPIDERS" ***Wolf, J.****, Salvato, M., Coffey, D., et al. (2020), Monthly Notices of the Royal Astronomical Society, 492, 3580.*

3.1 Summary

In this Chapter, we perform a principal component analysis (PCA) of the spectral features extracted from high signal-to-noise ratios (S/N) optical spectra for a sample of X-ray selected Type 1 AGN. The X-ray selection enables the construction of a relatively unbiased sample of AGN since the host X-ray-related processes are believed to be particularly weak compared to high-energetic AGN emission. The Type 1 AGN in our sample were detected by ROSAT (Boller et al. 2016; Voges et al. 1999). Their multi-wavelength counterparts were determined in Salvato et al. (2018) and spectroscopically followed-up in the SDSS-IV/SPIDERS program (Dwelly et al. 2017; Comparat et al. 2020). These sources were presented in Coffey et al. (2019, henceforth C19), where they were selected according to the width of their broad lines (FWHM $> 1000\,\mathrm{km\,s^{-1}}$). With respect to a previous PCA-based study on X-ray selected AGN (Grupe 2004), the sample used here is a factor ~ 20 larger and significantly extends the probed redshift and luminosity ranges (z $\sim 0.01 - 0.80$ and $L_{0.1-2.4\,\mathrm{keV}} \sim 2.0 \times 10^{41} - 1.0 \times 10^{46}\,\mathrm{erg\,s^{-1}}$). The PCA on the sources' optical and X-ray features allows us to determine the markers of spectral diversity in our sample. We specifically investigate the role played by λ_{Edd} and M_{BH} for the total variance in our data. The asymmetry of the broad Hβ contributes strongly to the orientation[1] of EV1 through the chosen spectral parameter space. We thus follow in the footsteps of Zamfir et al. (2010, henceforth Z10) and connect our global statistical results to the shape of the broad Hβ emission line, which is discussed in the context of BLR stratification and low-ionisation

---

[1] Here, *orientation* refers to the linear dependency of the Eigenvector on the investigated features.

outflows.

The Chapter is organized as follows: In Section 3.3, we briefly present the SPIDERS AGN value-added catalogue of C19 and how we created our sub-sample. We also outline how the additional properties, which were not provided in C19, were determined. In Section 3.4, we describe the emission line asymmetry index distribution. The core results of the direct correlation analysis and PCA on our sample are presented in Section 3.5. We demonstrate how the derived physical parameters λ_{Edd} and M_{BH} scale with the obtained principal components in Section 3.6. The distribution of our sample in the EV1 plane and the possible presence of a distinct kinematic BLR region is investigated in Section 3.7. We divided the sample into two distinct subsets according to the sign of their Hβ asymmetry indices. We detect an interesting contrast in the scaling relation between source luminosity and the equivalent width of the iron emission from the comparison of parameter correlations in the two sub-samples. This is discussed in the light of a self-shielding BLR model in Section 3.9. Conclusions are presented in Section 3.10.

Throughout this Chapter, we adopt a standard cosmology: $\Omega_{\mathrm{M}} = 0.3$, $\Omega_{\Lambda} = 0.7$ and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

3.2 Introduction: AGN unification and nuclear kinematics

In Chapter 1.1, we have seen how the exploration of the multi-wavelength emission of AGN, including their spectral features, motivated the development of unification schemes (e.g., Antonucci (1993); Urry & Padovani (1995); Netzer (2015); Padovani et al. (2017), but see also Elitzur (2012) for caveats of unification). Simplifying the complex historical AGN zoo is appealing since most of the identified species share common spectral properties over large ranges of luminosity and redshift. In this context, two main sub-classes have been defined based on their optical emission line properties: Type 1, which have both broad and narrow optical/UV emission lines, and Type 2, which show only narrow lines. These two classes are unified via the hypothesis that the broad lines in Type 2 sources are obscured (e.g., Netzer 2008; Hickox & Alexander 2018). There is nonetheless substantial variety in the properties within the Type 1 population. The broad lines encode information about the geometry and the kinematics of the inner regions of the system: the Broad Line Region (BLR, e.g., Rees et al. 1989; Peterson 2006; Gaskell 2009). As presented in Section 1.4, the width of the broad lines is used as virial broadening estimator to measure black hole masses. However, AGN broad lines show complex shapes, which may be affected by non-orbital contributions. It is, therefore, important to precisely understand the kinematics of the gas in the BLR to measure black hole masses reliably.

A potentially useful approach to constrain and explore the inner structure of AGN is to acquire an understanding of their spectral diversity. For instance, in their seminal paper Boroson & Green (1992) performed a PCA on a set of optical, radio and X-ray features of

87 Type 1 AGN. Through the orientation of the first principal component (the so-called Eigenvector1, or EV1) in their parameter space, they established an anti-correlation of the equivalent width of the narrow [OIII] lines at 4959Å and 5007Å with the relative strength of the iron emission $r_{\mathrm{FeII}} = F(\mathrm{FeII}\lambda4570)/F(\mathrm{H}\beta)$, and the FWHM of the H$\beta$ line as central markers of diversity in their sample. A large range of subsequent studies have investigated the EV1 correlation planes (e.g., Sulentic et al. 2000a,b; Marziani et al. 2001; Shang et al. 2003; Grupe 2004; Kuraszkiewicz et al. 2009; Mao et al. 2009; Tang et al. 2012; Shen & Ho 2014). In particular, Sulentic et al. (2000a,b) and Marziani et al. (2001) established the foundations of the four-dimensional EV1 (4DE1) formalism, including $\mathrm{FWHM}_{\mathrm{H}\beta}$ and r_{FeII} as two of the main correlates of EV1 in Boroson & Green (1992). These two quantities are respectively related to M_{BH} (since $\mathrm{FWHM}_{\mathrm{H}\beta}$ is used as virial broadening estimator) and the Eddington ratio $\lambda_{\mathrm{Edd}} = L/L_{\mathrm{Edd}}$ (e.g., Grupe et al. 1999; Marziani et al. 2001; Netzer & Trakhtenbrot 2007; Sun & Shen 2015; Du et al. 2016; Panda et al. 2019c). The 4DE1 parameter space was extended with the soft X-ray photon index Γ_{soft} and the centroid shifts from restframe wavelength of CIVλ1549 at 50% fractional) projected onto EV1. The points are colour-codedintensity (Sulentic et al. 2007a). The careful exploration of this parameter space led to a proposed two-population paradigm in the low-redshift Universe, with an empirical separation at $\mathrm{FWHM}_{\mathrm{H}\beta} \approx 4000\,\mathrm{km\,s^{-1}}$ (Sulentic et al. 2000a,b, see section 5 of Marziani et al. 2018 and references therein for a full review of the evidence supporting the population A/B classification). The optical plane of 4DE1 (i.e., $\mathrm{FWHM}_{\mathrm{H}\beta}$ vs r_{FeII}) was established as an analogue to the stellar Hertzsprung-Russel diagram for the Type 1 AGN population. The domain occupied by Type 1 AGN in this plane has been presented as a quasar main sequence (Marziani et al. 2001; Sulentic et al. 2011). The viewing angle and the accretion power have been investigated as potential drivers of this main sequence (Shen & Ho 2014; Panda et al. 2018). Z10 studied the characteristics of the Hβ emission of 477 optically selected AGN ($z < 0.7$) in the optical 4DE1 plane. They reported that the line shapes distinctively changed along the sequence, the asymmetry and shifts scaling with the optical dimensions, corroborating early results by Sulentic et al. (2002).

3.3 Data

The sample of Type 1 AGN analysed here is extracted from the SDSS-IV/SPIDERS AGN catalogue, presented in C19. The original catalogue compiles spectral information through SDSS DR14[2] (Abolfathi et al. 2018), for 7344 Type 1 AGN detected in the Second ROSAT All-Sky Survey ROSAT/2RXS, Boller et al. 2016) and 1157 Type 1 AGN in the XMMSL1 catalogue (Saxton et al. 2008), for which the multi-wavelength counterparts were determined in Salvato et al. (2018). Optical spectra of the counterparts were obtained from two different spectrographs, SDSS and BOSS. These instruments cover different optical wavelength ranges: $3800 - 9200\,$Å for SDSS (counting both channels, with a spectral resolution ranging from 1850 to 2200) and $3600 - 10400\,$Å for BOSS (spectral resolution of 1560-2270 in the blue channel, 1850-2650 in the red channel). A technical summary of the SDSS-IV

---

[2] SDSS DR14 VACs

survey is provided by Blanton et al. (2017). The Sloan Foundation 2.5-meter Telescope is presented in Gunn et al. (2006). Finally, a detailed description of the SDSS and BOSS spectrographs is given by Smee et al. (2013).

Details of the optical spectral fitting are presented in C19. The spectral regions centred around the Hβ and MgII lines ($4420 - 5500\,\text{Å}$ and $2450 - 3050\,\text{Å}$, respectively) were fitted with a multi-component continuum model and a series of Gaussian functions. The present Chapter will specifically focus on the broad Hβ emission line and the narrow [OIII]λ4959 and [OIII]λ5007 forbidden transition lines. C19's fitting algorithm used up to four Gaussian functions to fit the Hβ line. The best-fit model (and thus the required number of Gaussian components) was automatically selected based on the Bayesian Information Criterion (BIC). Of the four Gaussians, one accounts for the narrow core, while the remaining potential three components of Hβ are defined as broad if they fulfil the criterion: $\mathrm{FWHM}_{\mathrm{H}\beta} > 800\,\mathrm{km\,s^{-1}}$. This relatively complex model allows one to trace not only typical broad bimodal profiles above $\mathrm{FWHM}_{\mathrm{H}\beta} \sim 1000\,\mathrm{km\,s^{-1}}$, but also distinct broader components (see Section 8.1 in C19, and Section 3.7 of this Chapter). For instance, Marziani et al. (2010) propose a broad line decomposition into broad, very broad and blue components (respectively called 'BC','BLUE' and 'VBC', for details see section 6.2 of Marziani et al. 2018), which the model defined in C19 may individually trace.
Up to two Gaussians were used for each of the [OIII] lines: one fitting the narrow core and one tracing shifted wings. The continuum model was constructed from a power-law, a host galaxy component and an FeII template. The iron emission and the galaxy contribution are obtained respectively from a normalised Narrow-Line Seyfert 1 (NLS1) I Zw 1 template (Boroson & Green 1992) and an early type galaxy template[3]. Morphological studies (e.g., Grogin et al. 2005) have shown that AGN are mainly found in bulge dominated galaxies. Bulges are predominantly home to old star populations. C19 thus justify the use of an early-type galaxy template for the spectral host contribution by the fact that SDSS fibers sample only the central regions of targeted galaxies. The early-type approximation for the host contribution was also used in previous work (e.g. Calderone et al. 2017). The template has Hβ absorption features. Comparing a sub-sample with strong host contribution to a sub-sample with low host contribution at $z < 0.2$, we found no significant difference in their Hβ line shape diagnostics. The fitting parameters are listed in the catalogue of C19 along with monochromatic continuum and X-ray luminosities and derived parameters such as bolometric luminosities and estimates for M_{BH} and λ_{Edd}.

In C19, the single-epoch M_{BH} estimation method was used for the derivation of the black hole masses (Vestergaard 2002; McLure & Jarvis 2002; Vestergaard & Peterson 2006; Assef et al. 2011; Shen & Liu 2012 and Shen 2013 for a review). The masses stem from the Assef et al. (2011) calibration, which builds on the FWHM of Hβ and the BLR radius-luminosity relation from Bentz et al. (2009). Here we have reconstructed the best-fit models from the parameters listed in C19. One example of an Hβ fit reconstruction and details of the broad Hβ line decomposition are shown in Fig. 3.5.

---

[3] SDSS spectral cross-correlation templates, Template 24

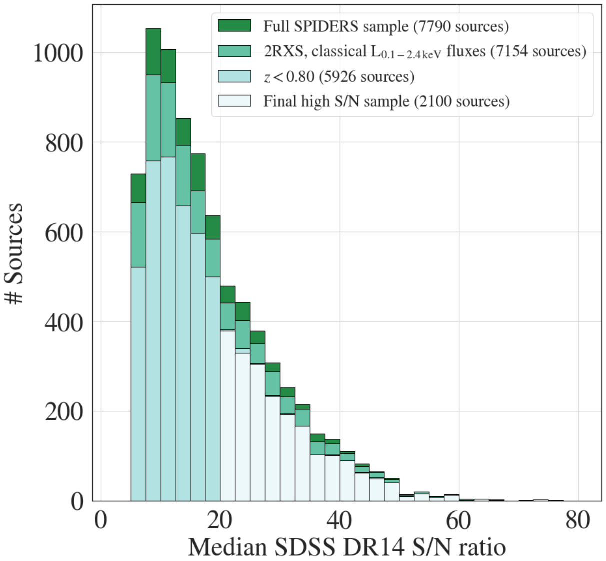


Figure 3.1: Distribution of the median S/N ratios per resolution element of the 2100 SDSS DR14 sources in our final sample(white). Also presented are sub-samples of the full SPIDERS sample in C19 after two consecutive cuts: requiring the presence of a $0.1-2.4$ keV flux and the redshift restriction $z < 0.8$.

3.3.1 Sample construction and selection pipeline

We are primarily interested in Type 1 AGN spectral properties, which can be associated with the physics of the BLR and the NLR. We thus assembled a parameter subset inspired by the classical quasar PCA papers by Boroson & Green (1992) and Grupe (2004). They included emission properties of Hβ, FeII, [OIII]λ5007[4] and HeIIλ1640, as well as X-ray and optical monochromatic luminosities to trace the emission of the hot electron corona and the accretion disk. We confined the analysis to sources detected in 2RXS only (7344 sources), as they were six times more numerous than XMMSL1 sources in the SPIDERS AGN sample. This simplifies the analysis, given the fact that the X-ray fluxes from the two surveys are measured in different bands and at different times (see Saxton et al. 2008; Boller et al. 2016). More specifically, we retained 2RXS sources, for which the X-ray flux in the $0.1-2.4$ keV band is provided in C19 (7154 sources). The missing 190 sources have no X-ray flux measurement listed. They correspond to sources with low photon count-rates.

Since we use the monochromatic luminosity at 5100 Å to trace the continuum emission, we limit the sample to $z < 0.8$ so that the restframe 5100 Å are covered by the spectral window. The redshift cut also ensures the presence of Hβ in the fits. These constraints yielded 5926 sources.

---

[4] We choose to work with the [OIII] transition line at 5007 Å and not 4959 Å, since this line will be less affected by interline fit contamination with Hβ. We argue that any blue asymmetries and shifts in [OIII]λ5007 should be detectable in [OIII]λ4959. Throughout this Chapter, [OIII]λ5007 will be referred to as [OIII]

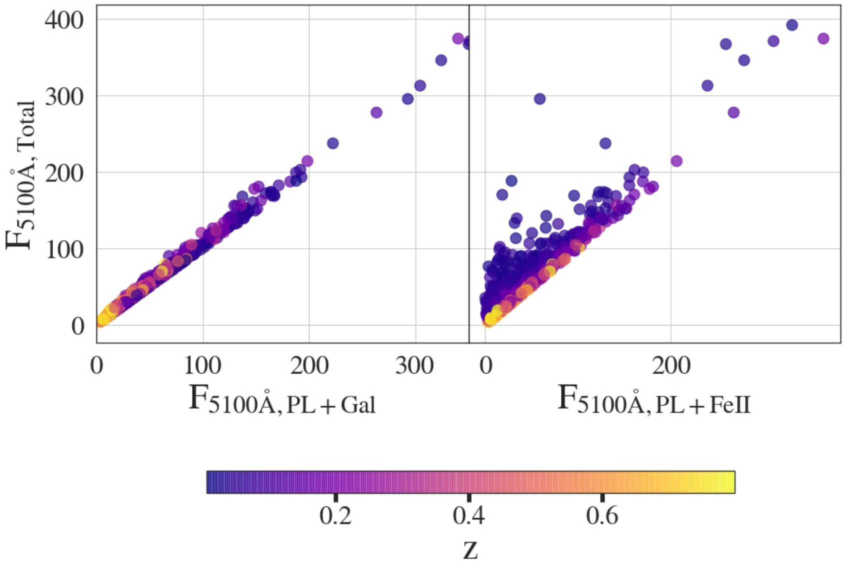


Figure 3.2: Component contribution to the monochromatic fluxes at 5100Å. The correlation between the monochromatic fluxes at 5100Å, as measured from the power-law model with FeII and host contributions ($F_{5100\mathrm{\mathring{A}},\mathrm{Total}}$), the power-law model with the FeII emission ($F_{5100\mathrm{\mathring{A}},\mathrm{PL+FeII}}$) and the power-law model with the host emission ($F_{5100\mathrm{\mathring{A}},\mathrm{PL+Gal}}$) is shown. The fluxes are in units of $10^{-17}\,\mathrm{erg\,cm^{-2}\,s^{-1}\,\mathring{A}^{-1}}$.

We further limit our analysis to sources with good SDSS S/N per resolution element in order to improve the significance of our results. Our analysis relies on the accurate measurement of line shapes, and low-quality spectra might adversely affect the fits. For this reason, we consider only sources with median S/N ≥ 20. This critical restriction drops our total count to 2124 objects. The SDSS S/N distribution of the original sample and its consecutive cuts are shown in Fig. 3.1. Finally, 24 sources were excluded due to the lack of M_{BH} estimate, large uncertainties on the fit parameters of Hβ or absent [OIII]. A visual inspection of these sources' spectra revealed that these issues might be linked to strong continuum emission.

For the remaining 2100 sources, the following subset of spectral properties was compiled (unless specified otherwise, all wavelengths, energies and equivalent widths are defined in the restframe):

- $\mathbf{FWHM_{H\beta}}$: full width at half maximum of Hβ (broad component) [$\mathrm{km\,s^{-1}}$]
- $\boldsymbol{F}$**([OIII])/**$\boldsymbol{F}$**(H**$\boldsymbol{\beta}$**)** : flux ratio of [OIII]λ5007Å and Hβ
- $\boldsymbol{r}_{\mathbf{FeII}} = \boldsymbol{F}$**(FeII)/**$\boldsymbol{F}$**(H**$\boldsymbol{\beta}$**)**: flux ratio of H$\beta$ and the total FeII (FeII emission in the $4434 - 4684\,\mathrm{\mathring{A}}$ range)

- $\boldsymbol{L_{5100\text{Å}}}$: optical continuum monochromatic luminosity at 5100$\mathring{A}$ from the power-law component (without host and FeII contributions) [$\text{erg}\,\text{s}^{-1}\text{Å}^{-1}$]
- $\boldsymbol{L_{\mathbf{X}}}$: observed X-ray luminosity in the 0.1-2.4 keV band [$\text{erg}\,\text{s}^{-1}$]
- $\boldsymbol{W}$**(FeII)** : equivalent width of FeII (4570 Å blend)[5] [Å]
- $\boldsymbol{W}$**([OIII])** : equivalent width of [OIII]λ5007 (full profile)[5][Å]
- $\boldsymbol{W(\mathrm{H}\beta)}$: equivalent width of the broad component of Hβ[5][Å]
- $\boldsymbol{W}$**(HeII)** : equivalent width of HeII at 4686 Å[5][Å]
- $\boldsymbol{\Delta\lambda_{[\mathrm{O\,III}]}}$: asymmetry index for [OIII]λ5007 [6]
- $\boldsymbol{\Delta\lambda_{\mathrm{H}\beta}}$: asymmetry index for the broad component of Hβ[6]

We further note the following:

i Sources for which the parameters relative to the FeII emission were not listed in C19, typically correspond to AGN with very weak FeII emission. We manually set their r_{FeII} and iron equivalent widths to zero. We thus implicitly assume that their iron emission is too weak to be disentangled from the AGN continuum emission.

ii All the spectral properties of our defined sub-set are available in C19 except for the equivalent widths and asymmetry indices which were derived as presented in Secs. 3.3.2 and 3.3.3.

iii The $L_{5100\,\text{Å}}$ provided in C19 includes the host contribution. However, in the present analysis, we aim to trace the accretion disk emission with $L_{5100\,\text{Å}}$. We thus derived the monochromatic luminosity from the reconstructed power-law model at restframe, removing the FeII and host contribution. The uncertainties are obtained from the errors in the normalisation of the power-law. The considerable degeneracy between the host, iron and power-law components at lower redshift ($z < 0.2$) due to the stronger host contribution makes it challenging to correctly perform an AGN-host decomposition. It is the principal limitation of deriving monochromatic luminosities directly from the fitted power-law model. The relative contributions of the FeII complex and the host galaxy to the monochromatic flux a 5100Å are illustrated in Fig. 3.2. While the FeII contribution at 5100Å to the total monochromatic flux is marginal over the complete redshift range, the host galaxy strongly contributes to the monochromatic flux at lower redshifts.

---

[5] cf. 3.3.3

[6] cf. Section 3.3.2

Sample cuts (Initial SPIDERS AGN sample: 7790 sources)	remaining sources
2RXS flux	7154
$z < 0.8$	5926
$S/N \geq 20$	2124
Robust asymmetry index/black hole masses	2100

Table 3.1: Summary of the sample construction pipeline

iv The X-ray fluxes were obtained by C19, following Dwelly et al. (2017), assuming an absorbed power-law ($\Gamma = 2.4$). The X-ray luminosities are derived from the fluxes assuming the same spectral model. In addition to these classical fluxes, C19 provide Eddington bias-corrected 2RXS fluxes using a Bayesian prior (Kraft et al. 1991; Laird et al. 2009; Georgakakis & Nandra 2011). The Bayesian flux estimates reach low values for sources with low count rates. C19 allow up to a factor of ten difference with respect to the classical fluxes. Sources for which the Bayesian fluxes do not meet this requirement have only classical flux measurements registered. We chose to use the uncorrected X-ray flux measurements in this Chapter since using Bayesian fluxes would reduce the size of our final sample by a factor of two. We note that a test run of the statistical pipeline presented in this Chapter on a smaller sample with Bayesian soft X-ray luminosities yielded similar results to those presented here for the entire sample using the nominal fluxes.

Fig. 3.3 presents the soft X-ray luminosity-redshift distribution of our sources and compares our sample to that of Grupe (2004). Our sample spans a range of soft X-ray luminosities from $1.9 \times 10^{41}\,\mathrm{erg\,s^{-1}}$ to $9.9 \times 10^{45}\,\mathrm{erg\,s^{-1}}$, with redshifts up to 0.8. Compared to Grupe (2004), we sample to lower luminosities and higher redshifts. The larger sample should allow the placement of more stringent constraints on the relation of EV1 and EV2 to physical parameters (see Fig. 3.11).

3.3.2 Measuring asymmetry in emission lines: motivation and method

Emission lines in AGN show asymmetries which encode precious information on the geometry and kinematics of the emitting region. In the case of broad emission lines, asymmetry can be defined as the relative displacement of the centroid wavelength of the peak and base components of the profile. The peak wavelength of the profile is, in practice, measured at a high fraction of the broad component intensity (without the narrow core) and is related to the classical broad component (Brotherton et al. 1994; Popovic et al. 2002; Adhikari et al. 2016). The base of the emission line profile is, as its name indicates, measured at lower fractional intensities and is expected to trace a distinct, red- or blueshifted, sometimes broader component. The shifted VBC is believed to originate from the presence of a distinct emitting region in the BLR: the Very Broad Line Region (VBLR). The VBLR is

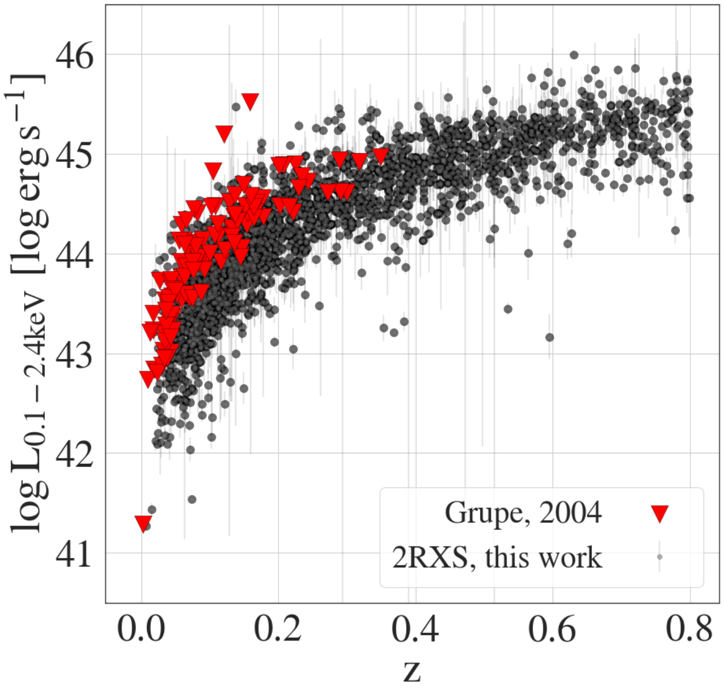


Figure 3.3: The observed soft X-ray luminosity-redshift distribution for our 2100 2RXS sources. The red triangles are the 110 AGN used in Grupe (2004), which all lie at the brighter end of our sample for equivalent redshifts. Our sample significantly extends both the redshift and the luminosity ranges.

expected to be composed of highly ionized gas situated even closer to the central black hole. It is further discussed in Section 3.7. The concept of BLR stratification arose from reverberation studies. Peterson & Ferland (1986) reported variability in the profiles of HeII4686 and Hβ in NGC 5548, most notably the appearance of broader line components while the continuum luminosity increased. Measuring the asymmetry of broad emission lines and including the indices in a statistical analysis, therefore, offers an approach to investigate how the relative kinematics of assumed layers in the BLR impacts the diversity of observed Type 1 AGN. One can also use the asymmetry index to trace the relative displacement of narrow emission lines and their shifted wings, to trace the presence of outflows in the NLR (e.g., Heckman et al. 1981; Zamanov et al. 2002; Zakamska et al. 2016; Wang et al. 2018; Rakshit & Woo 2018). In this case, the displacement of peak and base components is measured for the full profile (i.e. narrow core and shifted wings).

The asymmetry index was not included in C19 and was determined here starting from the reconstruction of the lines, using the Gaussian models listed in the catalogue. In this Chapter, we derived the following asymmetry indices:

- $\Delta\lambda_{\mathrm{H}\beta}$ for the broad component of Hβ. From the C19 catalogue we retain the Gaussians with FWHM $> 800\,\mathrm{km\,s^{-1}}$ as broad components. From the four Gaussian components used to fit the full Hβ profile, one was constrained by C19 to have FWHM $< 800\,\mathrm{km\,s^{-1}}$ to force the fit of the narrow emission line. Up to three Gaussians were thus used to fit the broad profile.

- $\Delta\lambda_{[\mathrm{O\,III}]}$ for the full profile of [OIII]λ5007. Here both narrow and broad components are used for the measure.

In this Chapter, we use the description of the asymmetry $\Delta\lambda$ measurement for any emission line first introduced by Heckman et al. (1981), and further detailed in Winkler & Chauke (2014). The method requires the identification of the wavelengths at which the line reaches 15% and 80% of its maximum intensity, which in turn requires the reconstruction of the multi-Gaussian line fits as described in Section 3.3.3. We determined the maximum of the full line profile (or alternatively the broad components for Hβ) and measured 15% and 80% of this maximum. A schematic view of this measurement is presented in Fig. 3.4. The asymmetry index ($\Delta\lambda$) is defined by:

$$\Delta\lambda = \frac{\alpha_\mathrm{b} - \alpha_\mathrm{r}}{\alpha_\mathrm{b} + \alpha_\mathrm{r}} = \frac{(\lambda_\mathrm{c,high} - \lambda_\mathrm{b,low}) - (\lambda_\mathrm{r,low} - \lambda_\mathrm{c,high})}{(\lambda_\mathrm{c,high} - \lambda_\mathrm{b,low}) + (\lambda_\mathrm{r,low} - \lambda_\mathrm{c,high})} \tag{3.1}$$

$$= \frac{\lambda_\mathrm{b,high} + \lambda_\mathrm{r,high} - \lambda_\mathrm{b,low} - \lambda_\mathrm{r,low}}{\lambda_\mathrm{r,low} - \lambda_\mathrm{b,low}} \tag{3.2}$$

Here $\lambda_\mathrm{b,high}$ and $\lambda_\mathrm{r,high}$ are measured at 80% fractional intensity and $\lambda_\mathrm{b,low}$ and $\lambda_\mathrm{r,low}$ at 15% fractional intensity. $\lambda_\mathrm{c,high}$ is the wavelength of the centre of the line at 80% fractional intensity, i.e., the midpoint between $\lambda_\mathrm{b,high}$ and $\lambda_\mathrm{r,high}$. The parameters α_b and α_r are the distances from $\lambda_\mathrm{c,high}$ to the blue and red edge of the line at 15% fractional intensity. A

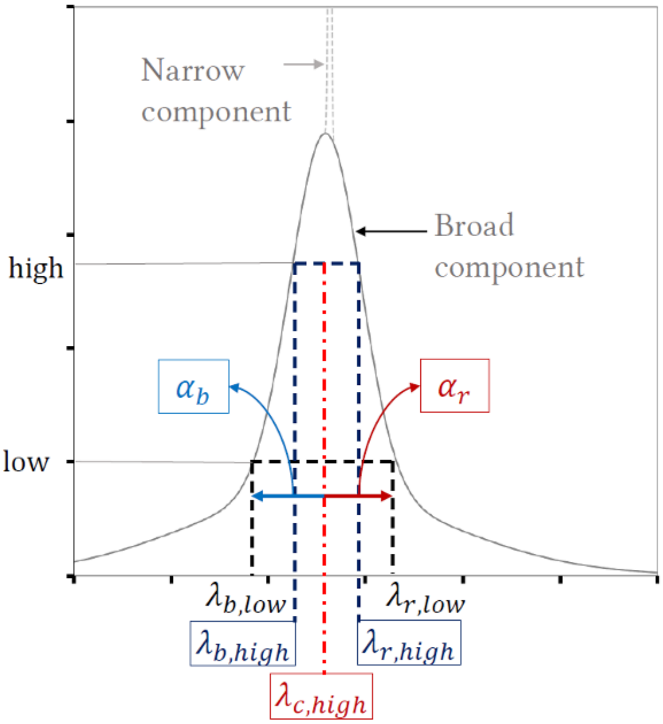


Figure 3.4: Diagram illustrating the determination of the asymmetry parameter using the Heckman et al. (1981) definition, adapted from Winkler & Chauke (2014). For Hβ the broad component is used without the narrow line. For [OIII], the asymmetry of the full profile (narrow and wing) is measured.

negative (positive) $\Delta\lambda$ corresponds to a blue-ward (red-ward) asymmetry, and its value ranges from -1 to 1. This measure is not sensitive to the continuum emission (cf. correlation matrix shown in Section 3.5.1). We propagated the uncertainties of the FWHM of the fits into the measure of asymmetry and obtained typical errors of $\sigma_{\Delta\lambda} \sim 0.05$. An important caveat is that using the uncertainties of the FWHM from the full profile does not account for the lower, distinct and possibly more continuum- and FeII-contaminated line components.

3.3.3 Equivalent width estimates

The equivalent width offers a measure of an emission line's strength relative to the continuum. Equivalent widths were not included in C19 and had to be derived.

The equivalent width of a spectral line is defined as:

$$W_\lambda = \int \frac{S_1(\lambda) - S_c(\lambda)}{S_c(\lambda)} d\lambda \tag{3.3}$$

S_1 denotes the total flux density at a given wavelength, and S_c the continuum flux density.

To obtain the interpolated S_1 and S_c fluxes, the lines were reconstructed from the catalogue's fitting parameters, the associated Gaussian fitting functions, the continuum power law, and the templates for the iron and galaxy contributions. Each Gaussian component $G(\lambda)$ was reconstructed according to:

$$G_{\text{line}}(\lambda) = N_{\text{line}} \cdot \exp\left[-\frac{1}{2}\left(\frac{\lambda - P_{\text{line}}}{W_{\text{line}}}\right)^2\right] \tag{3.4}$$

Here N_{line}, P_{line} and W_{line} denote the normalisation, peak wavelength and width for the fit, as listed in C19. For multiple Gaussian components $G_j(\lambda)$ the total flux in the line is then obtained from $G_{total}(\lambda) = \sum_j G_j(\lambda)$. In C19, the continuum model consists of a power law, a host galaxy component and the iron emission.

The power law model is reconstructed from :

$$PL(\lambda) = N_{\text{PL}} \cdot \lambda^{S_{\text{PL}}} \tag{3.5}$$

N_{PL} and S_{PL} are the normalisation and the slope of the power law for each spectrum, as listed in the catalogue. Following the catalogue's author's prescription, a Gaussian filter was applied to the FeII emission to reproduce the blending of the multiplets. The continuum emission is calculated by:

$$S_c(\lambda) = PL(\lambda) + Fe\text{II}(\lambda) + GAL(\lambda) \tag{3.6}$$

The total flux in a given line is: $S_1(\lambda) = S_c(\lambda) + G_{\text{total}}(\lambda)$.

Fig. 3.5 displays a fit centred around the Hβ/[OIII] complex, reconstructed using the above model. Interpolating D_1 and S_c over the fitting range of Hβ or MgII and performing the integration in Eq. 3.3 yields the equivalent width for the chosen lines.

3.4 Impact of fit contamination on the Hβ and [OIII] asymmetry index distribution

The distributions for Hβ and [OIII] asymmetry indices are shown in Fig. 3.6. The indices clearly follow very different distributions for the two emission lines. The following statistics were derived for the distributions of asymmetries with base centroids measured at 25% fractional intensity (i.e., the blue histograms in Fig. 3.6). The asymmetry indices of the [OIII] are predominantly blue-ward (skewness of distribution: $\gamma = -0.54$), which is consistent with the presence of a blueshifted wing detected in a two-Gaussian model. The mean value of $\Delta\lambda_{[\text{OIII}]}$ is ~ -0.10. 1546 sources with [OIII] asymmetry below zero and 547 above zero are counted, confirming the excess of blue asymmetric [OIII] lines. The distribution of the Hβ asymmetry appears positively asymmetric (skewness of distribution $\gamma = 0.36$). The mean value of $\Delta\lambda_{\text{H}\beta}$ is +0.03. An excess of red-asymmetric Hβ emitters is clearly measured (1244 sources out of 2100 with $\Delta\lambda_{\text{H}\beta} > 0$). The fractional intensity at which the lower velocity shift is measured for the asymmetry index has important repercussions on the overall distribution of asymmetries, as Fig. 3.6 demonstrate. The distribution of $\Delta\lambda_{\text{H}\beta}$

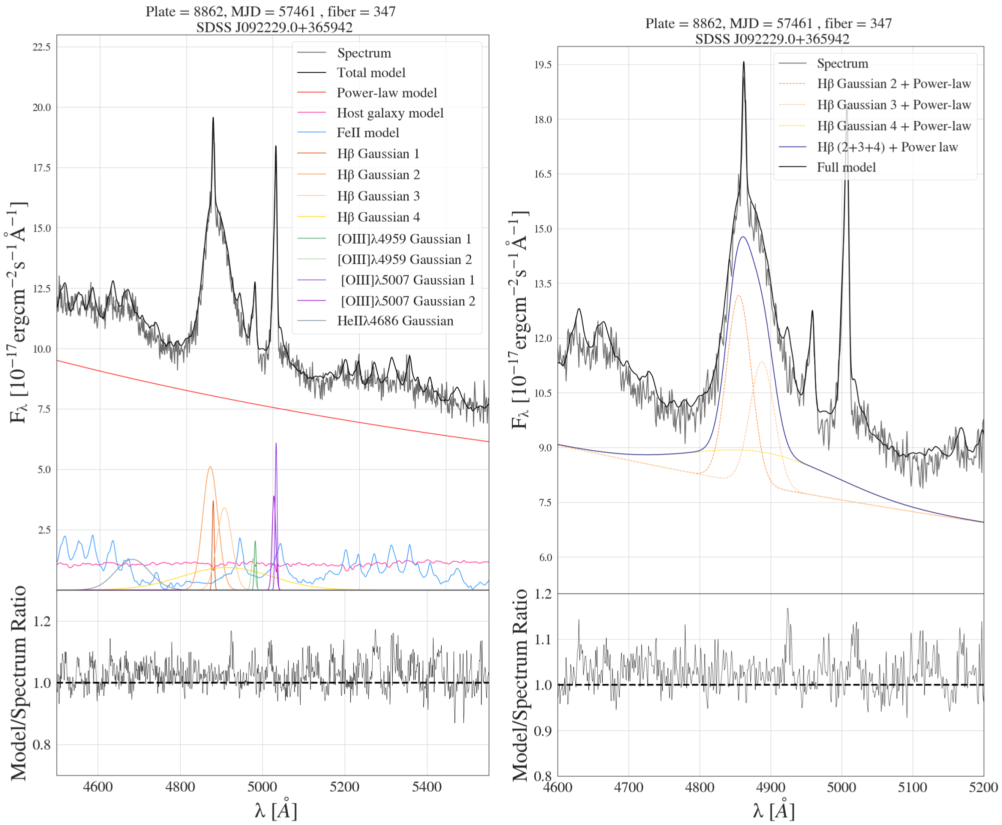


Figure 3.5: *Left panel*: Reconstructed Hβ fitting region using the model used by C19. The [OIII], HeII and Hβ model components are labeled. The components of the continuum model (a power-law, a host galaxy model and the iron pseudo-continuum) are also displayed. The originally observed spectrum is shown in grey. *Right panel:* A zoom onto the Hβ Gaussian line decomposition. In order to improve the visualisation, the power-law contribution was added to the Gaussian components. Only two broad components were required to fit Hβ.

and $\Delta\lambda_{[\mathrm{O\,III}]}$ are displayed for different fractional intensities at which $\lambda_{\mathrm{b,low}}$ and $\lambda_{\mathrm{r,low}}$ are set : 5%, 10%, 15%, 20%, 25%. While we always detect an excess in red-ward asymmetry for the Hβ profiles, the lower the percentage at which $\lambda_{\mathrm{b,low}}$ and $\lambda_{\mathrm{r,low}}$ are measured, the more the asymmetry index distribution is bimodal, with a clear separation between the symmetric and the red asymmetric Hβ modes. It is unclear if this red-ward excess in Hβ asymmetries is influenced by [OIII] and continuum contamination in the fit. The bimodality of the $\Delta\lambda_{\mathrm{H}\beta}$ distribution measured at lower base intensities, however, clearly indicates a preferred displacement between peak and base component, which could arise from the near systematic contamination by [OIII] or its blue wing.

A similar effect is observed for the distribution of the [OIII] asymmetries: it becomes double-peaked, with a distinct second mode appearing blue-ward of the symmetric sources for lower base intensities. For the measurement of the asymmetry index, we set the base of the lines at 15% fractional intensity in order to avoid the detection of overlaying line models. This choice is motivated by two aspects of the distributions seen in the two panels of Fig. 3.6: 15% fractional intensity, the peak of the symmetric sources is more sensitive to low intensity, shifted profile components are detected by the measurement than at 25% and 20%. Furthermore, for this measurement configuration (peak at 80% and base at 15% fractional intensity), the bimodality, which we identify as a signature of model degeneracy, is not yet detected.

To further investigate the risk of model degeneracy, the dependency of the [OIII] asymmetry index on the shifts of the Hβ base component c_{15} (the centroid shift at 15% fractional intensity with respect to the restframe wavelength) is presented in Fig. 3.7. We clearly observe an excess of sources in the quadrant of blue-asymmetric [OIII] lines and redshifted Hβ base components. The coloured markers correspond to high S/N ratio sources, which also appear to preferentially populate the lower right quadrant in the figure. 1196 sources out of 2100 (57%) have $\Delta\lambda_{[\mathrm{O\,III}]} < 0$ and $c_{15} > 0$. In the high S/N regime, i.e. S/N > 35, 278 out of 498 sources (56%) have $\Delta\lambda_{[\mathrm{O\,III}]} < 0$ and $c_{15} > 0$. A binomial test showed that the clustering in the lower right quadrant is indeed significant. For both c_{15} and $\Delta\lambda_{[\mathrm{O\,III}]}$, the null hypothesis that the sources distribute equally on both sides of zero was rejected with p-values $p < 10^{-6}$.[7] This result underlines the possibility of model degeneracy in the Hβ region. For example, in their study of Extremely Reddened Quasars (ERQs), Perrotta et al. (2019) presented rare sources for which Hβ and [OIII] appear to blend (see their Fig. 1). Disentangling the presence of a shifted VBC from fit contamination would offer a cleaner window on the kinematics of the BLR, however, the fitting procedure in C19 did not allow for such a decomposition.

3.5 Statistical Analysis

This section presents the various statistical tests performed on the selected optical properties of our sample of 2100 SPIDERS AGN, listed in Section 3.3.

---

[7] Throughout this Chapter, ρ denotes correlation coefficients, while p stands for p-value

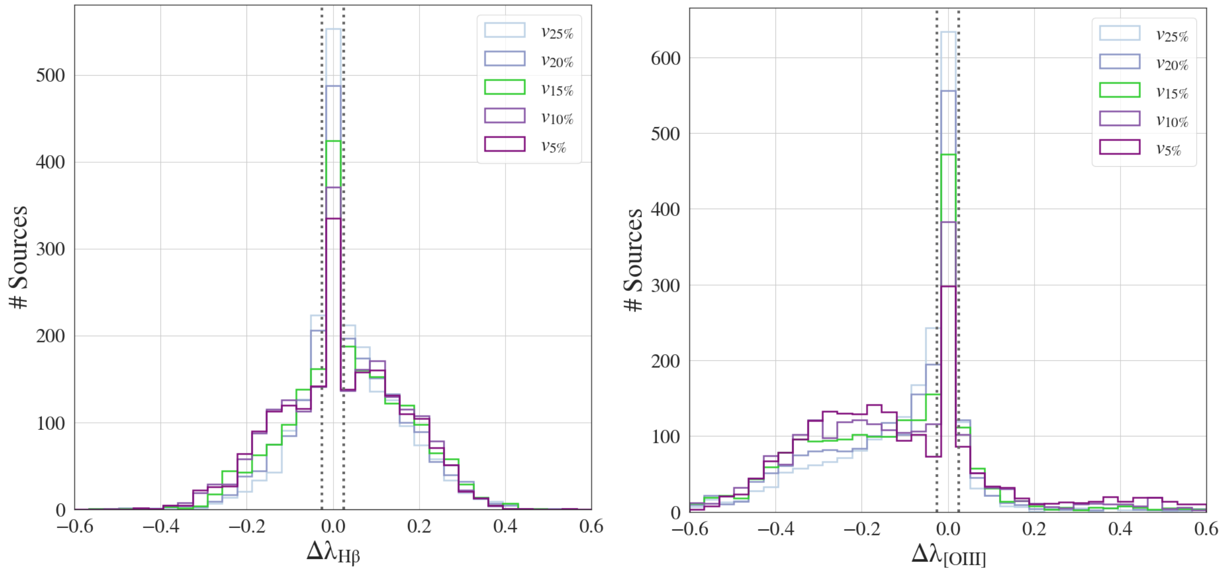


Figure 3.6: *Left*: The Hβ asymmetry distribution in the final source sub-sample for different levels of base fractional intensity: 5%, 10%, 15%, 20% and 25%. A clear inflexion redward of the symmetric Hβ profiles arises for bases measured at lower intensities. Typical 1σ errors are shown centred at zero by two dotted vertical bars. $\Delta\lambda_{\mathrm{H}\beta}$ values derived for a fractional base intensity of **15%** are used throughout this Chapter (green histogram). *Right*: The [OIII] asymmetry distribution for different levels of base fractional intensity: 5%, 10%, 15%, 20% and 25%. A clear inflexion can be detected here, bluewards of the symmetric [OIII] profiles for bases measured at lower intensities. Typical 1σ errors are shown centred at zero by two dotted vertical bars. The green histogram was obtained from $\Delta\lambda_{[\mathrm{OIII}]}$ measurements with base fractional intensity at **15%** as used in this Chapter.

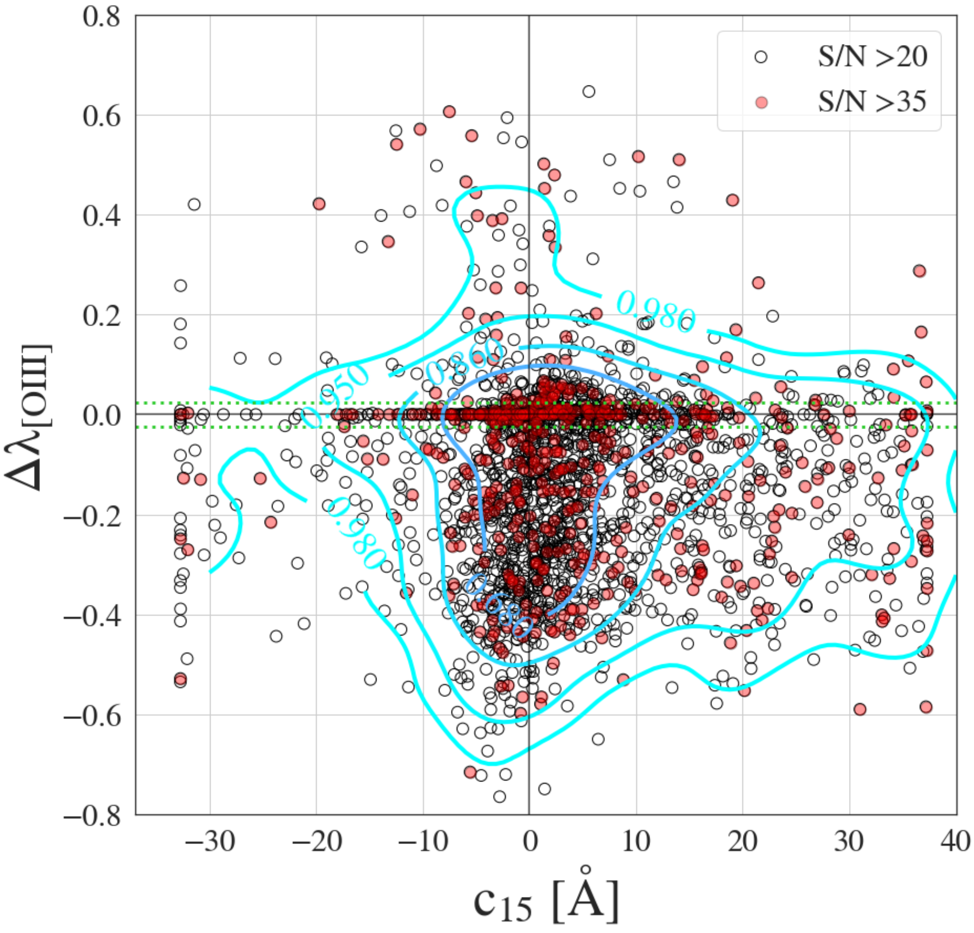


Figure 3.7: The asymmetry index of [OIII] as a function of the Hβ centroid shift at **15**% fractional intensity. The excess of sources in the lower right quadrant suggests inter-line contamination for redshifted Hβ components and blueshifted [OIII] components. The filled circles mark the sources with $S/N > 35$. The green dotted horizontal lines indicate typical 1σ uncertainties in the asymmetry parameter. The density contours containing 68%, 86%, 95% and 98% of all the sources in our sample were obtained from a Gaussian kernel density estimate.

3.5.1 Direct correlation

Our initial step is to generate a Spearman rank correlation matrix for our sample (Fig. 3.8). Statistically insignificant correlations, defined by imposing a threshold of $\mathrm{p_{max}} = 0.05/55$ on the p-values, were masked. The threshold $\mathrm{p_{max}}$ applies the Bonferroni correction for multiple statistical hypothesis testing (e.g., Haynes 2013). Positive (negative) coefficients indicate a positive (negative) monotonic relation between two parameters. We have included a dendrogram (tree diagram) which clusters our data hierarchically using correlation as distance metric (for a review, see Baron 2019). The clustered correlation matrix reveals significant ($\mathrm{p} < 10^{-6}$) structures in our parameter subset, some of which were also previously reported in Boroson & Green (1992), Grupe (2004) and Shen & Ho (2014). In particular:

- i A clear anti-correlation ($\rho = -0.53$) arises between $\mathrm{FWHM_{H\beta}}$ and r_{FeII}. This behaviour is the well-known anti-correlation between the first two dimensions of the 4DE1 (e.g., Sulentic et al. 2000a,b; Marziani et al. 2003a,b; Sulentic et al. 2007b, for a review, see Marziani et al. 2018 and Sulentic & Marziani 2015). In addition, W(FeII) and $\mathrm{FWHM_{H\beta}}$ are anti-correlated ($\rho = -0.38$), as expected.
- ii Similarly, an even stronger anti-correlation ($\rho = -0.61$) is measured between W([OIII]) and r_{FeII}, in agreement with the main EV1 anti-correlation.
- iii The X-ray and continuum luminosities, L_{X} and $L_{5100\text{Å}}$, are related to the emission line properties in a similar manner. There is a strong correlation ($\rho = 0.54$ for $L_{5100\text{Å}}$ and $\rho = 0.55$ for L_{X}) with $W(\mathrm{H}\beta)$. The correlation between $L_{5100\text{Å}}$ and $W(\mathrm{H}\beta)$ is consistent with photoionization models in which increasing emission of the central engine results in a more luminous BLR. L_{X} and $L_{5100\text{Å}}$ decrease strongly ($\rho = -0.44$ and $\rho = -0.51$ respectively) with increasing flux ratio F([OIII])/F(Hβ).
- iv The asymmetry index $\Delta\lambda_{\mathrm{H\beta}}$ is positively correlated with W([OIII]) ($\rho = 0.34$) and the $\mathrm{FWHM_{H\beta}}$ ($\rho = 0.35$). We also report its relatively strong anti-correlation with r_{FeII} ($\rho = -0.51$) and W(FeII) ($\rho = -0.40$). Despite using a different convention for the asymmetry index, Boroson & Green (1992) describe a similar correlation behaviour of $\Delta\lambda_{\mathrm{H\beta}}$.

The key result here is not the rediscovery of correlations mapped out by, e.g., Boroson & Green (1992); Grupe (2004); Shen & Ho (2014), but the fact that these relations hold for type 1 AGN in general, up to a redshift limit of at least $\mathrm{z} = 0.80$ and for luminosities up to $L_{\mathrm{X}} \sim 10^{46} erg.s^{-1}$. In addition, a new insight provided by the correlation matrix is the correlation behaviour of $\Delta\lambda_{\mathrm{[OIII]}}$ with parameters related to the BLR emission. Namely, the asymmetry index of [OIII] appears to be marginally related to the rest of the parameters, suggesting the absence of kinematic linkage between the inner region of the AGN and the NLR. This result is in contrast to Zamanov et al. (2002), who reported evidence for a correlation between the shifts of the high ionization lines CIVλ1549 and the shifts of the [OIII] lines, inferring a possible linkage between the NLR and BLR.

3.5.2 Partial correlation

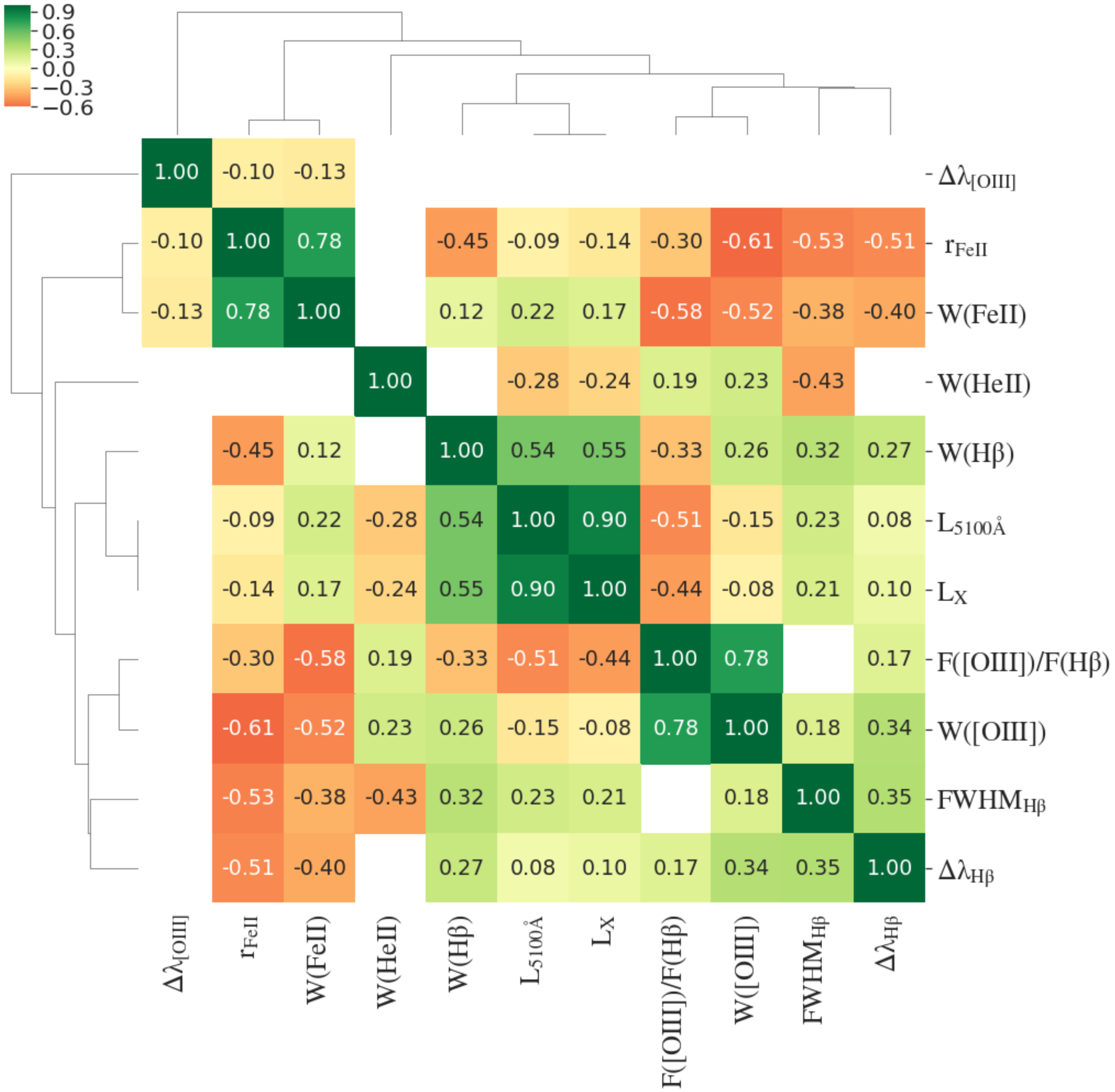
In order to study the correlations with the $\Delta\lambda_{\mathrm{H}\beta}$ parameter more closely we perform a partial correlation analysis of $\mathrm{FWHM}_{\mathrm{H}\beta}$, $W([\mathrm{O\,III}])$, $r_{\mathrm{Fe\,II}}$ and $\Delta\lambda_{\mathrm{H}\beta}$ (e.g., Baba et al. 2004). For this exercise, we first generate the Pearson correlation coefficients for this subset of parameters, then measure the strength of each of these correlations while controlling for the other confounding variables. The results are shown in Fig. 3.9. When we control for $\mathrm{FWHM}_{\mathrm{H}\beta}$ and $r_{\mathrm{Fe\,II}}$, the correlation between $W([\mathrm{O\,III}])$ and $\Delta\lambda_{\mathrm{H}\beta}$ significantly decreases (we measure a drop $\rho_P = 0.20$, $\mathrm{p} < 10^{-7}$ to $\rho_{P,partial} = 0.017$, $\mathrm{p} = 0.45$). The p-value of the latter partial correlation is above the corrected significance threshold: $\mathrm{p} = 0.05/6$. Similarly, when we control for $W([\mathrm{O\,III}])$ and $r_{\mathrm{Fe\,II}}$, the strength of the $\mathrm{FWHM}_{\mathrm{H}\beta}$ vs $\Delta\lambda_{\mathrm{H}\beta}$ correlation decreases ($\rho_P = 0.28$, $\mathrm{p} < 10^{-6}$ to $\rho_{P,partial} = 0.075$, $\mathrm{p} \sim 10^{-4}$). The anti-correlation between $\Delta\lambda_{\mathrm{H}\beta}$ and $r_{\mathrm{Fe\,II}}$ is, however, less affected when we control for $\mathrm{FWHM}_{\mathrm{H}\beta}$ and $W([\mathrm{O\,III}])$ ($\rho_{P,partial} = -0.44$, $\mathrm{p} < 10^{-6}$ instead of $\rho_P = -0.52$, $\mathrm{p} < 10^{-6}$). These results demonstrate that $\Delta\lambda_{\mathrm{H}\beta}$ parameter is principally related to the relative strength of the iron emission and is marginally linearly correlated with $\mathrm{FWHM}_{\mathrm{H}\beta}$, indicating that the broader profiles tend to be more red-skewed, as also suggested by Marziani et al. (2013a).

3.5.3 Principal Component Analysis

The second step consisted of running a PCA on the standardized data set to determine which parameters contribute most to the total variance of our sample. It is particularly important to scale the parameters to unit variance since PCA is sensitive to the variance of the parameter distributions. PCA is an orthogonal linear transformation which is often used as a dimensionality-reduction algorithm (for a review, see Jolliffe & Cadima 2016). It yields the eigenvectors in parameter space which point in the direction of total maximal variance in the dataset. For an $\mathrm{N} \times \mathrm{M}$ dataset, the first component is found by minimizing the distance between points in the M-dimensional parameter space and their orthogonal projections onto an M-dimensional vector, which simultaneously increases the variance of the projected points. The second component is chosen to be orthogonal to the first and is determined in the same manner given this condition. The associated eigenvalues measure the amount of explained variance in each principal component (eigenvector). The variables (our parameters) are linked to the principal components by linear coefficients. These coefficients yield the amount of a variable's variance explained by the component.[8] While PCA is a very common tool in astronomy, it comes with several disadvantages, such as the linearity of the dimensionality reduction, which might not capture the full complexity of the data. One must, in general, be careful with the choice of the particular dimensionality reduction one uses and its interpretation (e.g., see Baron 2019). The first two principal components are presented in Fig. 3.10. The bar diagrams show the values of the correlation coefficients which link the parameters (our variables) to the given component. The first and

---

[8] In practice, we made use of a SCIKIT-LEARN implementation of PCA (Pedregosa et al. 2011).



	$\Delta\lambda_{[OIII]}$	r_{FeII}	W(FeII)	W(HeII)	W(Hβ)	$L_{5100Å}$	L_X	F([OIII])/F(Hβ)	W([OIII])	$FWHM_{H\beta}$	$\Delta\lambda_{H\beta}$
$\Delta\lambda_{[OIII]}$	1.00	-0.10	-0.13								
r_{FeII}	-0.10	1.00	0.78		-0.45	-0.09	-0.14	-0.30	-0.61	-0.53	-0.51
W(FeII)	-0.13	0.78	1.00		0.12	0.22	0.17	-0.58	-0.52	-0.38	-0.40
W(HeII)				1.00		-0.28	-0.24	0.19	0.23	-0.43	
W(Hβ)		-0.45	0.12		1.00	0.54	0.55	-0.33	0.26	0.32	0.27
$L_{5100Å}$		-0.09	0.22	-0.28	0.54	1.00	0.90	-0.51	-0.15	0.23	0.08
L_X		-0.14	0.17	-0.24	0.55	0.90	1.00	-0.44	-0.08	0.21	0.10
F([OIII])/F(Hβ)		-0.30	-0.58	0.19	-0.33	-0.51	-0.44	1.00	0.78		0.17
W([OIII])		-0.61	-0.52	0.23	0.26	-0.15	-0.08	0.78	1.00	0.18	0.34
$FWHM_{H\beta}$		-0.53	-0.38	-0.43	0.32	0.23	0.21		0.18	1.00	0.35
$\Delta\lambda_{H\beta}$		-0.51	-0.40		0.27	0.08	0.10	0.17	0.34	0.35	1.00

Figure 3.8: Spearman rank correlation matrix for our parameters. A positive (negative) coefficient indicates a positive (negative) monotonic correlation. Statistically insignificant correlations (with p-values $> 0.05/55$) have been masked. The matrix was hierarchically clustered using correlation as a distance measure. A dendrogram displays the average linkage of the parameters. Previously reported correlations are confirmed for a larger range of redshift and luminosities.

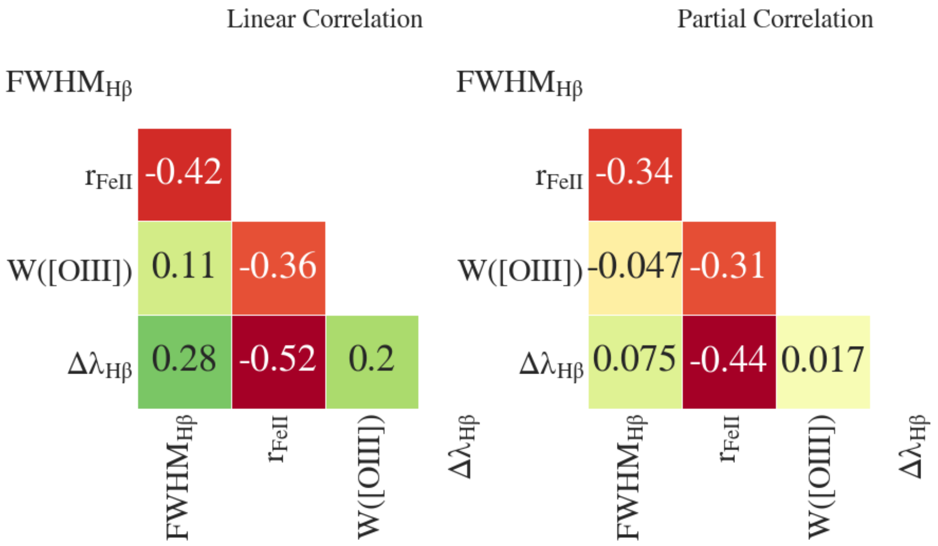


Figure 3.9: Left panel: Pearson Correlation coefficients for the parameter subset: $\mathrm{FWHM}_{\mathrm{H}\beta}$, $W([\mathrm{O\,III}])$, $r_{\mathrm{Fe\,II}}$, $\Delta\lambda_{\mathrm{H}\beta}$. Right panel: partial correlation coefficients (correlation coefficients when the effects of confounding variables are removed).

second principal components, respectively, explain 25.2% and 20.3% of the total variance in the selected parameter subset.

The first principal component (Eigenvector 1, EV1) is anti-correlated with the strength of the iron emission and correlated with $\mathrm{FWHM}_{\mathrm{H}\beta}$ and the (relative) strength of the [O III] emission, i.e., the diversity in the selected optical features of our SPIDERS AGN subsample is dominated by the anti-correlation of $W([\mathrm{O\,III}])$ and $W(\mathrm{Fe\,II})$. The Balmer profile asymmetries $\Delta\lambda_{\mathrm{H}\beta}$ are noticeably correlated with EV1, which corroborates the findings of Z10, who stress that the asymmetry parameter could, in essence, be used as a surrogate 4DE1 parameter. Z10 note, however, that $\Delta\lambda_{\mathrm{H}\beta}$ is also substantially dominating the loading factors Eigenvector 2 (EV2), i.e. its "orientation" in parameter space. This behaviour is not observed in our PCA results. Our EV1 is consistent with early results in Grupe (2004). The second principal component, EV2, is strongly dominated by the X-ray and optical continuum luminosities. The flux ratio $\mathrm{F}([\mathrm{O\,III}])/\mathrm{F}(\mathrm{H}\beta)$ is correlated to EV2. $W(\mathrm{H}\beta)$ once again significantly contributes to the orientation of the principal component. EV2 is also consistent with Grupe (2004).

3.6 Black Hole Mass and Eddington ratio

We extend the analysis to derived AGN properties, specifically the black hole mass and the Eddington ratio. In Boroson & Green (1992), EV1 was related to the Eddington ratio L/L_{Edd}, tracing the accretion power of the observed SMBH. Their EV2 was largely dominated by luminosity and related to the Baldwin effect (Baldwin 1977; Baldwin et al.

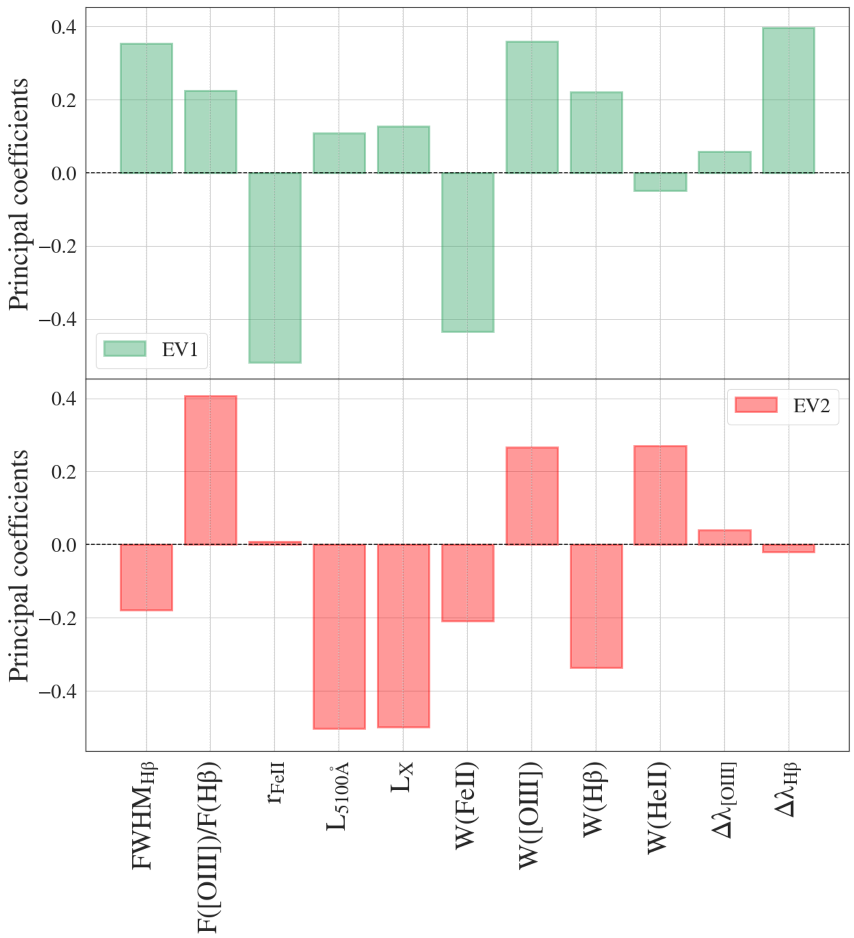


Figure 3.10: The component coefficients (factor loadings) that define the first and second principal components: EV1 (green bars) and EV2 (red bars). EV1 appears to be heavily dominated by the anti-correlation of FWHM$_{\mathrm{H}\beta}$ and the strength of the [O III] emission with r_{FeII}. A relatively large linear coefficient links the Hβ asymmetry to EV1. EV2 is heavily dominated in equal measure by the X-ray and optical luminosities.

1978; Dietrich et al. 2002). The Eddington ratio has indeed long been considered a prime candidate to explain the observed diversity in optical AGN features (e.g., Sulentic et al. (2000b); Yuan & Wills (2003); Grupe (2004); Kuraszkiewicz et al. (2009); Grupe (2011); Shen & Ho (2014); Bon et al. (2018)).

The black hole masses and the Eddington ratio for our sample are derived from Hβ using the calibration developed in Assef et al. (2011), who based their estimates on the BLR radius-luminosity relation of Bentz et al. (2009).

$$\log\left(\frac{M_{\rm BH}}{M\odot}\right) = A + B\log\left(\frac{\lambda L_\lambda}{10^{44}{\rm erg\,s^{-1}}}\right) + C\log\left(\frac{{\rm FWHM_{H\beta}}}{1000\,{\rm km\,s^{-1}}}\right) \tag{3.7}$$

L_λ corresponds to the monochromatic luminosity at 5100 Å and the parameters A, B and C are derived from reverberation mapping studies: $A = 0.895$, $B = 0.52$ and $C = 2$. From these black hole masses, C19 provided estimates of $\lambda_{\rm Edd} = L_{\rm bol}/L_{\rm Edd}$. The bolometric luminosities $L_{\rm bol}$ they used here are obtained from the bolometric corrections presented in Richards et al. (2006): $L_{\rm bol} = 9.26L_{5100\,\text{Å}}$. This factor is, in principle, dependent on $L_{5100\,\text{Å}}$, but for the purpose of this Chapter, this assumption is not crucial.
Fig. 3.11 displays how $\lambda_{\rm Edd}$ and $M_{\rm BH}$ scale with the projection of the sources onto our newly determined EV1 and EV2. We can strongly corroborate that EV1 correlates with $\lambda_{\rm Edd}$. An even stronger anti-correlation is that of EV2 with black hole mass. However, the log $M_{\rm BH}$-EV2 relation breaks down at lower redshifts ($z \sim 0.2$). In order to determine to which extent this correlation is a side-product of the $L_{5100\text{Å}}$ and FWHM$_{\rm H\beta}$ principal coefficients in EV2, we once again perform a partial correlation test. We start by measuring the Spearman rank correlation of EV2 with black hole mass while controlling for $L_{5100\,\text{Å}}$ and FWHM$_{\rm H\beta}$. In Fig. 3.12 the Spearman correlation coefficients measured for log $M_{\rm BH}$, EV2, $L_{5100\,\text{Å}}$ and FWHM$_{\rm H\beta}$ are displayed in the left panel. In the right panel of Fig. 3.12, the partial correlation coefficients results are presented. As already discussed in Section 3.5, this exercise consisted in performing a (multi) linear regression of EV2 and log $M_{\rm BH}$ with the confounding variables. In this version, we compute the Spearman correlation coefficients of the residuals of these two multi-linear regressions. EV2 is still strongly anti-correlated with $M_{\rm BH}$ ($\rho = -0.56$). This is not surprising since in Eq. 3.7, log $M_{\rm BH}$ is a non-linear function of FWHM$_{\rm H\beta}$ and $L_{5100\,\text{Å}}$, while the Eigenvectors obtained from PCA are obtained through linear orthogonal transformations of the initial parameter. The monotonic anti-correlation measured in the residuals might thus simply arise from the non-linearity of Eq. 3.7. More generally, this result implies that one can construct a linear combination of the parameters which dominate the orientation of EV2 (see Fig. 3.10) in order to estimate $M_{\rm BH}$ at redshifts $z > 0.2$.

We propose that EV2, through its anti-correlation to black hole mass, might be related to the evolution of the broad line AGN population. Over the observed redshift range, the median $M_{\rm BH}$ increases with decreasing redshift. This trend is a combination of two effects: the ROSAT/2RXS X-ray flux limit resulting in the Malmquist bias and the decreasing number density of high black hole masses at lower redshifts. This downsizing of black hole masses across cosmic time remains a matter of debate (cf. downsizing models,

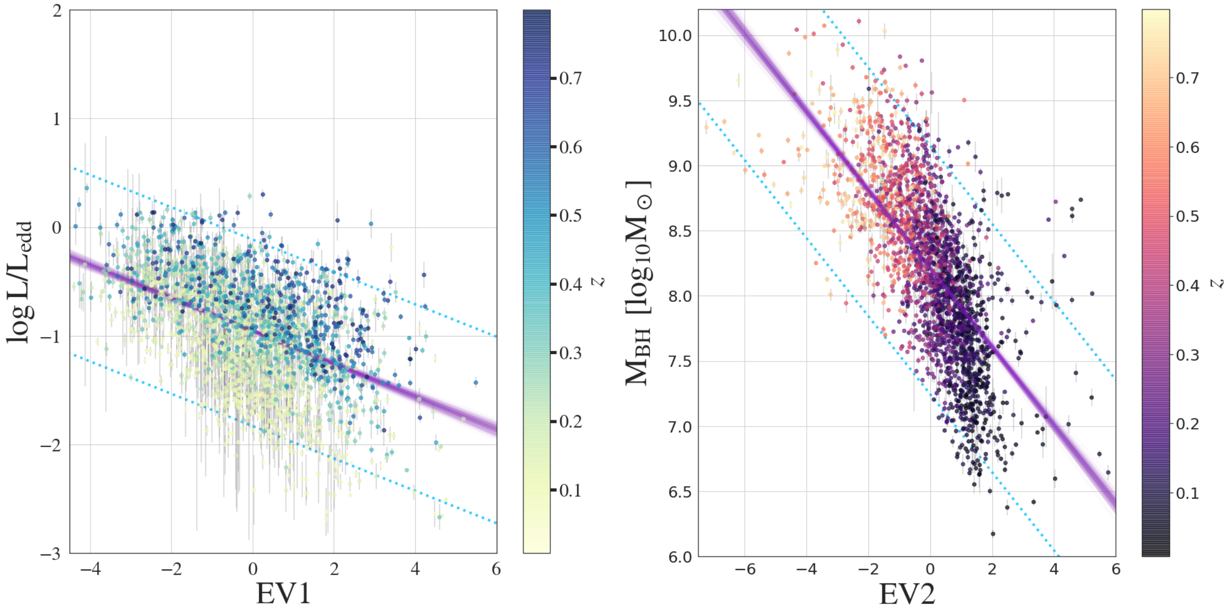


Figure 3.11: *Left panel*: $\lambda_{\rm Edd}$ has an exponential correlation with EV1. The 95% confidence contours are represented by the dashed blue lines. *Right panel*: $M_{\rm BH}$ is correlated with EV2. Again the 95% confidence contours are shown in blue. The linear regressions were performed with *Linmix* (Kelly 2007).

e.g., Fanidakis et al. 2012). We conclude that $\lambda_{\rm Edd}$ and $M_{\rm BH}$ are related to the principal components which explain $\sim$ 45% of the total variance in our data.

3.7 Asymmetry of the broad Hβ emission line; a marker of Type 1 AGN diversity

3.7.1 Asymmetries along the Eigenvector 1 plane

We investigate the role played by Hβ line asymmetries in this framework. Since Hβ asymmetries drive the total of variance (through their large factor loadings in EV1), we identify them as a dominant physical parameter along the quasar main sequence. Fig. 3.13 presents the sample in the FWHM$_{\rm H\beta}$ vs $r_{\rm FeII}$ plane, also called EV1 plane (it is spanned by the first two optical dimensions of the classical EV1 by Boroson & Green 1992). The sources have been separated into two subsets: blue and red asymmetric Hβ emission. More specifically these two sub-samples were constructed according to : $\Delta\lambda_{\rm H\beta} > 0.07$ and $\Delta\lambda_{\rm H\beta} < -0.07$. These criteria exclude symmetric sources and account for the typical uncertainties of the asymmetry index ($\sigma_{\Delta\lambda_{\rm H\beta}} \sim 0.05$). Contours of the bivariate Gaussian kernel density estimates (KDE, e.g., Silverman 1986) for each of the subsets are also indicated. The contours

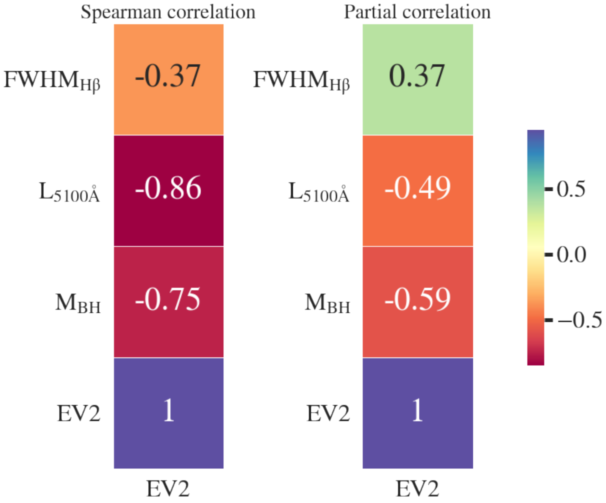


Figure 3.12: The monotonic correlation coefficients of EV2 with $M_{\rm BH}$, $L_{5100\,\text{Å}}$ and $\text{FWHM}_{\text{H}\beta}$ for our sample are shown: simple Spearman correlation in the left column and partial Spearman correlation in the right column. For the computation of the partial correlation of a parameter pair, we control for the other confounding variables. The $M_{\rm BH}$ vs EV2 anti-correlation slightly drops when we control for $L_{5100\,\text{Å}}$ and $\text{FWHM}_{\text{H}\beta}$, while partial $M_{\rm BH}$ vs EV2 anti-correlation remains significantly strong p $= -0.59$.

delimit the areas containing 68%, 86%, 95% and 98% of the data points.[9] The sample occupies an L-shaped form in the plane. Using simulations, C19 demonstrate that the absence of high $\text{FWHM}_{\text{H}\beta}$ and high r_{FeII} sources is to, an extent, due to limitations of the spectral fitting method. Most of the sources with blue asymmetric Hβ emission profiles and those with red-asymmetric profiles appear to occupy different domains of the $\text{FWHM}_{\text{H}\beta} - r_{\text{FeII}}$ plane. While the AGN with blue-ward asymmetric Hβ spread over the full r_{FeII} range and a large portion of the $\text{FWHM}_{\text{H}\beta}$ range, the red-ward asymmetric sources are concentrated at lower values of r_{FeII}, while dominating the higher segment of $\text{FWHM}_{\text{H}\beta}$. This result is essentially suggesting that high accretion rate sources, with relatively large r_{FeII} values, show mainly blue-ward asymmetric Hβ profiles. We also observe that both red and blue asymmetric line profiles are observed for sources with moderate $\text{FWHM}_{\text{H}\beta}$ and r_{FeII}.

In their in-depth study of Hβ line profiles of $\sim$ 470 low-z SDSS (DR5) quasars, Z10 show that low accretion rates sources (low r_{FeII}) possess a typically red-asymmetric Hβ profile, while high accretion rate sources tend to prefer blue-asymmetric Hβ profiles. The two panels of Fig. 3.14 can directly be compared to Fig. 6a and 9a presented in their work:

The size of our sample allows us to extend their results by affirming that blue asymmetric Hβ profiles can be found in sources with relatively high $\text{FWHM}_{\text{H}\beta}$ ($\sim 8000\,\text{km}\,\text{s}^{-1}$) and along the full EV1 sequence. There is, however, a clear decrease in blue asymmetric Balmer profiles at $\text{FWHM}_{\text{H}\beta} \sim 4000\,\text{km}\,\text{s}^{-1}$, consistent with the low-redshift separation in Population A/B introduced by Sulentic et al. (2000a). An excess of red-asymmetric Hβ

---

[9] The density contours shown throughout the rest of this Chapter are defined in the same way.

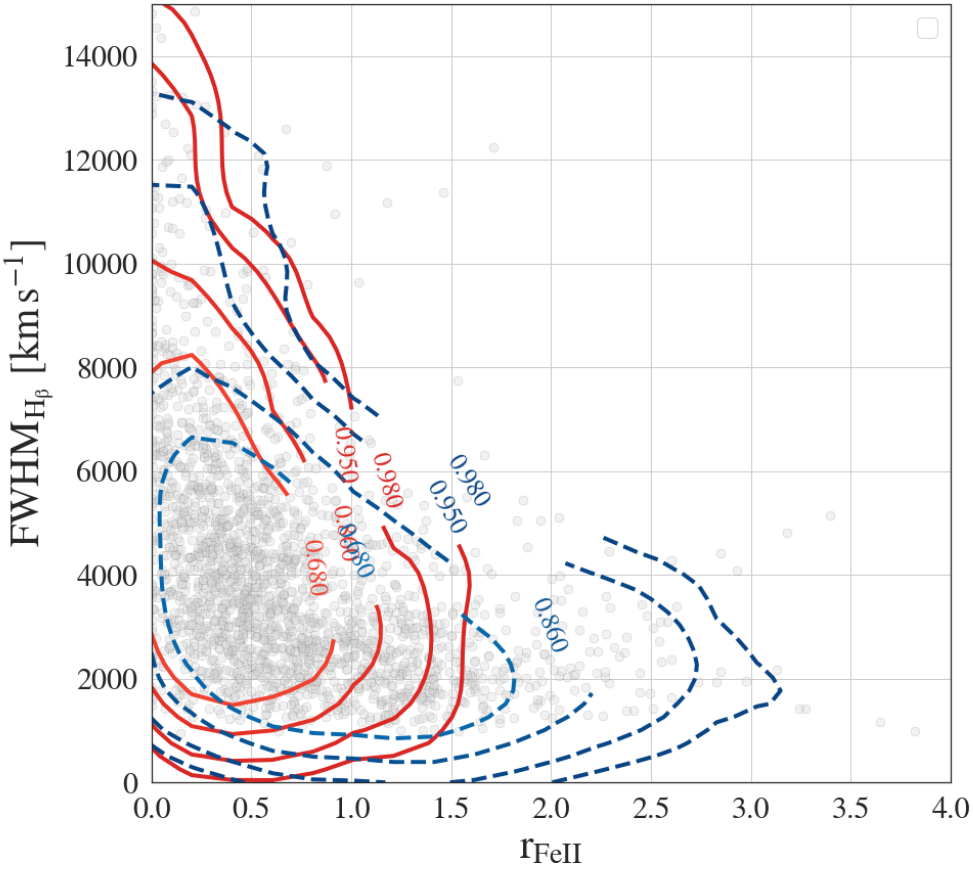


Figure 3.13: The horizontal trend of the Hβ asymmetry in the EV1 plane is made clear in this figure. The grey points represent the sources in our statistical sample. The contours delimit the areas in which 68%, 86%, 95% and 98% of the data points are confined. These contours reveal two very different occupation domains for red- and blue asymmetric Balmer emitters: sources with red-asymmetric Hβ are confined at low r_{FeII} values, while sources with blue asymmetric Hβ seem to extend over the full EV1 sequence.

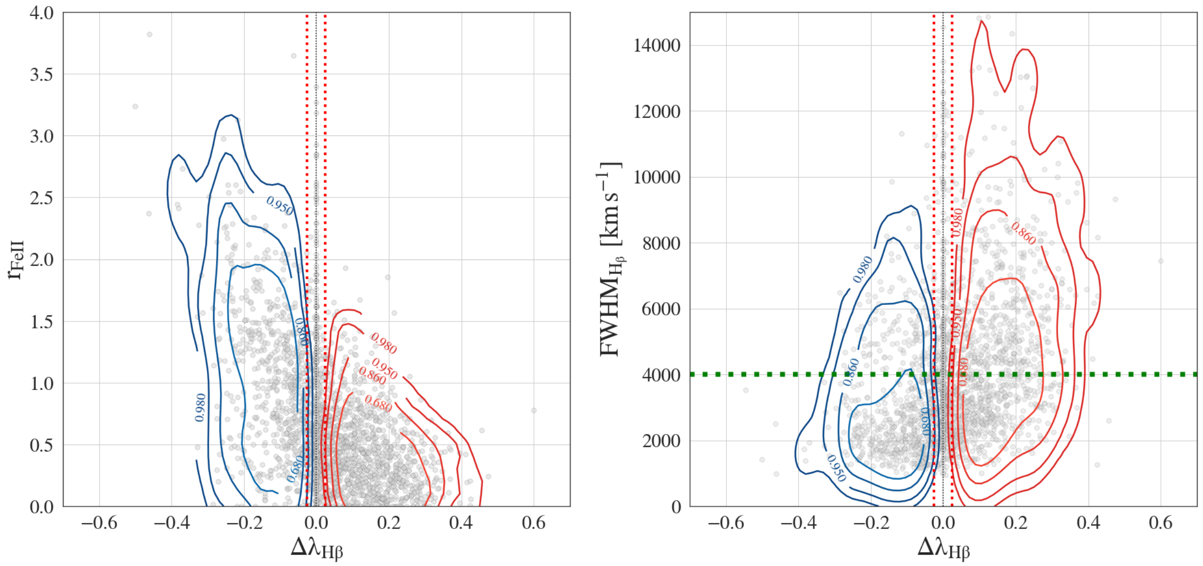


Figure 3.14: *Left Panel:* FWHM of Hβ as a function of the Balmer asymmetry index. The kernel density contours separate blue- from redward asymmetric Hβ emitting sources. The green dotted line indicates the separation in Pop.A/B at FWHM$_{\mathrm{H}\beta}$ $\sim 4000\,\mathrm{km\,s^{-1}}$. **Right Panel:** r_{FeII} as a function of the Balmer asymmetry index. The kernel density contours separate blue- from redward asymmetric Hβ emitting sources. For both figures, the typical 1σ errors are shown centred at zero by two dotted vertical bars.

profiles is indeed observed for larger widths. Sources with lower FWHM$_{\mathrm{H}\beta}$ do not, however, display more symmetric profiles. We thus argue that there is no evidence for a systematic relationship between FWHM$_{\mathrm{H}\beta}$ and $\Delta\lambda$. The excess of red asymmetric Hβ emission for larger widths might be the signature of an additional redshifted broad emission component.

The right panel of Fig. 3.14 strongly confirms the decreasing trend presented in figure 5 of Boroson & Green (1992) and figure 9a of Z10. Combined with the Spearman correlation coefficient of $\rho \sim -0.5$ reported in the previous section, we can state that, up to redshift of $z \sim 0.8$, relatively slow accretors tend to show more red asymmetric profiles, with a significant decrease of red-asymmetric Hβ emitters beyond $r_{\mathrm{FeII}} \sim 1.5$. While blue-asymmetric profiles are reported over the full r_{FeII} range, the shifts become more pronounced at higher values of the FeII-Hβ flux ratio.

Z10 discuss the VBLR-emission as a potential origin of typically red asymmetries in Hβ (3.3.2). In a layered model of the BLR, the asymmetry index traces the displacement between the Hβ BC and its VBC. Hβ lines often show a broad, redshifted, low-intensity feature in their profiles (Peterson & Ferland 1986; Corbin 1995; Brotherton 1996). The presence of a distinct emission region, the VBLR (Popović et al. 2004; Marziani et al. 2010), has been proposed to explain the characteristic red wing of broad Balmer emission. The VBLR is expected to be located in the inner region of the AGN: The typical widths of the VBC in Hβ indicate much higher Keplerian velocities than the classical broad line region.[10]

3.7.2 Tracing broad component displacements with centroid shifts

In order to further characterise the kinematics of two potentially distinct emitting regions, we compute the centroid shifts of the Hβ profiles at 15% (c_{15}) and 80% (c_{80}) fractional intensity. Fig. 3.15 displays the correlations between these shifts and the asymmetry index of Hβ, colour-coded according to the FWHM$_{\mathrm{H}\beta}$. The shifts with respect to the restframe at 80% fractional intensity are negatively related to the asymmetry index, i.e., the more blue shifted the top of the profile, the more red-ward asymmetric the complete profile and vice-versa. Symmetric profiles range over the full shift range. For the centroid shifts at c_{15} fractional intensity, an opposite and stronger trend is observed. c_{15} correlates positively with $\Delta\lambda_{\mathrm{H}\beta}$: a red-ward asymmetric Hβ profile can be associated with a redshift of the broad component. As expected, sources with symmetric Hβ span over the full c_{15} range. The highest values of FWHM$_{\mathrm{H}\beta}$ occur for red asymmetric Hβ consistent with the results in Fig. 3.14. A clear picture arises from the c_{15} distribution in the EV1 plane (Fig. 3.16). The absolute value of the shifts at the profile base appears to decrease along the EV1-sequence, while c_{15} shows clear trends with both r_{FeII} and FWHM$_{\mathrm{H}\beta}$. This picture is consistent with the presence of a redshifted VBLR. The base shifts of the Hβ profiles in our sample are tightly correlated with the asymmetry index $\Delta\lambda_{\mathrm{H}\beta}$. The largest shifts are observed for the most red-asymmetric profiles. The highest base component redshifts are found at the top of the Type 1 AGN main sequence, i.e., at the largest widths of Hβ; the broadest components in our sources' Hβ profiles are preferentially redshifted.

The correlation of c_{15} with EV1 provides additional information: as can be seen in Fig. 3.17, the positive base centroid shifts appear to correlate with EV1. Interestingly, the broad redshifted Hβ components are found for the sources with the highest black hole masses. The broad component redshifts are thus a source of variance in our sample. If the redshifted VBCs are interpreted as the signature of a VBLR, we can argue that the presence or absence of a distinct, inner emitting shell in the BLR (or, more precisely, its redshift) is a strong source of diversity in our sample of Type 1 AGN. Furthermore, we report that our results are consistent with Marziani et al. (2009), who find a systematic increase of FWHM$_{\mathrm{H}\beta}$ with source luminosity. More precisely, Marziani et al. (2009) find a scaling relation between the luminosity and the contribution of the VBLR to the full Hβ flux in population B ($> 4000\,\mathrm{km\,s^{-1}}$). The absence of red asymmetric Hβ lines at higher r_{FeII} shown in the right panel of Fig. 3.14 might be due to strong accretion disk winds preventing the formation of a VBLR.

In the following section, we explore a toy model of BLR obscuration, which accounts for the presence of a VBLR.

---

[10] The presence of a VBC in the profile shapes has to be corrected for when broad low ionization lines are used as a virial broadening estimator (Marziani et al. 2013a).

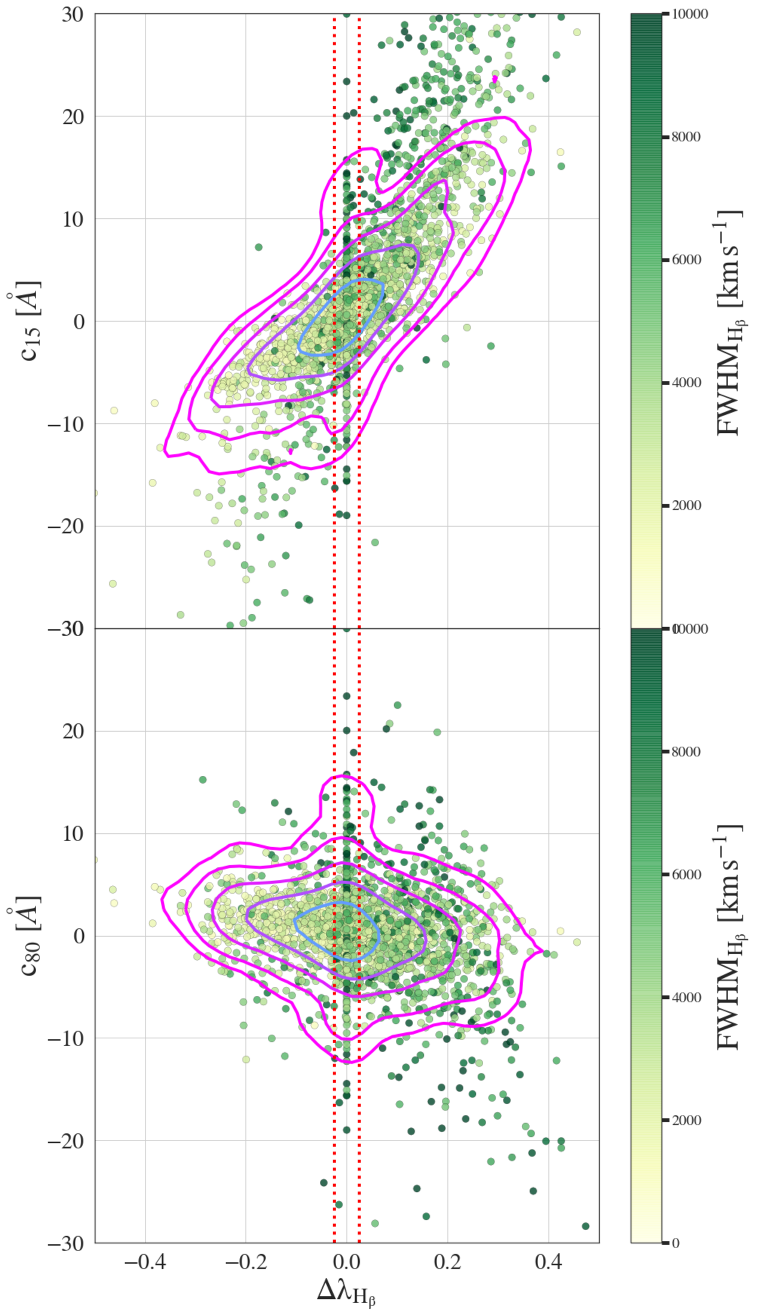


Figure 3.15: The Hβ centroid shifts at 15% and 80% fractional intensity as a function of Hβ asymmetry. $\mathbf{c_{15}}$ appears more tightly related to $\Delta\lambda_{\mathrm{H}\beta}$ than c_{80}. The kernel density contours show the sub-sample of sources, for which Hβ has not been fit with broad Gaussians (FWHM $<$ 10.000 km s^{-1}). Typical 1σ uncertainties in the measurement of $\Delta\lambda_{\mathrm{H}\beta}$ are indicated as vertical red dotted lines.

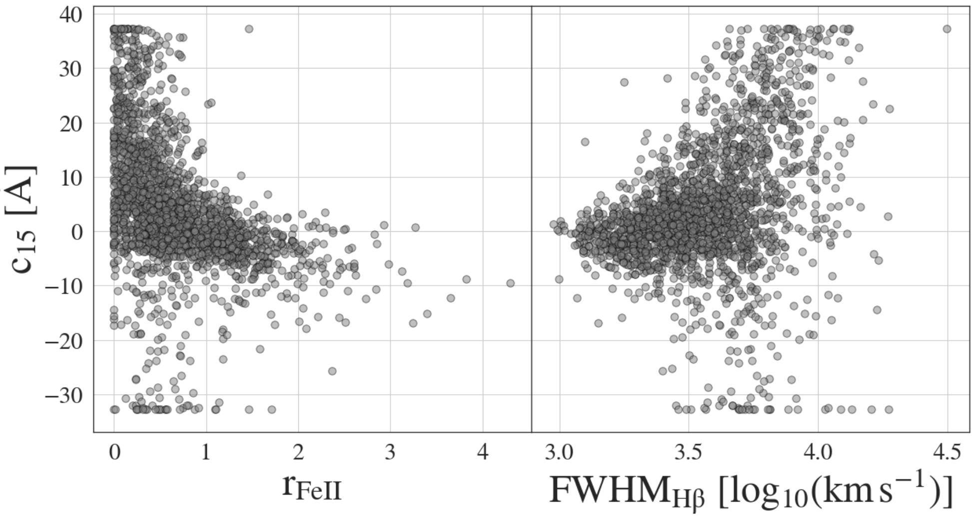


Figure 3.16: The individual correlations of $\mathbf{c_{15}}$ with the two optical 4DE1 dimensions. The absolute centroid shifts decrease along the EV1 sequence.

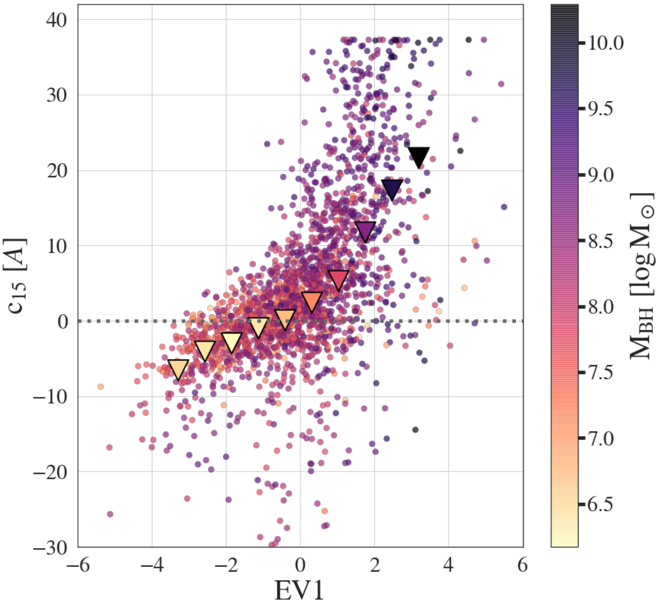


Figure 3.17: Hβ base centroid shifts ($\mathbf{c_{15}}$) projected onto EV1. The points are colour-coded according to the black hole masses estimated with the calibration from Assef et al. (2011). Positive centroid shifts (redshifts) correlate with EV1. Coloured triangles represent the median of $\mathbf{c_{15}}$ in EV1 bins and have been added to improve the visualisation of this trend. Their colour represents the median black hole masses in the bins, with darker colours corresponding to higher black hole masses. The highest black hole masses are found for redshifted base components.

3.8 Testing a simple obscuration scenario

Richards et al. (2002) proposed that the spatial configuration resulting from the BLR's inclination and obscuration might explain the asymmetries in the observed profile of CIVλ1549Å (see also 4.3. in Z10). In this picture, the asymmetry arises from the attenuation of redshifted photons by obscuring material. Similarly, Gaskell & Harrington (2018) developed a model of outflowing clouds of dust, blocking the line of sight to the inner disk-shaped BLR, which predicts broad emission profiles similar to those observed. An alternative view of the effect of the BLR and outflows is developed by Czerny et al. (2017) (see also Czerny & Hryniewicz 2010). In this model, the dusty clouds are driven out of an optically thick disk and are eventually exposed to radiation by the central source. The dust evaporates, removing the matter/radiation interaction and resulting in the fall-back of gas clouds. This scenario is referred to as Failed Radiatively Accelerated Dusty Outflow (FRADO). The physics of dust sublimation and the consequences for the BLR disk structure were further investigated by Baskin & Laor (2018).

Fig. 3.15 allows us to investigate a simple obscuration scenario in which the observed asymmetry of the line shapes of Hβ would result from partial obscuration of a flattened BLR in Keplerian motion around the black hole. If the receding or incoming region (with respect to our line of sight) of a disk-like structure is obscured by optically thick material, the emitted broad lines would be respectively red- or blue asymmetric (see, for example, Fig. 2 in Gaskell & Harrington 2018), since one peak of the double-peaked emission profile would be attenuated. In this scenario, the low-intensity, 'excess' emission measured on the blue (or red) side of the line would correspond to the attenuated blue (or red) horn of the disk-like emission.

Furthermore, if we assume no further kinematics in our BLR model, the attenuation of the red of blue peak of the diske-like emission would lead to a centroid shift configuration :

i Blue-ward asymmetric Hβ ($\Delta\lambda_{\mathrm{H}\beta} < 0$): Redshift of the peak component ($c_{80} > 0$) and blueshift of the base component ($c_{15} < 0$)

ii Red-ward asymmetric Hβ ($\Delta\lambda_{\mathrm{H}\beta} > 0$): Blueshift of the peak component ($c_{80} < 0$) and redshift of the base component ($c_{15} > 0$)

Fig.3.15 supports this scenario. The results in the previous section suggest that the broad Balmer shapes might be affected by the presence of a distinct emitting region: the VBLR. The next step is to test if the obscuration scenario is compatible with the presence of a very broad component related to the stratification of the BLR. To this effect, we identify sources for which the broad Hβ has been fit with very broad Gaussians.

Fig. 3.18 presents the peak wavelength of all the Gaussians which have been used to fit the components of Hβ as a function of their FWHM. We clearly identify a sequence of very broad Gaussians at velocities $> 10.000\,\mathrm{km\,s^{-1}}$, which are mostly redshifted with respect to the systemic redshift. We investigated the reduced χ^2 distribution of the fits of the sources which have at least one Gaussian component of FWHM > 10.000 km.s^{-1}.

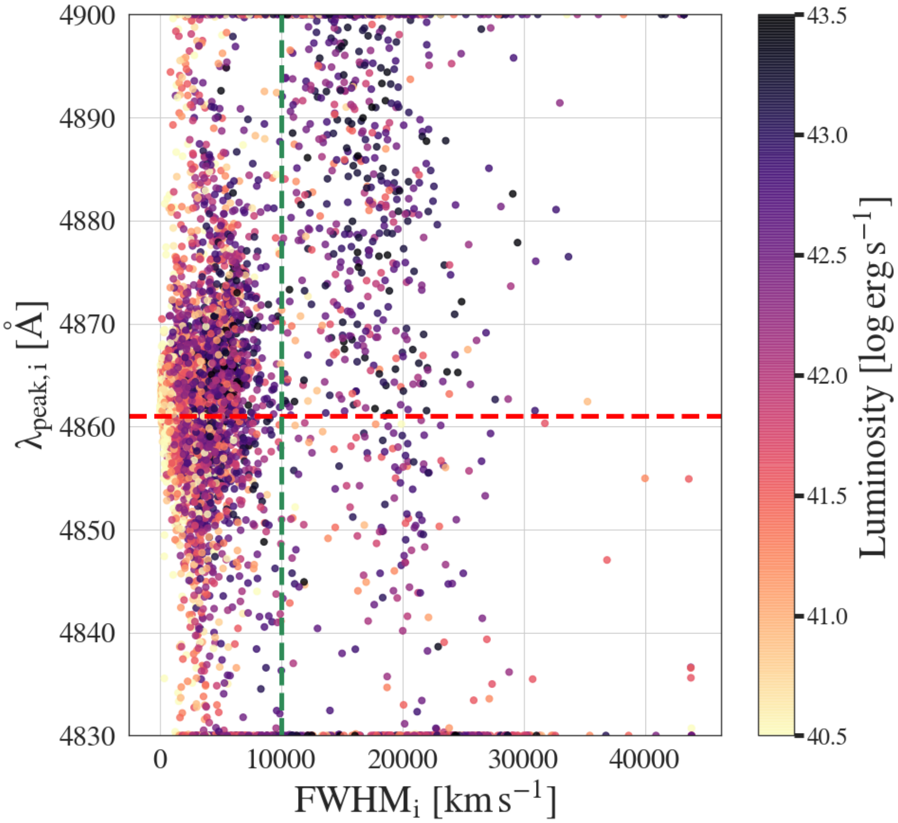


Figure 3.18: The peak wavelength of the Gaussians used to fit the Hβ components as a function of their FWHM. The red-dotted line indicates the systemic wavelength of Hβ. The green dotted line shows an empirical separation of broad and very broad Gaussian components at 10.000 km s^{-1}. The right side of this separation clearly reveals a sequence of very broad, mainly redshifted Gaussians. The points are colour-coded according to the luminosity from each Gaussian. Each source in our sample has four points in this figure (one for each of its Gaussian).

The sources which have a very broad component are not biased towards higher reduced χ^2 (and thus higher BIC) values, indicating that they are not preferentially found in less secure fits.

The FWHM of broad Gaussian components ($> 800\,\mathrm{km\,s^{-1}}$) that were fit to the H$\beta$ emission lines show a bimodal distribution. The two modes appear to distinguish broad components (800 $\mathrm{km\,s^{-1}} < \mathrm{FWHM} < 10.000\,\mathrm{km\,s^{-1}}$) and very broad components ($\mathrm{FWHM} > 10.000\,\mathrm{km\,s^{-1}}$). Out of the 8400 Gaussian components used to fit the 2100 sources of the sample, 4936 (58%) can be considered broad, with $\mathrm{FWHM} > 800\,\mathrm{km\,s^{-1}}$ and 1080 (13%) are very broad components with $\mathrm{FWHM} > 10.000\,\mathrm{km\,s^{-1}}$. The remaining components are narrow ($\mathrm{FWHM} > 800\,\mathrm{km\,s^{-1}}$). The statistical significance of a second mode in the distribution of the Gaussian widths at $\mathrm{FWHM} > 10.000\,\mathrm{km\,s^{-1}}$ was assessed with a Silverman test of multimodality (Silverman 1981, for its calibration, see, e.g., Hall & York 2001; Ameijeiras-Alonso et al. 2016). We present the outline of the test here.

i We extract all the Gaussians used in the fits of our sample with the condition: $\mathrm{FWHM} > 800\,km\,s^{-1}$ & $\lambda_{peak} \neq 4830\text{Å}$ & $\lambda_{peak} \neq 4900\text{Å}$. This cut removes narrow Gaussians, as well as Gaussians which have central wavelengths shifted to the edge of the fitting window. 4275 Gaussian components remain. We show the resulting density histogram of standardised FWHMs[11] in Fig. 3.19.

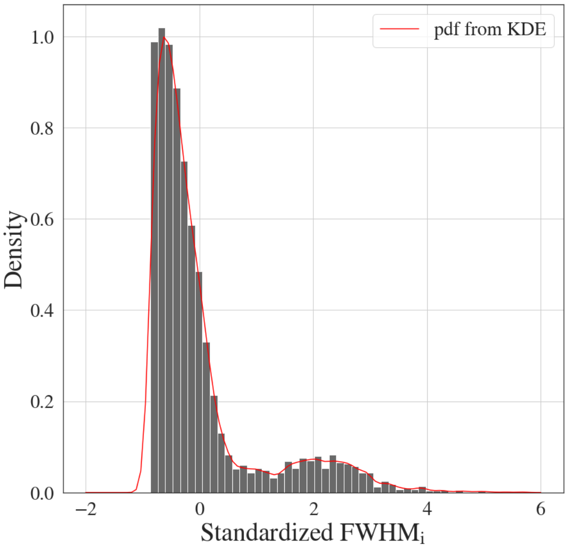


Figure 3.19: Density of standardised Gaussian component FWHM.

ii For a range of bandwidths (or smoothing scales) $h = [0.01, 2.00]$, we performed a Gaussian kernel density estimate (KDE) of the distribution of FWHM (centred at zero and scaled to unit variance). For n observations in our sample, the bandwidth h

---

[11] Here "standardised" refers to mean-subtracted values, normalised by the standard deviation.

smooths the Gaussian kernel estimate as follows:

$$\hat{f}(x,h) = \frac{1}{nh}\sum_{i=1}^{n}\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-X_i)^2}{2h^2}} \tag{3.8}$$

An example probability density function (PDF) from KDE is displayed in Fig. 3.19 (red line).

iii We then detect the critical bandwidths $h_{c,j}^{\mathrm{DATA}}$, at which new local maxima appear in the pdf, when h is decreased. For $1 < j < n$ (where n corresponds to the number of observations), the critical bandwidths $h_{c,j}^{\mathrm{DATA}}$ are formally defined as the minimum h, at which the Gaussian density estimate has no more than j maxima. In Fig. 3.20, the bandwidth h and the corresponding number of detected local maxima j are displayed. Critical widths are symbolized by vertical lines.

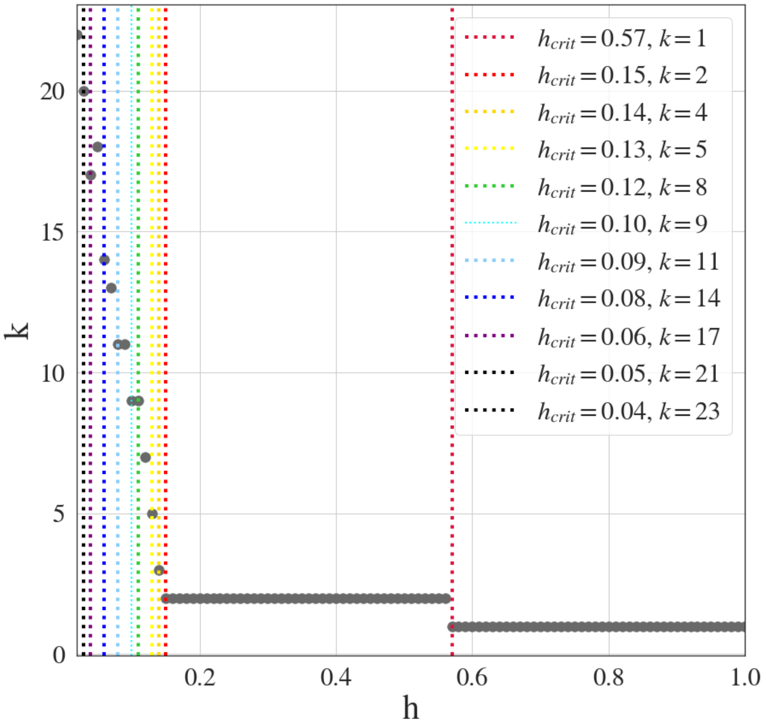


Figure 3.20: Number of local maxima k of the estimated PDF, as a function of the smoothing h, the bandwidth of the Gaussian KDE. The critical values $h_{c,j}^{\mathrm{DATA}}$ are marked as vertical lines.

iv In order to assess the significance of a mode detection, we follow Fischer et al. (1994) and set up a null hypothesis test, which is based on the observation that the number of modes of $\hat{f}(x)$ decreases with increasing bandwidth h. If k is the true number of modes:

$\mathrm{h_j^{DATA}}$	Number of local maxima	p-value
0.57	1	0.100
0.15	2	0.001
0.14	3	0.008
0.13	5	0.041
0.12	8	0.095
...	...	...

Table 3.2: p-values for mode detections for decreasing critical bandwidths.

$$\text{Null hypothesis } H_0 : k = j$$
$$\text{Alternative } H_1 : k \geq j + 1 \text{ for } j = 1, 2...$$

For each $h_{c,j}^{\mathrm{DATA}}$ detected in our data, we perform a Gaussian kernel defromnsity estimate on 100 smoothed bootstrapped samples, using $h_{c,j}^{\mathrm{DATA}}$ as bandwidth. These resamplings are obtained by sampling smoothed FWHM$_s$, using bootstrapping (with replacement). The exact computation of FWHM$_s$ is as follows:

$$\mathrm{FWHM_s} = \frac{1}{\sqrt{1 + {h_{c,j}}^2/\sigma^2}} (\mathrm{FWHM_r} + h_{c,j}\epsilon) \tag{3.9}$$

were FWHM$_\mathrm{r}$ are resampled FWHM, σ, the standard deviation of the bootstrapped sample and ϵ is a random standard variable. For the null-hypothesis test Silverman (1981) motivates a p-value defined for a given number of maxima j as:

$$p = P\{h_{c,j}^{\mathrm{NEW}} > h_{c,j}^{\mathrm{DATA}}\} = P\{\hat{f}(x, h_{c,j}^{\mathrm{DATA}}) \text{ has } > j \text{ modes}\}, \tag{3.10}$$

where $h_{c,j}^{\mathrm{NEW}}$ is the critical value for j maxima derived from the resamples, $h_{c,j}^{\mathrm{DATA}}$ the initial critical width from the data, $\hat{f}(x, h_{c,j}^{\mathrm{DATA}})$ the kernel density estimate of the resamples using $h_{c,j}^{\mathrm{DATA}}$ as critical density. The null hypothesis is rejected if $\mathrm{P\{h_{c,j}^{NEW} > h_{c,j}^{DATA}\}} < \alpha$, where α is the chosen significance threshold. For the present application we assume a threshold $\alpha = 0.05$. For each $h_{c,j}^{\mathrm{DATA}}$, and each associated number of modes we compute a p-value. We assess how many times the density estimates with fixed bandwidths of the 100 resamples have at most j detected modes. The ratio of these realizations and the total number of simulations are taken as p-value. The resulting p-values are listed in Table 3.2.

v We read the Table 3.2 as a *set of successive significance tests*: For $\mathrm{h_j^{DATA}} = 0.57$, there is only one local maximum and the p-value is p= 0.100. Similarly of for $\mathrm{h_j^{DATA}} = 0.15$, there are two local maxima and the p-value is p= 0.001. Following Silverman (1981), we select the lowest $\mathrm{h_j^{DATA}}$, for which the significance threshold $\mathrm{p} < \alpha = 0.05$ is met. As for two modes we have $\mathrm{p} < 0.05$, the test supports the bimodality of the FWHM distribution, i.e. the very broad mode of Gaussian components is significant.

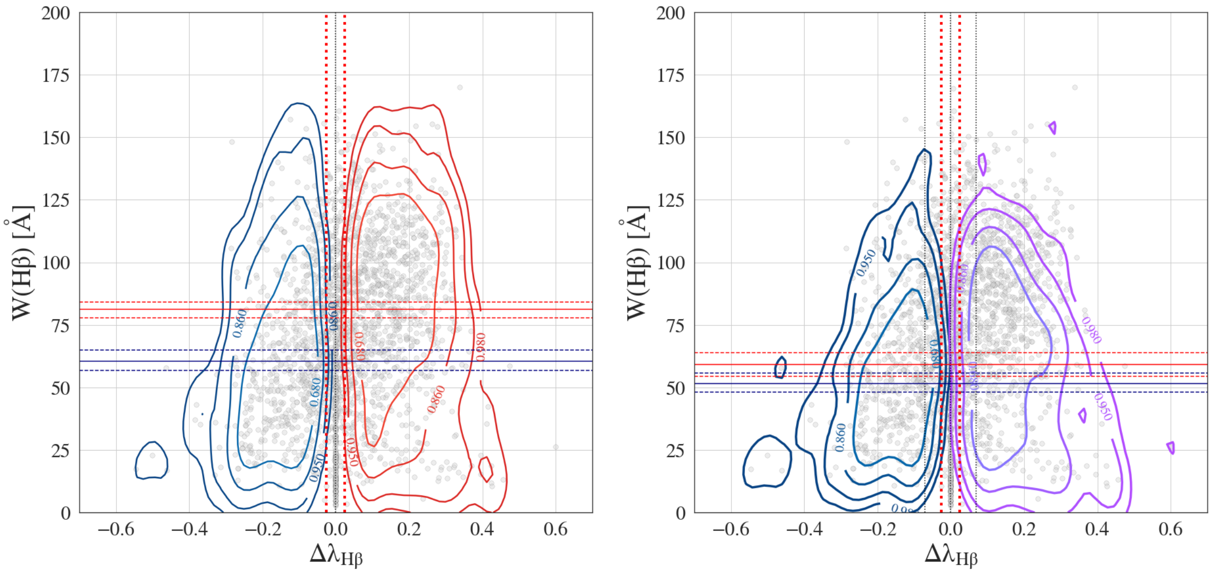


Figure 3.21: *Left panel*: The equivalent width of the broad Hβ emission as a function of its asymmetry Index. The blue and red density contours are included to improve the visualisation of the red- and blue asymmetric sub-samples. An offset between the two populations is observed. The horizontal red (blue) line marks the mean of the bootstrapped mean W(Hβ) of the red- (blue-)ward asymmetric sources. The red and blue horizontal dashed lines correspond to the 3σ confidence contours, as derived from the percentile method at 0.13% and 99.87% of the sampling distribution of the mean. *Right panel*: Same format as on the left. The blue and purple contours respectively represent the sources with blue- and red-asymmetric Hβ, for which the Gaussian models had FWHM $< 10.000\,\mathrm{km\,s^{-1}}$.

Fig.3.15 includes the density contours containing the sources for which Hβ has not been fitted with a very broad Gaussian (FWHM $< 10.000\,\mathrm{km\,s^{-1}}$, i.e., the sources to the left of the vertical dashed line in Fig.3.18). The trend observed for all objects in our sample is preserved for this subset, supporting the possibility of a partially obscured flattened BLR + VBLR model. The red- and blueshifts of the Hβ base and peak components are symmetrically distributed, which is consistent with the obscuration scenario.

In a simple model in which wing attenuation via partial obscuration is the sole source of asymmetry of the broad emission lines, we expect the Hβ line fluxes to be symmetrically distributed for $\Delta\lambda_{H\beta} < 0$ and $\Delta\lambda_{H\beta} > 0$, as no wing of the emission line profile should be preferentially attenuated. The left panel in Fig.3.21 shows the equivalent width of Hβ plotted against its asymmetry index for our full sample. We observe an offset in W(Hβ) between the population of red- and blue-asymmetric emitters. We argue that this offset might once again be the signature of the presence of a VBLR in some of our sources. If we display the kernel density contours for the subset of sources which were not fitted with a very broad Gaussian, a more symmetric distribution of equivalent widths is produced (i.e., line fluxes), as shown in the right panel of Fig. 3.21.

These simple tests have demonstrated that our sample is consistent with a model of partial obscuration coupled to a stratified BLR (one that might or might not contain a VBLR).

3.9 Blue-asymmetric Hβ: Outflows in a flattened and stratified BLR model

3.9.1 Blue asymmetries, outflows and self-shielding

If we consider that blue- and red-ward asymmetric Hβ profiles are signatures of two distinct kinematic states of the BLR, we can divide our sample according to the asymmetry indices of the sources and investigate the properties of each sub-sample. As in the previous section, we formed two sub-samples with $\Delta\lambda_{\mathrm{H}\beta} < -0.07$ (blue asymmetric) and $\Delta\lambda_{\mathrm{H}\beta} > 0.07$ (red asymmetric). We constructed a Spearman rank correlation matrix for the parameter subset, as defined in Section 3.3.1, of each sub-sample and directly compared their correlation spaces. A striking difference is observed in the correlation of the optical and X-ray luminosity and the iron emission. For blue asymmetric Hβ emitters, L_{X} and $L_{5100\mathrm{\AA}}$ correlate positively with the equivalent width of the iron emission W(FeII) (Spearman correlation coefficients: $\rho_S = 0.47$ and $\rho_S = 0.55$, respectively). The sources with red asymmetric Hβ have a much weaker correlation between the source luminosities and W(FeII) ($\rho_S = 0.023$ for L_{X} and $\rho_S = 0.097$ for $L_{5100\mathrm{\AA}}$). The correlation of $L_{5100\mathrm{\AA}}$ and W(FeII) for red-asymmetric sources is $p = 0.53$ and can safely be considered insignificant. This contrasting behaviour is presented in the left panel of Fig. 3.22. In order to confirm the different clustering of red- and blue-asymmetric Hβ emitters in the plane spanned by $L_{5100\,\mathrm{\AA}}$ and W(FeII), we binned log W(FeII) in the range $1.0 - 1.9$ and bootstrapped the sub-samples in each bin (10000 resamples). We obtained the mean $L_{5100\,\mathrm{\AA}}$ for each bin from the sampling distribution of the means. The 68% confidence intervals were derived using the percentile method (from the 16.0% and 84.0% percentiles). This method should yield relatively good estimates of the error, given the symmetric shape of the sampling distribution of the mean in each bin (shown in the right panel of Fig. 3.22). The positive correlation between $L_{5100\mathrm{\AA}}$ and W(FeII) for sources with blue asymmetric Hβ is confirmed. No trend is found for sources with red-asymmetric Hβ. The error bars, showing the standard deviation of the bootstrap samples' means, do not overlap. The correlation of $L_{5100\mathrm{\AA}}$ and W(FeII) for the blue asymmetric Hβ population withstood a partial correlation test, where we controlled for the confounding $W(\mathrm{H}\beta)$. Similarly, when we control for the FWHM$_{\mathrm{FeII}}$ of the Gaussian kernel used to fit FeII, $L_{5100\mathrm{\AA}}$ and W(FeII) remain positively correlated. The dependency of the FeII flux on the continuum luminosity in blue-asymmetric Hβ emitters is only marginally related to the broadening of the iron lines. This effect also manifests itself in the correlation of the luminosity parameters with the flux ratio r_{FeII}. Blue-asymmetric Hβ sources have significantly correlated r_{FeII} and source luminosities ($\rho_S = 0.20$ for $L_{5100\mathrm{\AA}}$ and $\rho_S = 0.19$ for L_{X}). For red asymmetric Hβ sources, these quantities are anti-correlated ($\rho_S = -0.16$ for $L_{5100\mathrm{\AA}}$ and $\rho_S = -0.25$ for L_{X}). Another noticeable difference is observed for the $\Delta\lambda_{[\mathrm{O\,III}]}$ parameter. In the sub-sample of blue-asymmetric Balmer emitters, $\Delta\lambda_{[\mathrm{O\,III}]}$ and $F([\mathrm{O\,III}])/\mathrm{F}(\mathrm{H}\beta)$ are anti-correlated ($\rho_S = -0.26$). If blue shifted wings are interpreted as the signature of NLR outflows, this result implies that outflow velocities increase with increasing $F([\mathrm{O\,III}])/\mathrm{F}(\mathrm{H}\beta)$. In the red-asymmetric sub-sample, these two parameters are positively correlated ($\rho_S = 0.17$).

Blue-ward asymmetries of low-ionization lines have been related to radiation-driven outflows (Marziani et al. 2013b). They occur in the high $r_{\mathrm{Fe\,II}}$ bins of population A (e.g. Ganci et al. 2019), which contain the highest accretion rate sources along the main sequence (e.g. Sun & Shen 2015; Sulentic et al. 2017; Panda et al. 2019c). The origin of FeII emission in the BLR has been a long-standing matter of research. Several lines of evidence support an emission of FeII in the outer parts of the BLR, while Hβ might be emitted closer to the black hole (e.g. Rodríguez-Ardila et al. 2002; Barth et al. 2013; Marinello et al. 2016). The physical conditions for iron ionization have been investigated in detail. Wills et al. (1985) showed that photoionization models might not suffice to account for the total observed FeII flux in AGN. For a review of necessary conditions for the formation of the observed FeII in photoionization models, see Collin & Joly (2000). Models of collisional excitation (e.g. Baldwin et al. 2004; Joly et al. 2007), in addition to continuum and line fluorescence (e.g. Sigut & Pradhan 1998, 2003; Marinello et al. 2016), have been considered. These models make predictions on the physical conditions of the FeII emitting region, such as shielding from the continuum source and high densities. A more in-depth study of FeII emission strength should also account for the orientation of the flattened BLR with respect to the observer. This has been recently investigated for Narrow Line Seyfert 1 objects by Panda et al. (2019b,a), whose preliminary analysis leads to conclusions on the density of the emitting region.

In a flattened, horizontally stratified cloud distribution in Keplerian motion around the central black hole, the difference in response to the continuum radiation of the FeII emitting region might arise from different degrees of exposure to the central continuum. Radiation-driven winds produce outflows of Hβ emitting gas, thereby exposing the previously shielded FeII regions more directly to continuum emission. Using the photoionization code CLOUDY (Ferland et al. 2013), Panda et al. (2017) found a decrease of the flux ratio $F_{\mathrm{Fe\,II}}/F_{\mathrm{H}\beta}$ with increasing disk temperature. A similar experiment accounting for BLR stratification and outflowing components would allow testing the scenario we propose. The scaling between $W(\mathrm{Fe\,II})$ and the $L_{5100\text{Å}}$ might be due to the contribution of UV fluorescence to the excitation of Fe+ levels. Investigating FeII fluorescence in HII regions, Rodríguez (1999) uses the sensitivity to the fluorescence of [FeII]λ4287, as well as the continuum insensitive [FeII]λ8617 to investigate the role played by the UV radiation field in the formation of this line. If the proposed dichotomy of shielded vs non-shielded FeII region holds, the intensity ratio $I([\mathrm{Fe\,II}]\lambda4287)/I([\mathrm{Fe\,II}]\lambda8617)$ should scale differently with the UV/optical continuum luminosity for the red- and blue-asymmetric Hβ populations. This may be investigated in further work. In this scenario, the distance of the FeII region to the SMBH should not depend on the presence of outflows. We performed a two-sample Anderson-Darling test on the distributions of the $\mathrm{FWHM}_{\mathrm{Fe\,II}}$ of the Gaussian kernel, which was convolved with the FeII template in the red- and blue-asymmetric Hβ subsets. 55

sources in our sample have no FeII flux detection and were not considered for this test. The obtained p-value is $p > 0.25$. We thus cannot reject that the $\mathrm{FWHM}_{\mathrm{FeII}}$ for red- and blue-asymmetric Hβ sources were sampled from the same distribution. Under the assumption of a Keplerian velocity field, the radial distribution of the FeII emitting clouds is not affected by the presence of outflows.

We conclude that our data are consistent with a flat, stratified and self-shielding model of the BLR, in which the FeII and Hβ emissions originate at different radii. When the inner clouds are driven out of the plane by outflows during strong accretion events, the previously shielded, neutral and dense FeII clouds are more exposed to the ionizing continuum radiation, which is clearly seen in the FeII flux-luminosity scaling.

3.9.2 Evidence for model degeneracy: FeII vs Hβ

We note that C19 fitted the emission of the iron complex over the full Hβ centred region, using the I Zw 1 template (Boroson & Green 1992). Since all the emission features of FeII in the spectral window were taken into account during the fits, the effect of contamination by Hβ should be marginal on the total FeII flux, i.e. W(FeII) should not be strongly affected by increasing Hβ line flux. However, we can investigate if the FeII contaminates the Hβ fit on its blue side. For blue asymmetric Hβ emitters, we seek to test if the W(FeII) correlates with $\Delta\lambda_{\mathrm{H}\beta}$ asymmetry index, i.e. we want to determine if the excess flux on the blue side of Hβ is due to contamination by iron. On the sub-sample with, $\Delta\lambda_{\mathrm{H}\beta} < -0.07$, we measure the Spearman rank correlation coefficient of $\mid \Delta\lambda_{\mathrm{H}\beta} \mid$ and W(FeII). We obtain $\rho = 0.087$ with a p-value of $p = 0.08$. Even if we do not correct the significance threshold for multi-hypothesis testing, one can state that there is no significant correlation between $\Delta\lambda_{\mathrm{H}\beta}$ and W(FeII). We note however that $W(\mathrm{H}\beta)$ is a confounding variable for both: $\Delta\lambda_{\mathrm{H}\beta}$ ($p = -0.26$) and W(FeII) ($p = 0.29$). We thus have to once again perform a partial correlation analysis, marginalising over $W(\mathrm{H}\beta)$ in order to obtain the real correlation behaviour of the blue-asymmetries and the iron strength. The partial correlation coefficient is $p_P = 0.17$ with a p-value of $p \sim 10^{-4}$. There is a significant correlation between $\Delta\lambda_{\mathrm{H}\beta}$ and W(FeII) once we control for $W(\mathrm{H}\beta)$. We thus cannot exclude that the iron complex contaminates the blue wing of Hβ.

3.10 Conclusions

In this Chapter, we have performed a statistical analysis of a high S/N (> 20) sub-sample of the SDSS-IV/SPIDERS DR14 VAC of X-ray selected Type 1 AGN. The sample included 2043 sources spanning an X-ray luminosity range of $\mathrm{L_X} = 1.9 \times 10^{41} - 9.9 \times 10^{45}\,\mathrm{erg\,s^{-1}}$, up to redshift $z = 0.80$. It is the largest sample of high S/N sources used for the PCA-based statistical study of Type 1 AGN spectral properties to date.

i We used PCA as a central tool to determine the source of variance in our data and have mapped a correlation space which is remarkably consistent with previous studies

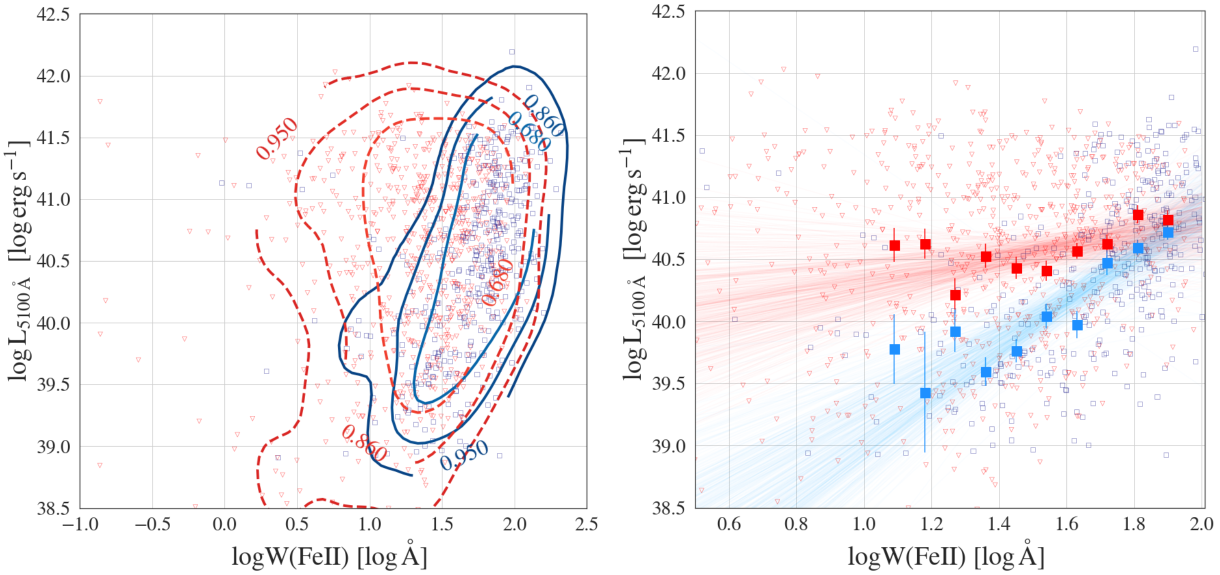


Figure 3.22: *Left panel* : The monochromatic luminosity at 5100 Å as a function of the equivalent width of the FeII-blend at 4570Å. The red triangles and blue squares distinguish red- ($\Delta\lambda > 0.07$) from blue-asymmetric ($\Delta\lambda < -0.07$) Hβ emitters. Kernel density contours have been included. A tighter relation is observed for sources with blue-asymmetric Balmer emission. A similar picture arises in the case of $L_{0.1-2\,\mathrm{keV}}$-W(FeII). *Right panel* : The L_{5100}-W(FeII) plane binned in 10 log W(FeII)-bins between 1.0 and 1.9. For each bin, 10000 bootstrap samples were obtained, for which the mean $L_{5100\,\text{Å}}$ was derived. The points and error bars represent the mean of the sampling distribution of the mean. The error bars correspond to the 1σ error derived from the sampling distribution of the mean using the percentile method. The linear regression was performed with *Linmix* (Kelly 2007).

of the optical Eigenvector 1 (e.g., Boroson & Green 1992; Sulentic et al. 2000a; Grupe 2004; Shen & Ho 2014), while probing a larger cosmological volume.

ii We confirm that the Eddington ratio and the black hole mass are significantly related to the observed diversity of Type 1 AGN through their correlation to EV1 and EV2.

iii Studying Hβ line shapes in this context, we find blue asymmetric emission profiles for the full Type 1 AGN main sequence, while red asymmetries only appear for low accretors. Z10, investigating Hβ profile asymmetries in the 4DE1 context, have suggested that profiles with lower $\mathrm{FWHM_{H\beta}}$ tend to be more symmetric while profiles with larger $\mathrm{FWHM_{H\beta}}$ are preferentially red-asymmetric. The larger number of sources in our sample enabled us to complete this picture: While we do observe a larger portion of sources with red asymmetry index at high $\mathrm{FWHM_{H\beta}}$, lower width profiles appear to cover the full range of asymmetries observed in our sample, i.e., lower $\mathrm{FWHM_{H\beta}}$ do not lead to more symmetric profiles. We can, however, confirm a strong trend between $\mathrm{r_{FeII}}$ and $\Delta\lambda_{\mathrm{H}\beta}$. In their discussion of the physical origins of Hβ profile shapes, Z10 identify VBLR-emission and disk winds as main candidates.

iv A sub-class of our sources does indeed show very shifted, broad components in their Hβ emission, possibly due to the presence of a distinct emitting gas distribution in the inner regions of the BLR. The redshift of this VBLR correlates with Eigenvector 1 and might thus be related to Type 1 AGN diversity.

v Exploring parameter correlations for blue- and redward asymmetric Hβ emitters separately, we observed that FeII line flux correlates differently with source luminosity for red- and blue-ward asymmetric Hβ emitters. We discussed this effect in the light of a flattened, self-shielding BLR, in which the FeII emitting clouds are located, are larger radii than the ones emitting the Balmer lines. The Hβ outflows might in such a configuration may deplete the inner regions of the BLR, exposing the FeII-emitting outer region more directly to the accretion disk continuum emission. However, we find evidence for inter-line contamination between FeII and Hβ, which might play a confounding role in this effect.

Chapter 4

Constraining the AGN X-ray luminosity function at $z \sim 6$

The results presented in this chapter were published in "First constraints on the AGN X-ray luminosity function at z $\sim$ *6 from an eROSITA-detected quasar"* ***Wolf, J.****, Nandra, K., Salvato, M., et al. (2021), Astronomy and Astrophysics, 647, A5.*

4.1 Summary

In this Chapter, we exploit the contiguous area of $\sim 140\,\mathrm{deg}^2$ observed by eROSITA during the calibration and performance verification phase with the purpose of demonstrating the science capabilities of the all-sky survey after 4 years (eRASS:8, $\sim 2.3\,\mathrm{ks}$): eFEDS. We report here the blind detection of a high-redshift X-ray source (eFEDSU J083644.0+005459) that we could identify as the well-known quasar SDSS J083643.85+005453.3 (z=5.81, Fan et al. 2001, henceforth J0836+0054). This quasar was initially discovered by i-band dropout selection in the main SDSS. Taking advantage of the synergy between the eROSITA detection and new radio data from the LOw-Frequency Array (LOFAR; van Haarlem et al. 2013) and the Australian Square Kilometre Array Pathfinder (ASKAP, Johnston et al. 2008, Hotan et al. submitted) Survey With ASKAP of GAMA-09 + X-Ray (SWAG-X, Moss et al. in prep) programme, we investigate the origin of the X-ray emission from J0836+0054 and find further evidence for a confined jetted radio structure. We consequently discuss how the detection of the quasar in a contiguous survey constrains models of the space density of X-ray-emitting AGN.

The optical counterpart determination procedure for eFEDS sources is outlined in Section 4.3. The identified quasar and the extraction and reduction of the eROSITA spectral data are presented in Section 4.4. In Section 4.5, we investigate the multi-wavelength properties of eFEDSU J083644.0+005459/J0836+0054 using ancillary data. The new radio data from LOFAR and ASKAP are presented in Section 4.6. In Section 4.7, we derive constraints on the X-ray AGN space density at high redshift from the detection of J0836+0054. We discuss the origin of the X-ray emission in Section 4.8. After comparing the detection to expected source counts from optical surveys, we conclude by making a prediction for the expected number of $z > 5.7$ quasars which will be found in eRASS:8.

Dr. Duy Hoang documented the LOFAR observation of the source presented in 4.6.1. In addition, Dr. Vanessa Moss provided and discussed the ASKAP observation of J0836+0054 (Section 4.6.2).

In this Chapter a flat ΛCDM cosmology (Planck Collaboration et al. 2020): $\Omega_m = 0.31$, $\Omega_\Lambda = 0.69$ and $H_0 = 68\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Unless stated otherwise, uncertainties are given at the 68% confidence level.

4.2 Introduction: High-redshift AGN space density from X-ray surveys

The detection of quasars at z > 5.5 in the past two decades is an intriguing development because their associated black hole masses challenge our understanding of the formation and initial growth of SMBHs. Testing black hole seed models requires a complete census of high-redshift AGN encoded in well-constrained luminosity functions. In this redshift regime, X-rays in the soft band ($0.2 - 2.3\,\mathrm{keV}$ for eROSITA) probe the restframe hard X-ray emission of the distant sources ($\sim 1.3 - 15\,\mathrm{keV}$ at $z > 5.5$). The soft X-ray selection of AGN, therefore, suffers less from absorption biases than restframe optical/UV selection. So far, however, optical and infrared dropout-selected AGN at $z > 5.5$ (e.g., Fan et al. 2001; Willott et al. 2009; Venemans et al. 2013; Reed et al. 2015; Bañados et al. 2016; Matsuoka et al. 2016; Wang et al. 2017; Bañados et al. 2018) significantly outnumber X-ray selected ones because of the lack of sufficiently wide and deep X-ray surveys backed up by homogeneous ancillary multi-wavelength data. To date, > 400 high-redshift quasars ($z > 5.5$) have been discovered in dedicated optical/near-infrared(NIR) surveys. Chandra and XMM-Newton pointed observations of known quasars in the range $z = 5.7 - 7.54$ have led to the detection of an X-ray signal for only approximately 30 of these objects (e.g. Brandt et al. 2002; Nanni et al. 2017; Vito et al. 2019a; Pons et al. 2020). However, such X-ray follow-up samples suffer from the selection biases of the optical selection because of tight colour–magnitude constraints and absorption.

For the study of the accretion history, absorption biases can be reduced by constructing the X-ray luminosity function (XLF, Hasinger et al. 2005; Ueda et al. 2014; Vito et al. 2014; Miyaji et al. 2015; Aird et al. 2015; Georgakakis et al. 2015; Buchner et al. 2015; Khorunzhev et al. 2018; Ananna et al. 2019) from a purely X-ray-selected sample. However, only three X-ray-selected AGN have been identified at z > 5 so far (in the Chandra Deep Fields, COSMOS and XMM-XXL Barger et al. 2003; Marchesi et al. 2016; Menzel et al. 2016) with the most distant being at z=5.3. These surveys suffer from the small cosmological volume they probe. The full-sky survey currently being carried out with the eROSITA on board the SRG mission (Predehl et al. 2021) will allow us to overcome these limitations and probe the bright end of the XLF at high redshifts with a limiting flux of $\sim 8 \times 10^{-15}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$. For comparison, the second ROSAT all-sky survey catalogue

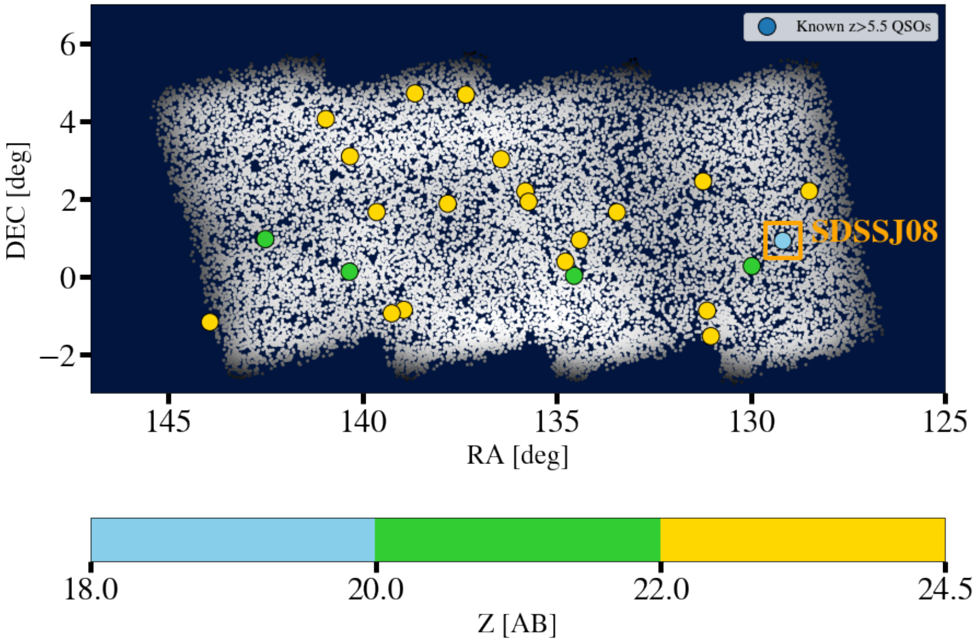


Figure 4.1: Detected point-like sources in eFEDS colour-coded according to their spatial density (brighter is denser). The visible difference in source density is due to the non-uniform exposure of the eFEDS field (Brunner et al. 2022). Known $z > 5.5$ QSOs in the footprint are shown as circles and are colour-coded according to their z-band magnitude. The detected quasar is marked by an orange square.

(2RXS Boller et al. 2016) reached a depth of $\sim 10^{-13}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$. Indeed, in the early months of the first eROSITA all-sky survey, Medvedev et al. (2020, 2021) reported the detection of the X-ray ultra-luminous source SRGE J142952.1+544716, which is matched to the z=6.18 quasar CFHQS J142952+544717 (henceforth J1429+5447).

4.3 Optical counterparts to eFEDS sources

The eFEDS catalogue is presented in Brunner et al. (2022). In Fig. 4.1, the footprint of the survey is presented. The resulting eFEDS source catalogue contains 27910 sources, from which 27369 are classified as point-like.

The field is embedded in the footprint of the DESI Legacy Survey DR8 (LS8) survey, which provides photometry in the g, r, and z bands and in mid-infrared wavebands via forced photometry at the optical positions on WISE images (unWISE data release, Wright et al. 2010; Schlafly et al. 2019). With close-to uniform 5σ depths $g \sim 24.0$, $r \sim 23.4$, and $z \sim 22.5$ (AB magnitudes), LS8 aids the determination of secure counterparts for eFEDS sources to a high level of completeness.

The eFEDS point-like sources were cross-matched to LS8 optical counterparts in a two-method approach, which will be detailed further in Salvato et al. (2022). All LS8 sources within 30″ of an X-ray source are considered potential counterparts. There are an average of approximately 20 LS8 sources within this radius for each X-ray source at the depth of eFEDS.

A careful treatment of astrometry and additional photometric information is needed to associate each source with its correct optical counterpart. The counterpart identification was performed using the Bayesian cross-matching algorithm NWAY[1] (Salvato et al. 2018). In addition to the positional offset and positional uncertainty, it uses a multi-dimensional photometric prior, which was modelled using a Random Forest classifier (see Salvato et al. 2022). The prior was defined using 23058 X-ray sources from the XMM-Newton serendipitous survey (3XMM DR8, Rosen et al. 2016) and the Chandra Source Catalogue (CSC 2.0, Evans et al. 2020) with comparable fluxes to the eFEDS sources and secure counterparts. In parallel, a similar multi-dimensional photometric prior was applied to the classical Likelihood Ratio technique[2] (Sutherland & Saunders 1992). The reconciliation of both approaches delivers a highly reliable set of optical counterparts to the eFEDS sources. Tests on a validation set of simulated eFEDS-like sources indicate that the chosen approaches reach ∼ 96% purity and ∼ 96% completeness (Salvato et al. 2022).

4.4 X-ray properties of J0836+0054

By matching eFEDS LS8 optical counterparts to our compilation of all spectroscopic entries in the field, we were able to determine that the X-ray emission from the eROSITA source eFEDSU J083644.0+005459 is associated with the z=5.81 SDSS quasar SDSS J083643.85+005453.3 (hereafter J0836+0054). The quasar lies at a distance of 6.3" from the eROSITA position and is the brightest optical source within a radius of 30". The LS8 counterpart to the eROSITA source matches J0836+0054 within 0.1″. The eROSITA-eFEDS and Hyper-Suprime Cam (HSC, Aihara et al. 2018) images of the matching region are shown in Fig. 4.2. A summary of the match is presented in Table 4.1. The source belongs to an up-to-date list of 24 spectroscopically confirmed $z > 5.5$ quasars in the eFEDS footprint. which were all discovered in dedicated optical searches (Fan et al. 2001; Venemans et al. 2015; Matsuoka et al. 2018b; Matsuoka et al. 2018, 2019). In Fig. 4.1, we have colour-coded these sources according to their z-band magnitude, which clearly reveals a decrease in space density with increasing optical brightness. J0836+0054 is by far the brightest of all $z > 5.5$ quasars in the field. It has previously been observed in X-rays in a follow-up program of high-redshift Sloan quasars (Brandt et al. 2002). The measured flux is higher than the average soft flux limit of eROSITA in the field ($\sim 8 \times 10^{-15}\,\mathrm{erg\,cm^{-2}s^{-1}}$). This high-redshift quasar is also the only radio-loud one on the list.

In the eFEDS catalogue, eFEDSU J083644.0+005459 has 14.4±5.0 source model counts. These model counts are obtained by fitting the point-spread function (PSF) to a count rate image (the ratio of the spatial count distribution and on-axis exposure time corrected

---

[1] https://github.com/JohannesBuchner/nway

[2] https://github.com/ruizca/astromatch

eFEDS ID	-	eFEDSU J083644.0+005459
$\mathrm{RA}_{\mathrm{eFEDS}}$	[deg]	129.1834 these limit
$\mathrm{DEC}_{\mathrm{eFEDS}}$	[deg]	0.9164
$\sigma_{\mathrm{RADEC,eFEDS}}$	[arcsec]	4.1
$\mathrm{DET}_{\mathrm{LIKE}}$	-	11.00
Counts (0.2-2.3 keV)	-	14.4 ± 5.0
LS8 objID/brickID	-	926/336644
Sep.X/LS8	[arcsec]	6.3
QSO ID	-	SDSS J083643.85+005453.3
QSO Redshift	-	5.81
Sep.QSO/LS8	[arcsec]	<0.1

Table 4.1: Basic source and counterpart information. The coordinates of the eFEDS source are equatorial, with $\sigma_{\mathrm{RADEC,\,eFEDS}}$ being the 1σ X-ray positional uncertainty. The net counts and errors are obtained via photon-mode PSF fitting (Brunner et al. 2022). The Sep. X/LS8 measures the separation between the centroid of the detected X-ray source and the position of the LS8 counterpart. Sep. QSO/LS8 corresponds to the separation between the quasar optical position and the LS8 counterpart.

for vignetting). Background and exposure maps are used in the fitting procedure. The corrected exposure time for the source is `ML_EXP` $= 1179\,s$.

4.4.1 Manual eROSITA spectrum extraction

The eSASS task `srctool` was used to extract source and background spectra, along with instrumental responses. The source coordinates, as well as calibrated event files, are passed to the extraction algorithm, with the background and source-extraction regions determined manually. The background extraction region was defined as an annulus of inner and outer radii $(60'', 198'')$. The source region is delimited by a circle $30''$ in radius centred at the X-ray position. Detected sources in the background region were excluded. The photons are collected over the full eROSITA band (0.2-10 keV). We obtain a total of 20 counts in the spectrum (source and background).

4.4.2 X-ray spectral analysis

A spectral analysis was performed to infer the primary X-ray properties of the quasar (see Liu et al. 2022a, for more details about the spectral analysis of eFEDS sources). We used the analysis software `BXA` (Buchner et al. 2014), which connects the X-ray spectral fitting tool `XSPEC` (v12.11, Arnaud 1996) to the nested sampling algorithm MultiNest (Feroz et al. 2009). The fit was performed in the 0.3 - 8.0 keV energy range. A simple redshifted power-law model only accounting for Galactic absorption was chosen to fit the extracted spectrum: *tbabs*zpowerlw*. In addition, we used a background model, which was trained

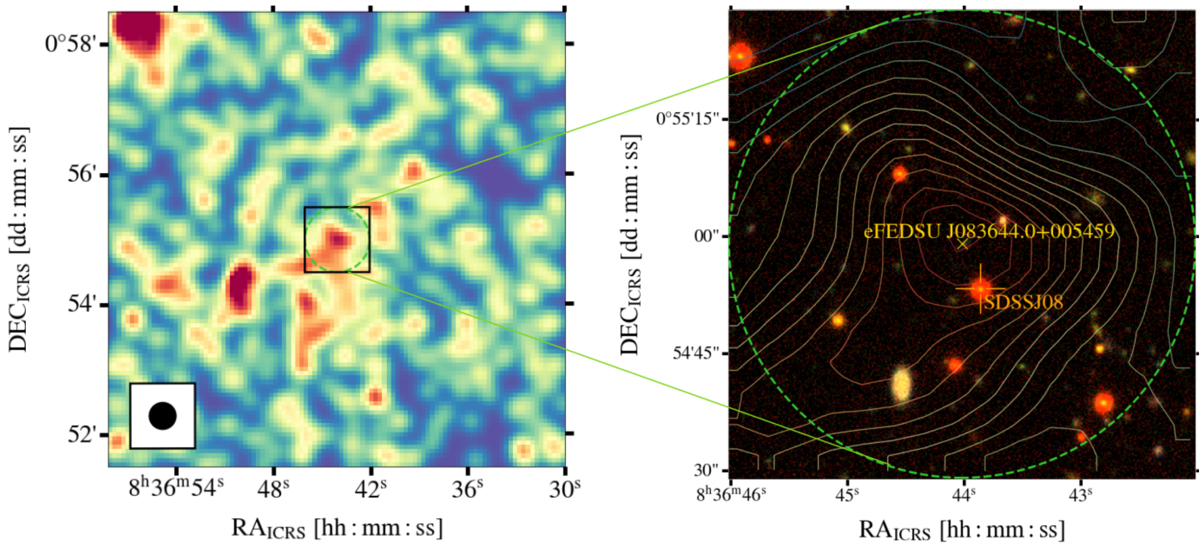


Figure 4.2: (Left) 7′ image of the eROSITA events centred at the optical position of J0836+0054 (full band 0.2-10 keV). The circle has a radius of 30”. The square shows the size of the field shown in Fig. 4.8 (left). (Right) G, I, and Y 60” x 60” HSC (PDR2) image centred at the coordinates of the J0836+0054 associated eFEDS source. The orange cross shows the optical position of the quasar. The eROSITA contours are derived from the eFEDS image smoothed with a Gaussian kernel. The black dot in the lower left corner shows the eROSITA FWHM of the eROSITA PSF in survey mode (12”).

on eFEDS AGN spectra using a principal component analysis and scaled to the source and background extraction sizes following Simmonds et al. (2018, see their appendix A). The corresponding Galactic absorbing column density is taken from HI4PI Collaboration et al. (2016): $N_{\rm H} = 4.8 \times 10^{20}\,{\rm cm}^{-2}$. The power law is shifted to the spectroscopic redshift of the quasar. The low photon counting statistics limit our ability to accurately retrieve X-ray spectral parameters. Nevertheless, we allow the photon index Γ, the normalisation of the power law, and the normalisation of the background model to vary freely in the fit in order to retrieve realistic error bars on the measured X-ray fluxes. The best fit was determined with the C-statistic (Cash 1979). We assumed a flat uniform prior for the photon index, restricting the range to $\Gamma = 1 - 3$. The resulting posterior parameter distributions are shown in Fig. 4.3. We obtain a photon index of $2.20^{+0.49}_{-0.60}$. While Γ is not well constrained, it tends to typical values of X-ray-detected, radio-quiet quasars in this redshift regime. From their joint spectral analysis of X-ray-detected $z > 6$ quasars, Vito et al. (2019a) derived an average photon index of $\Gamma = 2.20^{+0.22}_{-0.20}$, a value consistent with earlier results by Nanni et al. (2017) who performed the same exercise for a $z > 5.7$ quasar sample. The X-ray spectrum of the source is presented in Fig. 4.4.

We compute the soft band flux and intrinsic luminosity from the fitted model. The errors are propagated with `XSPEC` using the posterior samples. Table 4.2 lists the resulting

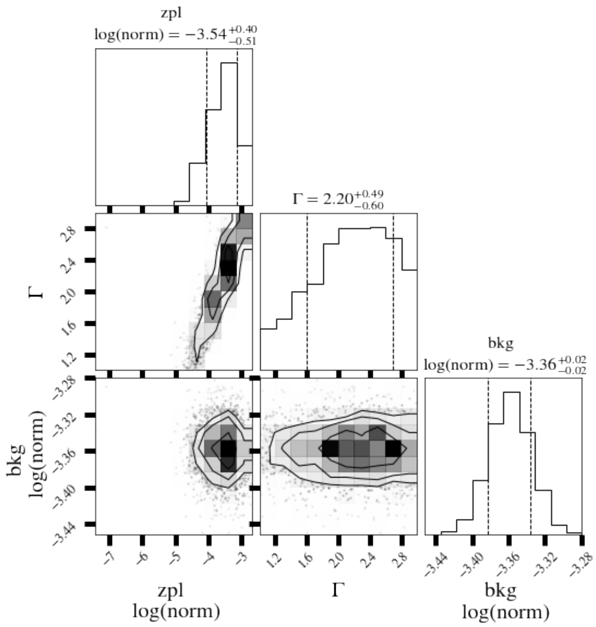


Figure 4.3: Posterior marginal distributions of Γ, normalisation of the power law (zpl), and normalisation of the PCA background model (bkg). The photon index remains poorly constrained but is consistent with typical X-ray spectral slopes of the radio-quiet quasar population.

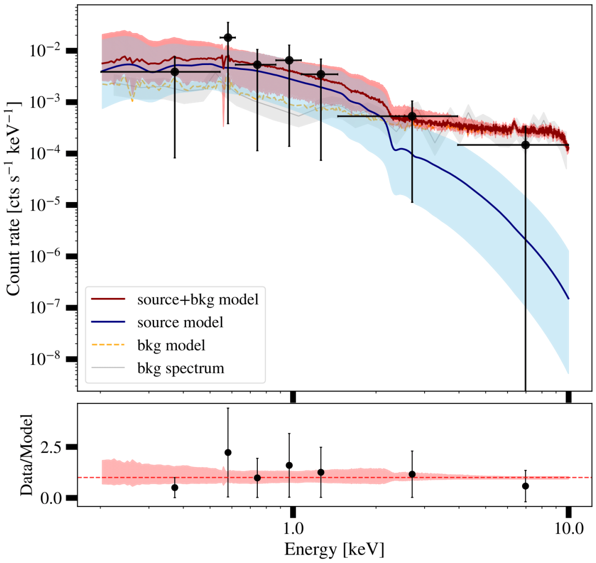


Figure 4.4: X-ray spectrum for eFEDSU J083644.0+005459. The observed count rates are shown in black. The fit was performed in the range 0.3 - 8.0 keV. The fitted source model (blue) and combined source and background model (red) are also presented. The residuals are shown in the lower panel.

X-ray property	Units	Value
$^1F_{0.5-2\,\mathrm{keV}}$	$[10^{-14}\,\mathrm{erg\,cm^{-2}s^{-1}}]$	$9.9^{+3.7}_{-3.2}$
$^2L_{2-10\,\mathrm{keV}}$	$[10^{45}\,\mathrm{erg\,s^{-1}}]$	$4.7^{+2.2}_{-1.6}$
$^3L_\nu$	$[10^{27}\,\mathrm{erg\,s^{-1}\,Hz^{-1}}]$	$7.0^{+6.0}_{-3.9}$
$^4\alpha_{\mathrm{OX}}$	-	$-1.57^{+0.10}_{-0.13}$
$^5\Gamma$	-	$2.20^{+0.49}_{-0.60}$

Table 4.2: Derived X-ray properties of J0836+0054. ^{1}Soft X-ray flux corrected for Galactic absorption. ^{2}X-ray 2-10 keV rest-frame luminosity. ^{3}Monochromatic rest-frame luminosity at 2 keV. ^{4}: The X-ray to optical spectral slope. ^{5}: Photon index.

soft-band flux-corrected for Galactic absorption, the intrinsic $2-10$ keV luminosity, the monochromatic luminosity at 2 keV, the photon index and the two-point spectral X-ray to optical spectral index (see Section 4.4.3).

J0836+0054 has an X-ray detection from a Chandra follow-up observation in 2002 (Brandt et al. 2002). Using a frozen power-law model with $\Gamma = 2$ and Galactic absorption $N_{\mathrm{H}} = 4.4 \times 10^{20}\,\mathrm{cm^{-2}}$ (Stark et al. 1992) these authors obtained a soft-band flux of $F_{0.5-2.0\,\mathrm{keV}} = 1.05 \times 10^{-14}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$. The broad-band Chandra image taken at the optical position of the quasar reveals a single point-like source, which is strong evidence against any contamination in eROSITA from X-ray emission from any other source within a radius of 30”. Re-analysing the Chandra data with a slightly lower photon index ($\Gamma = 1.9$), Nanni et al. (2017) derived the rest-frame intrinsic luminosity $L_{2.0-10.0\mathrm{kev}} = 4.2^{+1.0}_{-1.4} \times 10^{45}\,\mathrm{erg\,s^{-1}}$. With respect to the complete sample of X-ray detected z¿5.7 quasars (Nanni et al. 2017; Vito et al. 2019a; Pons et al. 2020), this makes J0836+0054 one of the four most X-ray luminous high-redshift quasars known to date, together with J1429+5447 (Willott et al. 2010; Medvedev et al. 2020), SDSS J010013.02+280225.8 (Wu et al. 2015; Ai et al. 2016), and PSO J030947.49+271757.31, the blazar discovered by Belladitta et al. (2020). This can be seen in Fig. 4.5, where we present the redshift–luminosity plane for X-ray-detected $z > 5.7$ quasars. The luminosity and flux derived for J0836+0054 are consistent with those reported by Brandt et al. (2002) and Nanni et al. (2017). For a direct comparison with the results of Brandt et al. (2002), we have also computed the unabsorbed flux in the 0.5-2 keV band, fixing $\Gamma = 2$ and using the Galactic absorption quoted in Stark et al. (1992). We obtain $F_{0.5-2\,\mathrm{keV}} = (1.01^{+0.42}_{-0.34}) \times 10^{-14}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$, a value which is consistent with the previous Chandra results. We therefore find no evidence for X-ray variability in J0836+0054 over a timescale of ~ 20 years.

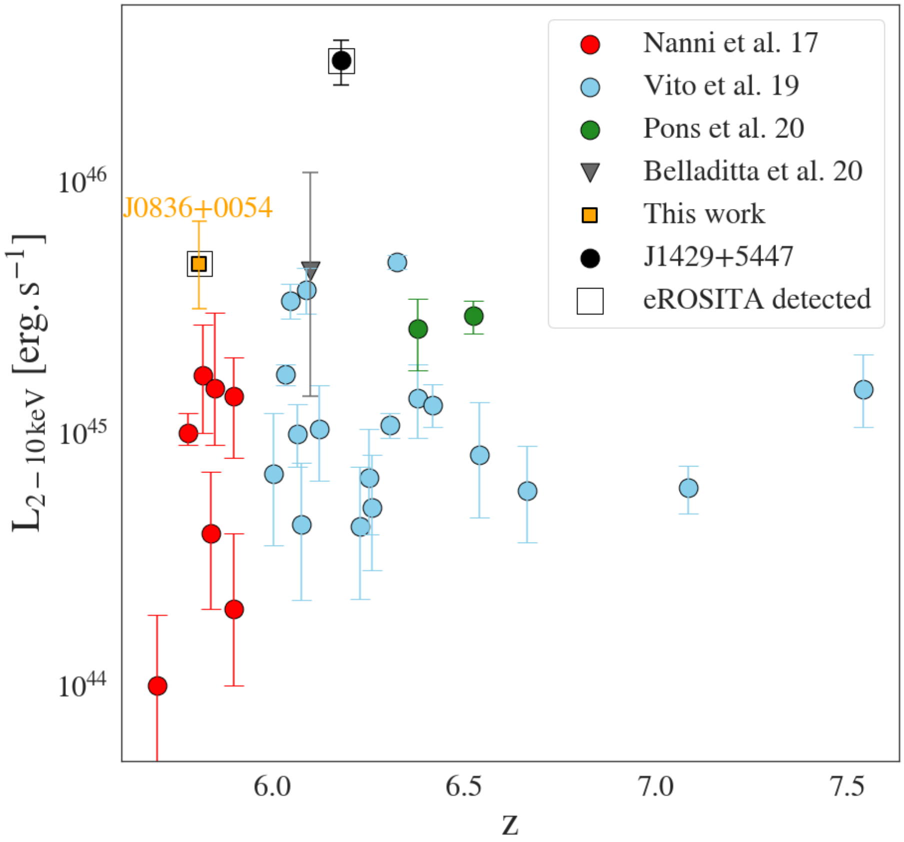


Figure 4.5: Intrinsic hard X-ray luminosity as a function of redshift for X-ray detected $z > 5.7$ quasars. J0836+0054 lies at the X-ray luminous end of the sample. The luminosity of J0836+0054 is computed from the eROSITA data. The quasar lies at the X-ray luminous end of the sample. eROSITA detected sources are marked by an empty square. In the case of overlapping sources in the sample of Nanni et al. (2017) and Vito et al. (2019a), only the data points from the latter are shown.

4.4.3 X-ray loudness

The non-linear relation between X-ray and optical emission of quasars has been studied via the $\alpha_{\rm OX}$ parameter (Tananbaum et al. 1979), the optical to X-ray spectral index. This quantity measures the relative strength of UV continuum and coronal X-ray emission in the active core:

$$\alpha_{\rm OX} = 0.384 \times \log_{10}\left(\frac{L_{2\,\rm keV}}{L_{2500\,\text{Å}}}\right), \tag{4.1}$$

where $L_{2\,\rm keV}$ and $L_{2500\,\text{Å}}$ are the rest-frame monochromatic luminosities at 2 keV and 2500 Å. We computed the 2 keV monochromatic luminosity and its uncertainties from the posterior distribution of the $L_{2-10\,\rm keV}$ rest-frame luminosity and the associated photon index Γ for each solution:

$$L_{2\,\rm keV} = \frac{L_{2-10\,\rm keV}}{\int_{\nu_{2\,\rm keV}}^{\nu_{10\,\rm keV}} \nu^{1-\Gamma}\,d\nu}\nu_{2\,\rm keV}^{1-\Gamma}. \tag{4.2}$$

$L_{2500\,\text{Å}}$ was extrapolated from the UV absolute magnitude $M_{1450\,\text{Å}}$ listed by Jiang et al. (2016). An optical spectral slope of $\alpha = -0.3$ was assumed (e.g. Vito et al. 2019a), corresponding to a correction $M_{2500\,\text{Å}} \approx M_{1450\,\text{Å}} - 0.18$. The value of the X-ray to optical slope for J0836+0054 is given in Table 4.2. For the same quasar, Nanni et al. (2017) measured $\alpha_{\rm OX} = -1.61^{+0.03}_{-0.06}$, consistent with our estimated $-1.57^{+0.10}_{-0.13}$. There is a well-established anti-correlation between $\alpha_{\rm OX}$ and $L_{2500\,\text{Å}}$ for $z < 5$ AGN, also investigated directly in the $L_{2\,\rm keV} - L_{2500\,\text{Å}}$ plane (e.g. Avni & Tananbaum 1986; Strateva et al. 2005; Just et al. 2007; Lusso et al. 2010). Performing a linear regression in the $\alpha_{\rm OX} - L_{2500\,\text{Å}}$ plane, Nanni et al. (2017) argue that their sample of 29 X-ray-detected $z > 5.7$ quasars followed the anti-correlation well. From their sample of $z > 6$ radio-quiet quasars, Vito et al. (2019a) find no significant evolution of this trend with redshift.

The eFEDS quasar shows a typical 'X-ray loudness' with respect to the $\alpha_{\rm OX}$-luminosity relation. For J1429+5447, Medvedev et al. (2020) reported a significant deviation from the relation, arguing that this flattening of the spectral slope could be caused by an excess X-ray emission, possibly related to the inverse Compton scattering of cosmic microwave background photons off the jet (iC-CMB, Tavecchio et al. 2000; Celotti et al. 2001). Despite being radio-loud (Bañados et al. 2015), J0836+0054 does not display such an X-ray excess. We show how the two eROSITA-detected $z > 5.7$ quasars are distributed with respect to this relation in Fig. 4.6. The $\alpha_{OX} - L_{2500\text{Å}}$ relation derived by Lusso et al. (2010) on a sample of *XMM*-COSMOS AGN is shown, as is the one obtained by Nanni et al. (2017) on their sample of $z > 5.7$ quasars.

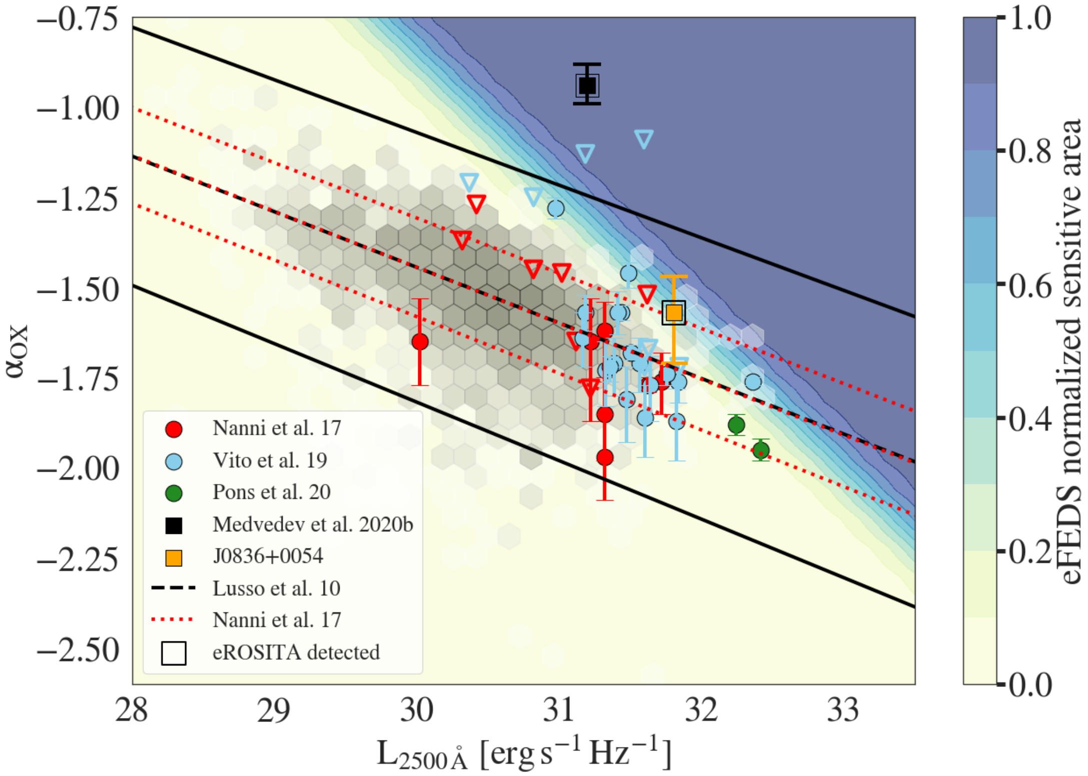


Figure 4.6: X-ray-to-optical slope anti-correlates with the UV monochromatic luminosity at 2500 Å. J0836+0054 (orange circle) is consistent with the best-fitting relation of Lusso et al. (2010, dashed line) and Nanni et al. (2017, red dotted line). The $z > 5.7$ X-ray-detected sources from Nanni et al. (2017, restricted to $z < 6$ and ignoring J0836+0054), Vito et al. (2019a) and Pons et al. (2020) are shown. The empty triangles denote upper limits from undetected sources from Nanni et al. (2017); Vito et al. (2019a) and Pons et al. (2020) (same colour code as detected sources). The grey density scale shows a sample of 2685 XMM-Newton-detected $z < 5$ SDSS quasars (Lusso & Risaliti 2016). The contours show the eFEDS normalised, sensitive area derived from synthetic power-law spectra at $z = 6$ with $\Gamma = 2$ and Galactic absorption. We note that J1429+5447 (black point) tends towards flatter α_{OX} values than X-ray-detected quasars at similar UV luminosities. Its α_{OX} was computed from the luminosity reported in Medvedev et al. (2021).

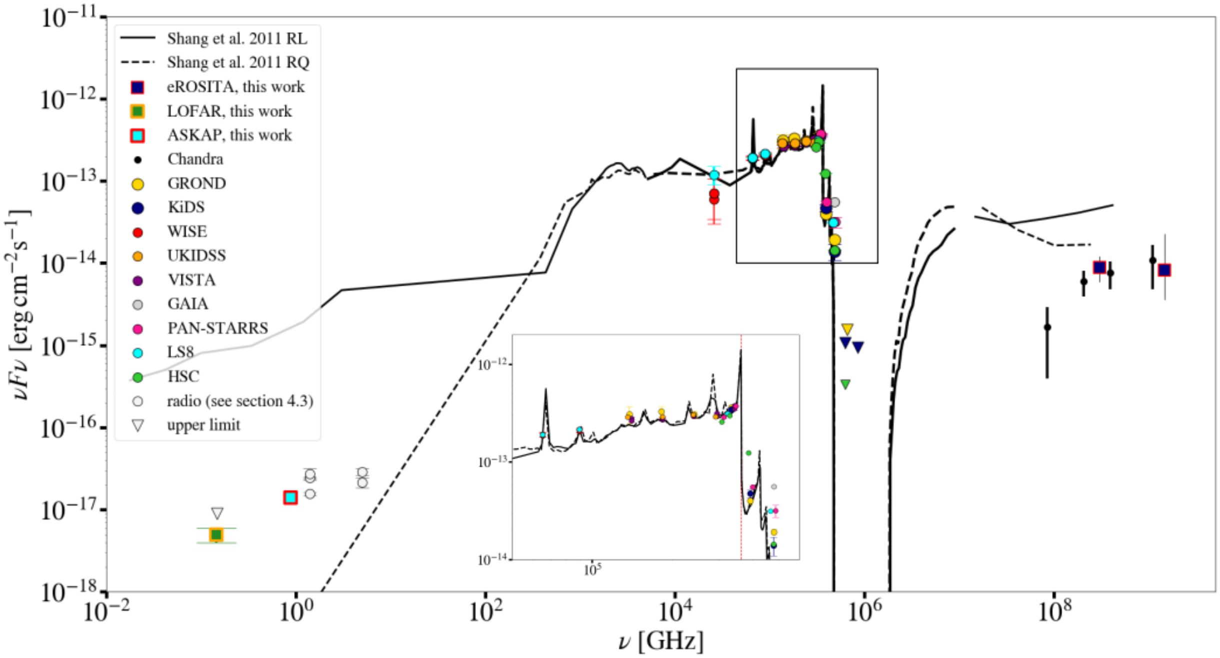


Figure 4.7: SED of J0836+0054 constructed using archival multi-wavelength data, together with the new eROSITA (observed broad band fluxes in 0.5-2 keV and 2-10 keV), LOFAR 145 MHz, and ASKAP 888 MHz measurements. The markers show photometric points from various optical, X-ray, and radio surveys. Triangles denote upper limits. Composite SEDs from Shang et al. (2011) for radio-loud and radio-quiet quasars are fitted to the SED at $z = 5.81$ and corrected for absorption by the intergalactic medium (Madau & Rees 2000). J0836+0054 does not present an X-ray excess typically observed in radio-loud quasars.

4.5 Archival multi-wavelength properties

4.5.1 Optical selection and spectroscopy

J0836+0054 was initially discovered through i-dropout selection and consecutive spectroscopic confirmation in $\sim 1550\,\mathrm{deg}^2$ of the SDSS main survey (Fan et al. 2001). It is part of a sample of 52 $z > 5.4$ quasars that were found by exploiting imaging data in the SDSS main survey (Jiang et al. 2016). The $i - z > 2.2$ dropout criterion selects $z > 5.8$ quasars because the neutral hydrogen absorption bluewards of the Lyα line is shifted in the i-band. A redshift of z = 5.81 was measured by Kurk et al. (2009) using VLT/ISAAC NIR observations. Fan et al. (2001) report a strong and broad Lyα and $\mathrm{N_V}$ complex with an equivalent width of $\sim 70\,\text{Å}$. The quasar is extremely luminous with an absolute AB magnitude $M_{1450\text{Å}} = -27.86$ (Jiang et al. 2016). Its black hole mass of $(2.7 \pm 0.6) \times 10^9\,\mathrm{M}_\odot$, was derived from the width of the broad MgII λ2800Å line (Kurk et al. 2009). Stern et al. (2003) observed J0836+0054 with the FLAMINGOS multi-object, NIR spectrograph at the 8 m Gemini-South Observatory and reported an optical power-law index of $\alpha = 1.55$ (measured over the rest-frame wavelength range $\lambda = 1480 - 2510\,\text{Å}$). This red spectral slope is indicative of the presence of substantial amounts of dust in the environment of J0836+0054.

4.5.2 Spectral energy distribution

J0836+0054 has been covered by a number of imaging surveys over the entire spectral energy distribution (SED)[3]. In addition, we carried out simultaneous g′, r′, i′, z′, J, H, and K_s-band photometry of J0836+0054 with the Gamma-Ray Burst Optical/Near-Infrared Detector (GROND, Greiner et al. 2008) at the MPG 2.2 m telescope at the ESO La Silla observatory. The resulting SED is presented in Fig. 4.7. We fitted the composites of radio-loud and radio-quiet quasars by Shang et al. (2011) with the photometric code Le Phare (Arnouts et al. 1999; Ilbert et al. 2006). This figure can be directly compared to the SED of J1429+5447 presented by Medvedev et al. (2020). For the fit, the redshift was fixed to its spectroscopic value. We note that J0836+0054 does not display the X-ray excess typically observed in radio-loud quasars (Wilkes & Elvis 1987; Shastri et al. 1993; Reeves et al. 1997) and is more consistent with the radio-quiet template in the radio and X-ray part of the SED. The overall radio output is well below that of the radio-loud template. We summarise all available archival radio data for J0836+0054 in the following section.

4.5.3 Archival radio properties

All the radio flux measurements and upper limits associated with J0836+0054 are reported in Table 4.3. Coppejans et al. (2017) ascribe the tension between some of the 1.4 GHz flux-density measurements to resolution effects (e.g. NVSS has a resolution of 45″) but could not entirely exclude variability. Assembling all available radio data on this quasar, these latter authors estimated a radio spectral slope of $\alpha_r = -0.89 \pm 0.29$. In this Chapter, sources with $\alpha < -0.5$ were classified as steep-spectrum sources. Coppejans et al. (2017) further note that their computed radio spectrum slope for J0836+0054 would predict a flux density of $\sim$ 12.0 mJy at 148 MHz, but Coppejans et al. (2016) argue that most steep-spectrum high-redshift quasars must have turnovers in their synchrotron spectra effectively making them MHz-peaked or GHz-peaked sources. The Giant Metrewave Radio Telescope (GMRT) 150 MHz All-sky Radio Survey (TGSS-ADR1; Intema et al. 2017) does not detect J0836+0054. Using the TGSS-ADR1 data, Coppejans et al. (2017) derived an upper limit of 6.1 mJy on the flux density of the source. If variability and resolution effects can be excluded, these latter authors conclude that this lower flux density could indicate a

---

[3] Gaia (Gaia Collaboration et al. 2018), Pan-STARRS1 (Chambers et al. 2016), HSC SSP, LS8. It has also been detected in NIR and mid-infrared (MIR) surveys: United Kingdom Infrared Telescope (UKIRT) Infrared Deep Sky Survey (UKIDSS, Lawrence et al. 2007), Visible and Infrared Survey Telescope (VISTA) Kilo-degree Infrared Galaxy Survey (VIKING Kuijken et al. 2019), and WISE (AllWISE data release, Wright et al. 2010; Cutri et al. 2013)

Telescope	Frequency [GHz]	Flux density [mJy]	Survey/Ref.
GMRT	0.150	< 6.1 (3σ)	[1]TGSS-ADR1
VLA	1.4	1.11 ± 0.15	[2]FIRST
VLA	1.4	2.1	[3]NVSS, Petric et al. (2003)
VLA	1.4	1.75 ± 0.04	Petric et al. (2003)
VLA	1.4	1.96 ± 0.31	Frey et al. (2005)
EVN	1.6	1.1	Frey et al. (2003)
VLA	5	0.580 ± 0.057	Petric et al. (2003)
VLA	5	0.43 ± 0.06	Frey et al. (2005)
EVN	5	0.34	Frey et al. (2005)
MAMBO	250	< 2.9 (3σ)	Petric et al. (2003)

Table 4.3: Archival radio observations of J0836+0054. [1]:Intema et al. (2017), Coppejans et al. (2017); [2]: Becker et al. (1995),White et al. (1997); [3]: Due to the low resolution of NVSS ($45''$), the reported flux density of 2.5 ± 0.5 in Condon et al. (1998) is contaminated by a source located $10''$ to the south of J0836+0054. We subtracted the contribution of the contaminated source (0.44 mJy; Petric et al. 2003) from the reported value.

spectral turnover below 150 MHz in the observed frame (~ 1 GHz rest-frame). Using high-resolution imaging of J0836+0054 from the European Very Long Baseline Interferometry (VLBI) Network at 5 GHz and simultaneous VLA observation, Frey et al. (2005) were able to demonstrate the compactness of the source down to an accuracy of 2 mas, thereby showing that the radio emission is concentrated within the central 40 pc of the AGN.

4.6 Confirmation of a sub-GHz spectral flattening with LOFAR and ASKAP

Within 6 months of the eROSITA observations, the entire eFEDS field was observed with LOFAR 145 MHz and ASKAP 888 MHz with dedicated programs. In the sections below, we describe the observations and report the detection of J0836+0054 in the sub-GHz radio bands.

4.6.1 LOFAR 145 MHz observations

The eFEDS field was observed with LOFAR in the 120–168 frequency band (Project: LC13_029) between February and May of 2020. The LOFAR data were processed with the data reduction pipelines (`PREFACTOR`[4]; van Weeren et al. 2016; Williams et al. 2016;

---

[4] `https://github.com/lofar-astron/prefactor`

de Gasperin et al. 2019 and `ddf-pipeline`[5]; (Tasse 2014a,b; Smirnov & Tasse 2015; Tasse et al. 2018) that were developed by the LOFAR Surveys Key Science Projects (Shimwell et al. 2017, 2019). Direction-independent (bandpass, instrumental delays) and direction-dependent (ionospheric, beam) effects were corrected for in the calibration. The flux scale of the LOFAR final image in the region of interest (i.e. 1.5 degree square region centred on J0836+0054) was found to be consistent within ~ 15 percent with the flux scale of the TGSS-ADR1 150 MHz data (Intema et al. 2017). Here, we use a conservative uncertainty of 20 percent for the flux scale of the LOFAR data. For details of the data reduction, we refer to (Ghirardini et al. 2021) and (Hoang et al. 2022).

In Fig. 4.8 we report the detection of J0836+0054 with LOFAR at 145 MHz. The emission peak is detected up to 14σ, where $\sigma = 200\,\mu$Jy/beam. The flux density of J0836+0054 at 145 MHz is 3.35 ± 0.7 mJy. In addition, our LOFAR image confirms the presence of a second fainter radio source with a flux density of 1.62 ± 0.32 mJy to the south of J0836+0054 (i.e. with an angular separation of $\sim 10''$). This source was found with high-frequency (1.4 and 5 GHz) observations and is associated to a lower redshift galaxy (Petric et al. 2003; Frey et al. 2005). Combining our measurement with the 1.4 GHz flux density (0.44 mJy) reported in Petric et al. (2003), we find that the spectrum of the source has a spectral slope of -0.57.

4.6.2 ASKAP SWAG-X 888 MHz observations

eFEDS was also observed by the ASKAP telescope as part of the SWAG-X Observatory Project (Moss et al. in prep) almost simultaneously with the eROSITA observations. In this Chapter, we characterise J0836+0054 based on the continuum-only SWAG-X data observed at 888 MHz in October 2019. This dataset comprises six ASKAP tiles for complete coverage of the eFEDS region, with 8 hr integration per tile. Each tile was processed using ASKAPsoft with standard continuum settings (Guzman et al. 2019, Whiting et al., in prep), including bandpass calibration, flagging, and self-calibration. Imaging was carried out using multi-scale, multi-frequency synthesis, resulting in average sensitivities across the full field of $\sim 50\,\mu\mathrm{Jy\,beam}^{-1}$. Sources and fluxes were extracted using Selavy (Whiting & Humphreys 2012). The resolution of the image containing J0836+0054 is 13.3 × 12.1 arcsec, and at this resolution the two components seen with LOFAR are confused. The source extraction performed as part of the pipeline processing fitted a single extended Gaussian component. A subsequent fit was performed that forced the size of the Gaussian components to be that of the PSF, and two components were fitted. These are spatially coincident with the components seen in LOFAR, and have fluxes of 1.575 ± 0.008 mJy (J0836+0054) and 0.926 ± 0.027 mJy.

---

[5] `https://github.com/mhardcastle/ddf-pipeline`

4.6.3 Low-frequency spectral flattening

In Fig. 4.8 we present the radio spectrum of J0836+0054 combining the measurements from the new ASKAP and LOFAR data presented here, as well as ancillary radio data. The spectral slope becomes slightly flatter at lower frequencies, being -0.30 ± 0.13 from 145 MHz to 1.4 GHz, compared to -1.02 ± 0.16 from 1.4 GHz to 5 GHz. The SED confirms the spectral flattening at sub-GHz frequencies and hints at the peaked nature of the spectrum. However, no conclusion as to whether or not there is a spectral turnover at sub-GHz frequencies can be drawn from these data.

An alternative picture arises when considering the discrepancy between the high signal-to-noise ratio 1.4 GHz measurements of Petric et al. (2003) and Frey et al. (2005) with the reported 888 MHz ASKAP flux density, which may be indicative of flux variability at $\sim 1\,\mathrm{GHz}$. Indeed, spectral variability has been observed for most steep spectrum sources (e.g. Orienti et al. 2007; Mingaliev et al. 2012). The expansion of a young radio source can produce spectral variations that are observable over the course of a decade (for a review see Section 2.2. in Orienti 2016). Such variability behaviour was observed by Orienti & Dallacasa (2008) in the GHz-peaked source RXJ1459 + 3337. By compiling VLA observations over a period of 17 years, these latter authors report a shift of the turnover frequency towards lower frequencies, which they show to be well-explained by the adiabatic expansion of a homogeneous component. The expansion of this component results in a decrease of the source opacity, which effectively shifts the turnover frequency. If we consider the contemporaneous ASKAP 888 MHz and LOFAR 145 MHz flux measurements (2019) separately from the Frey et al. (2005) VLA 1.4 GHz and 5 GHz observations (2003) for example, we obtain spectral slopes: $\alpha_{145-888\,\mathrm{MHz}} = -0.42$ and $\alpha_{1.4-5\,\mathrm{GHz}} = -1.02$. This significant flattening observed between two observations separated by roughly 16 years could also be explained in the adiabatic expansion scenario. However, to answer questions relative to the turnover frequency, future observations at low frequencies, for example with LOFAR Low Band Antennas operating at 10–80 MHz, will be necessary. It is worth mentioning that the hypothesis of variability due to adiabatic expansion could also already be tested with contemporaneous flux measurements at 1.4 GHz.

4.7 AGN space density at $z \sim 6$

The X-ray-selected AGN with the highest and second-highest redshift to date are J1429+5447 and J0836+0054 detected by eROSITA. By ‘X-ray-selected’, we mean the blind, serendipitous detection of a source in a contiguous survey (in contrast to dedicated followup or count extraction at known quasar coordinates). The extremely low expected space density of luminous X-ray sources at high redshifts requires wide surveys to reach within the epoch of re-ionisation. In the Chandra Deep Fields, COSMOS and XMM-XXL, only three X-ray-selected quasars have been identified at z > 5 (Barger et al. 2003; Marchesi et al. 2016; Menzel et al. 2016), the highest redshift being $z = 5.3$ (Capak et al. 2011). The absence of $z > 5.5$ X-ray-selected quasars can arise from technical difficulties such as the definition

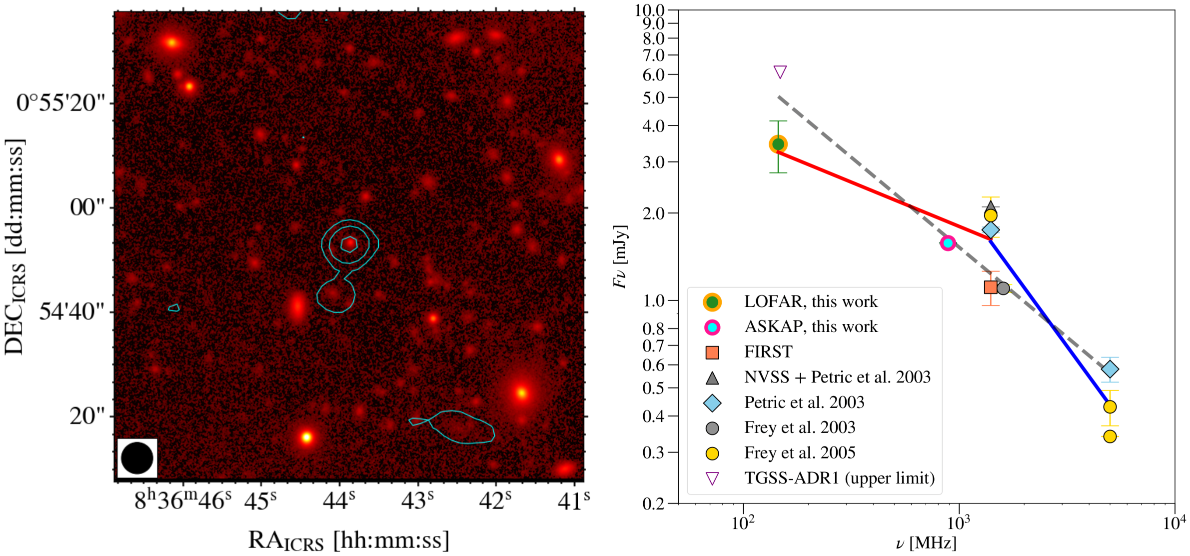


Figure 4.8: *Left*: LOFAR 145 MHz contours on top of the HSC z-band image (right). The contour levels are $[1, 2, 4] \times 3\sigma$, where $\sigma = 200\,\mu$Jy/beam. The beam size of $6'' \times 6''$ is shown in the bottom left corner. The field corresponds to the square in Fig. *Right*: Radio spectrum of J0836+0054. The coloured markers present radio measurements from the literature. The empty triangle shows the TGSS-ADR1 upper limit which is not used for spectral fitting. At frequencies above $\sim$ 1 GHz, the spectrum appears steep. The LOFAR 145 MHz (green and orange point) and the ASKAP 888 MHz (cyan and red point) flux densities reported in this Chapter confirm a flattening of the spectrum at frequencies below $\sim$ 1 GHz. The red and blue lines show the best-fitting spectra with the indices of -0.30 ± 0.13 and -1.02 ± 0.16 in the frequency ranges below and above 1.4 GHz, respectively. The best-fitting line for all data points with an index of -0.62 ± 0.12 is shown with the grey dashed line. 4.2.

of appropriate source-extraction parameters, i.e., adapted to the detection of faint sources (such as detection likelihood thresholds that are too conservative for low-count sources, too soft detection bands). However, a dominant factor is the probed cosmological volume which must be large enough to sample the low density and luminous sources beyond the knee of the XLF at higher redshifts.

The evolution of AGN selected in X-ray surveys has been extensively studied up to $z = 5$ (e.g. Hasinger et al. 2005; Ueda et al. 2014; Miyaji et al. 2015; Buchner et al. 2015; Georgakakis et al. 2015; Aird et al. 2015; Ananna et al. 2019). Vito et al. (2018) investigated the AGN space density in the CDF-N and CDF-S up to $z = 6$, focusing on the $\log\left(L_{2-10\,\mathrm{keV}}/(\mathrm{erg/s})\right) < 44$ regime, and particularly addressing the question of the evolution of the slope at the low-luminosity end. The highest spectroscopically confirmed redshift in their sample is $z = 5.186$ (Vignali et al. 2002). The detection of eFEDSU J083644.0+005459 in a contiguous and flux-limited survey of near-homogeneous exposure allows us to impose constraints on the space density of more luminous X-ray-selected AGN at $z \sim 6$ based on secure spectroscopic data.

In the following section, we derive a binned estimate of the XLF in the range $5.7 < z < 6.4$. In a complementary step, we compute number estimates from extrapolated fits to the XLF from the literature and verify whether they are consistent with the detection of eFEDSU J083644.0+005459.

4.7.1 Binned estimate of the XLF

We have defined the binned estimate of the XLF in Eq. 1.8 (Page & Carrera 2000). The statistical uncertainty on ϕ_{est} is given by the Poisson error on N (e.g. Gehrels 1986) normalised by the comoving sensitive volume (Eq. 1.9). We compute ϕ_{est} in the bin $\Delta z = 5.7 - 6.4$ and $\Delta\log\left(L_{\mathrm{X}}/(\mathrm{erg/s})\right) = 45.5 - 46$, which contains J0836+0054. The chosen redshift range corresponds to the selection function of bright SDSS quasars (Jiang et al. 2016) and arises from colour-selection criteria. The resulting binned XLF point is shown in Fig. 4.10. The sensitive area $A(\log L_{\mathrm{X}}, z)$ is derived from the eFEDS sensitivity curve which expresses the sensitive area as a function of counts in the range $0.2-2.3\,\mathrm{keV}$. It was generated with the eSASS task `APETOOL` (for more details on the sensitivity determination please see Georgakakis et al. (2008)). In Eq. 1.8, we integrate over redshifts and luminosity. In order to convert a given redshift–luminosity in soft band counts, we simulated X-ray spectra with `XSPEC` using a model *clumin*tbabs*zpowerlw*. This model corresponds to a redshifted power law with Galactic absorption. We have frozen the photon-index to $\Gamma = 2$, a value which is consistent with results from spectral analysis carried out at $z > 5.7$ Nanni et al. (2017); Vito et al. (2019a). The Galactic column density was fixed to the values: $3 \times 10^{20}\mathrm{cm}^{-2}$. The convolution model *clumin* was used to fix the values of redshift and luminosity for a given $z - L_{\mathrm{X}}$ configuration. ARF and RMF files for a typical eFEDS source were used[6]. We have generated a grid of spectra over the ranges $z = 4-8$ and $\log\left(L_{2-10\,\mathrm{keV}}/(\mathrm{erg/s})\right) = 44-47$. For each spectrum, the count-rate in the range $0.2-2.3\,\mathrm{keV}$, was computed and the normalised area sensitivity was evaluated with the `APETOOL` sensitivity curve. This sensitivity grid is shown in Fig. 1. The synthetic $(L_{2-10\,\mathrm{keV}}, z, A(L_{2-10\,\mathrm{keV}}, z))$ data display a sharp break and are distributed as a multivariate sigmoid. We fitted $A(L_{2-10\mathrm{keV}}, \log(1+z))$ using logistic regression. We note that we fit in $\log(1+z)$ and not z in order to capture the evolution of the flux limit with $\sim L/(1+z)^{-4}$. The regression was implemented with a generalised linear model (GLM, e.g.Nelder & Wedderburn (1972)) using the `statsmodels` Python library Seabold & Perktold (2010). A GLM is a regression model for which the probability density function of the outcome variable can be specified. In the case of $A(L_{2-10\,\mathrm{keV}}, \log(1+z))$, we select the binomial distribution with *logit*[7] as link function. In a GLM the link function specifies the relation of a variable to the linear model. The resulting fitted surface is shown in Fig. 4.9 This function can be used in the integral of Eq. 1.8. The statistical uncertainty on ϕ_{est} is given by the Poisson error on N e.g. Gehrels (1986) normalised by the comoving sensitive

---

[6] The actual eROSITA response matrices for the source were not available when the analysis was performed.

[7] This function is used in statistics to express the natural logarithm of odds corresponding to a probability p. It is given by $logit(p) = \ln(p/(1-p))$.

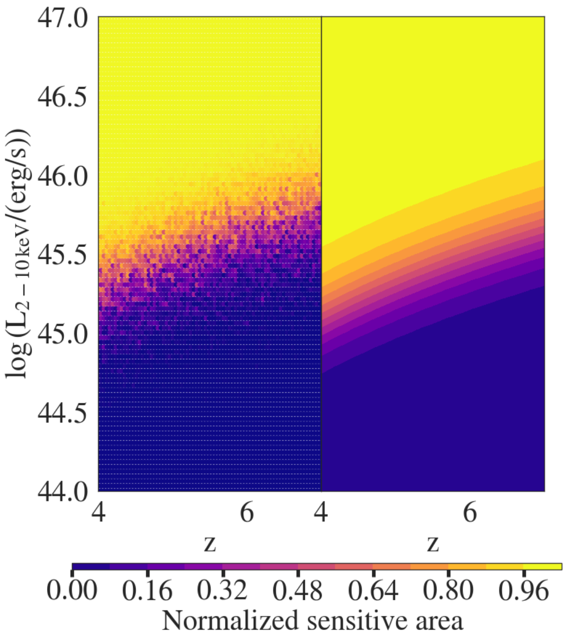


Figure 4.9: *Left*: Sensitive area as a function of luminosity and redshift from synthetic spectra. *Right*: Fit to the $(L_{2-10\,\mathrm{keV}}, z, A(L_{2-10\,\mathrm{keV}}, \log(1+z)))$ surface (used for efficient integration).

volume. We compute ϕ_{est} in the bin $\Delta z = 5.7 - 6.4$ and $\Delta \log\left(L_\mathrm{X}/(\mathrm{erg/s})\right) = 45.5 - 46$, which contains J0836+0054. The chosen redshift range corresponds to the selection function of bright SDSS quasars Jiang et al. (2016) and arises from colour-selection criteria. The resulting binned XLF point is shown in Fig. 4.10 (left).

The XLF can be parametrised as a double power law:

$$\phi_m = \frac{K}{(L_\mathrm{X}/L_*)^{\gamma_1} + (L_\mathrm{X}/L_*)^{\gamma_2}}, \tag{4.3}$$

where K is the normalisation, γ_1 and γ_2 the slopes of the power-law components, and L_* the break luminosity. This double power law is modified by a redshift evolution-term which can either be applied to the normalisation or the break luminosity (or both).

We extrapolated fitted parametric models of the XLF from various authors to redshift $z = 6.05$ (i.e. the central redshift of Δz), in order to compare ϕ_{est} with the expected number density from studies at lower redshift. For the comparison with our binned estimate, we have chosen to extrapolate XLFs which were measured up to $z < 5$ by Ueda et al. (2014, U14), Vito et al. (2014, V14), Buchner et al. (2015, B15), Miyaji et al. (2015, M15), Aird et al. (2015, A15), and Georgakakis et al. (2015, G15). In these works, the XLF was measured from samples of soft-X-ray-selected AGN found in Chandra[8], XMM-Newton[9],

---

[8] CDF-S, CDF-N, AEGIS, ECDF-S,C-COSMOS

[9] XMM-XXL, XMM-COSMOS, SXDS

SWIFT/BAT, MAXI, ASCA, and ROSAT data. The best-fitting model from each of these works is retained here.

V14 and G15 reported that a pure density evolution (PDE) best fits their data. This model assumes an evolutionary term, parametrised as a multiplicative factor to the normalisation K. U14 assumed a luminosity-dependent density evolution (LDDE). Based on previous observations of a decline in the comoving number of $\log\left(L_{2-10\,\mathrm{keV}}/(\mathrm{erg/s})\right) > 44$ AGN at higher redshifts (Brusa et al. 2009; Civano et al. 2011; Hiroi et al. 2012), two cut-off redshifts were introduced in the parametrisation of the evolutionary term. M15 also used LDDE, accounting for the probability distribution of photometric redshifts in addition to absorption effects. A15 introduced a flexible double power law (FDPL) as parametrisation of the XLF. This allows any parameter in Eq. 4.3 to evolve with redshift. This redshift dependence is modelled by polynomials of $\log(1+z)$. The extrapolations are presented in Fig. 4.10. For U14, V14, M15, and G15, uncertainties are derived by sampling from the 1σ confidence intervals of the parameters entering the models fitted by these authors. We note that we are not accounting for the correlation between the parameters, which may result in an over-estimation of the computed uncertainties. In B15, the XLF is measured in a non-parametric approach based on the Bayesian spectral analysis of individual sources and a smoothness prior connecting bins of the XLF. The associated interval presented in Fig. 4.10 is obtained from a tabulated version of the non-parametric XLF, allowing intrinsic absorption to vary over $\log N_{\mathrm{H}} = 21 - 26$. This function was initially derived in the $z = 4 - 7$ range within which it is constant. We have re-scaled it to the cosmological volume in the range $z = 5.7 - 6.4$. The hard luminosity cut-off displayed in Fig. 4.10 is due to the luminosity range on which B15 evaluated the XLF. The model by A15 shown in Fig. 4.10 corresponds to the best fit to the unobscured sample ($20 < \log N_{\mathrm{H}} < 22$) from this work.

The error bars of ϕ_{est} based on the detection of J0836+0054 appear only marginally consistent with the extrapolations of the functions of V14, U14, M15, and G15, while allowing noticeably higher space densities than all model extrapolations. The function by B15 shows the best agreement with our data. ϕ_{est} is furthermore not corrected for redshift-completeness. The 30% spectroscopic completeness of eFEDS makes this single detection a lower limit on the number of bright high-redshift quasars in the field. The 1σ Poisson uncertainty on ϕ_{est} indicates that expected number counts for $z \sim 6$ X-ray-selected AGN based on the integration of current XLF models (for eROSITA expected counts, see Kolodzig et al. 2013) may in fact represent quite a conservative estimate of the true number of luminous, high-redshift AGN that eROSITA will detect.

4.7.2 Comparison to eFEDS expected number counts

We can compare the single detection in eFEDS to the number of counts expected from extrapolated parametric XLFs (for the same type of object, in the same field). The expected number of detected AGN in a survey for an XLF model ϕ_m with parameters θ can be written as:

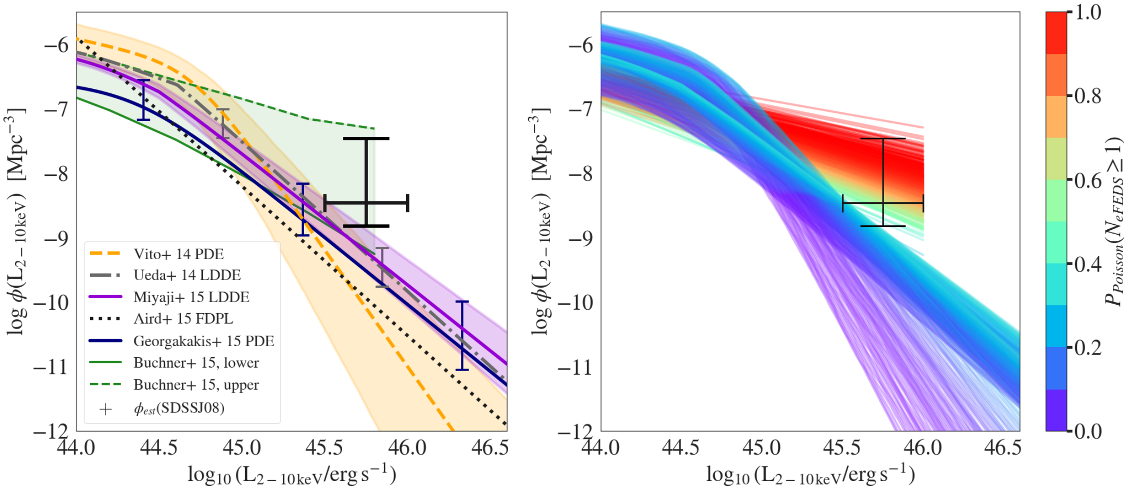


Figure 4.10: *Left*: Parametric models of the XLF extrapolated to $z = 6.05$. The coloured error bars and shaded areas correspond to 1σ uncertainties on the nominal models of the same colour. The nonparametric XLF from B15 re-scaled to the comoving volume in the range $z = 5.7 - 6.4$ is also shown in green. The binned estimate of the AGN space density as derived from the detection of J0836+0054 is shown by the black error-bars. *Right*: Samples drawn from the 1σ uncertainties on the XLFs from V14, U14, M15, G15, and B15, colour-coded according to their Poisson probability of supporting at least one detection in eFEDS. Flatter slopes of the XLF are preferred.

XLF	$\bar{N}^{nominal}_{\mathrm{eFEDS}}$	method	$N^{50-th}_{\mathrm{eFEDS}}$	$N^{84.1-th}_{\mathrm{eFEDS}}$
V14	0.0	sampled	0	0
U14	0.2	sampled	0	1
M15	0.4	sampled	0	1
G15	0.2	sampled	0	1
B15 lower	0.2	nominal	0	1
B15 upper	15.9	nominal	15	19
A15	0.0	nominal	0	0

Table 4.4: High-redshift source count predictions. $\bar{N}^{nominal}_{\mathrm{eFEDS}}$ are expected source counts in eFEDS obtained from the integration of various nominal best-fitting XLF models. We also list the 50th and 84.1th percentiles of the distributions of expected Poisson counts in eFEDS (accounting for the XLF fit uncertainties). The investigated intervals are $z = 5.7 - 6.4$ and $\log{(L_{2-10\,\mathrm{keV}}/(\mathrm{erg/s}))} = 45.5 - 50$. The *method* indicates whether the 1σ uncertainties on the XLF parameters were accounted for (sampled) or if the nominal model was used (nominal) for the derivation of the percentiles of the count distributions.

$$N = \int\int A(log\, L_{\mathrm{X}}, z)\frac{\mathrm{d}V}{\mathrm{d}z}\phi_m(\theta)\,\mathrm{d}\,z\,\mathrm{d}\,\mathrm{log}L_{\mathrm{X}}. \tag{4.4}$$

To account for the area sensitive to a luminosity L_{X} at a redshift z, we used the model presented in Section 4.7.1. We compute N in the bin $z = 5.7 - 6.4$ and allow luminosities $\log{(L_{2-10\,\mathrm{keV}}/(\mathrm{erg/s}))} = 45.5 - 50$. The lower bound $\log{(L_{2-10\,\mathrm{keV}}/(\mathrm{erg/s}))} = 45.5$ is roughly the 1σ lower edge of the confidence interval calculated for J0836+0054 in Section 4.4.2. With these bounds we effectively probe whether or not the extrapolated XLFs are consistent with at least one detection of a quasar, which is at least as X-ray luminous as J0836+0054, in the redshift range covered by the SDSS selection in eFEDS (given its area and sensitivity). We first perform the integration on the nominal models of U14, V14, M15, B15, A15, and G15. For B15, we use the upper and lower bounds of the non-parametric XLF, because these delimit the 90% credible interval between their constant-slope model (upper bound) and constant-value model (lower bound). The uncertainties of the expected number of AGN counts from U14, V14, M15, and G15 are derived from the 1 σ errors of the parameters of the XLFs. Sampling from these, we generate 5000 counts with Eq. 4.4. For each integration result $\bar{N}$, we draw a random integer from a Poisson distribution with rate $\lambda = \bar{N}$. Similarly, we perform 5000 Poisson draws using the expected counts from the B15 lower and upper bounds and the nominal A15 model. The generated count distributions are well approximated by the log-normal distribution as can be verified using quantile-quantile plots (Waller & Turnbull 1992). The resulting count distributions are shown in Fig. 4.11. They are compared to the lower limit imposed by the detection of J0836+0054 in eFEDS. The expected counts from the nominal as well as the 50th and 84.1th percentiles of the Poisson count distributions are presented in Table 4.4.

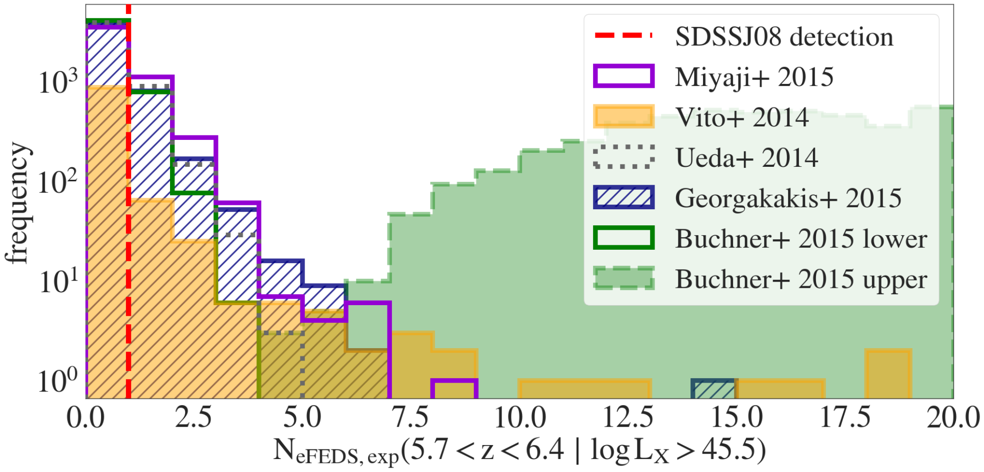


Figure 4.11: Frequency of expected eFEDS AGN source counts. The black bar is the lower limit imposed by the detection of eFEDSU J083644.0+005459.

Except for the higher bound of the XLF of B15, the nominal parametric XLFs investigated in this Chapter all predict less than one count in the probed redshift–luminosity bin. Accounting for the parameter uncertainties, the count expectations computed from the extrapolated XLFs from M15, G15, and U14 are consistent with the detection of J0836+0054 at the 84.1th percentile. The PDE model by V14 is rejected at the 88th percentile. This indicates that a shallower slope on the XLF beyond the break luminosity is favoured at $z \sim 6$. This result is best illustrated in Fig. 4.10 (right), in which we show samples drawn from the 1σ uncertainties on the XLF models from B15, M15, G15, V14, and U14. For each of these functions, expected counts were computed with Eq. 4.4 over the ranges $z = 5.7 - 6.4$ and $L_{2-10\,\mathrm{keV}}/(\mathrm{erg/s})) = 45.5 - 46$. We colour-code the sampled XLFs according to the Poisson probability of detecting at least one source in this bin in eFEDS: $1 - P_{Poisson}(0, k_{pred})$, where k_{pred} is the expected value obtained from Eq. 4.4. It clearly appears that a milder decrease in AGN space density beyond L_* (e.g. B15) is more consistent with our data.

4.8 Discussion

In this Chapter, we report the X-ray detection of J0836+0054 at z=5.81 in eFEDS. Medvedev et al. (2020) reported the detection of J1429+5447 at z=6.18 in the eROSITA all-sky survey. While both quasars were previously known from optical surveys, they were both detected as X-ray sources in 'blind' X-ray surveys, distinguishing them from distant quasars detected in dedicated follow-up observations (e.g. Brandt et al. 2002; Vito et al. 2019a; Pons et al. 2020). As such these two objects were considered the highest redshift X-ray-selected AGN discovered at the time (but see Chapter 5), and demonstrate the power of eROSITA to push the boundaries of our knowledge of accretion power at high redshift.

Even with the detection of this single object, we are able to set constraints on the evolution of the high-redshift XLF, given the very well-characterised selection function in the eFEDS field. We furthermore argue in this Chapter that J0836+0054 has X-ray emission dominated by the X-ray corona rather than, e.g., a jet. As a result, the X-ray luminosity should be representative of the bolometric luminosity. Extrapolating from the results in Section 4.7, we predict the number of X-ray luminous $z > 5.7$ quasars that will be detected by eROSITA in the full-sky survey to be ~ 100. The resulting constraints on the XLF can provide important information on the accretion power in the early Universe.

Here, we first discuss the physical nature of J0836+0054 as inferred from our results, and thereafter the implications for AGN demographics and in particular the eROSITA all-sky survey.

4.8.1 The radio core of J0836+0054

Bañados et al. (2015) classified J0836+0054 as radio loud ($\mathrm{R} = 11.9 \pm 0.3$), although close to the threshold of the definition by Kellermann et al. (1989): $\mathrm{R} = \mathrm{f}_{5\,\mathrm{GHz}}/\mathrm{f}_{4400\text{Å}} > 10$. To date, nine radio-loud $z > 5.5$ quasars have been discovered (e.g. Bañados et al. 2015; Coppejans et al. 2016), of which only three have been detected in X-rays: J0836+0054, J1429+5447 (Medvedev et al. 2020, 2021) and PSO J030947.49+271757.3 (Belladitta et al. 2020). Like J0836+0054, J1429+5447 is also classified as a steep-spectrum radio source (Coppejans et al. 2017), while PSO J030947.49+271757.3 has a flat radio spectrum, typically observed in blazars. Besides its spectral shape, J1429+5447 is also similar to J0836+0054 in terms of morphology. Indeed, Frey et al. (2011) showed with VLBI images that its radio core is confined to scales < 100 pc.

The steep radio spectral slopes of J0836+0054 and J1429+5447, as well as their compact morphologies are characteristic of compact steep-spectrum sources and peaked spectrum sources (i.e. GHz-peaked sources and MHz-peaked sources). The steep radio spectra argue against the possibility of a jetted AGN seen at a small inclination, i.e. relativistic beaming (Padovani & Urry 1992). While variability cannot be excluded, the spectral flattening at lower frequencies reported here for J0836+0054 hints at the peaked nature of its radio spectrum. MHz-peaked sources and GHz-peaked sources are thought to be at the very beginning of their evolution into large-scale radio sources (Fanaroff-Riley I or II, Fanaroff & Riley 1974), a stage at which their jets are still contained within the $\sim$ 1 kpc of their narrow line region (e.g. Orienti 2016, and references therein). Measurements of the hot-spot separation velocities of compact sources with steep spectra and the associated kinematic age support the youth hypothesis (Giroletti & Polatidis 2009). We note here that radio spectral variability in J0836+0054 would not necessarily contradict this scenario, as the adiabatic expansion of young jets would result in the shift of the peak towards lower frequencies (see Section 4.6.3). An alternative to the young radio source scenario is confinement through the surrounding dense interstellar medium (e.g. O'Dea 1998; O'Dea & Saikia 2020, and references therein). The turnover observed in the spectra of these objects is thought to be due to synchrotron self-absorption or free-free absorption through shocks in the dense environment surrounding the quasar (for a review see O'Dea & Saikia 2020), to name just a few. However, these mechanisms have difficulty in describing MHz-to-GHz spectra of both compact radio sources and extended structures of radio galaxies (e.g. Tingay & de Kool 2003). More recently, jet energy dissipation and a change in acceleration mechanism have been put forward as alternative explanations (Godfrey et al. 2009). Harris et al. (2019) discovered spectral curvature in a blazar using LOFAR long-baseline observations and showed how radio observations in the MHz energy range can improve estimates of source parameters such as the equipartition magnetic field.

4.8.2 Origin of the X-ray emission

For J0836+0054, our tentative estimation of the photon index yielded $\Gamma = 2.20^{+0.49}_{-0.60}$. Medvedev et al. (2021) performed a 20ks *XMM-Newton* DDT follow-up observation of J1429+5447. Their absorbed power-law fit to J1429+5447 yielded $\Gamma = 2.5 \pm 0.2$. The source PSO J030947.49+271757.3 was observed with *SWIFT*/XRT and its spectral analysis returned $\Gamma = 1.6 \pm 0.6$, consistent with typical blazar photon indices. While the uncertainty on Γ is too large to unambiguously demonstrate the absence (or presence) of an additional X-ray component in the spectrum of J0836+0054 due to a potential jet contribution, the posterior distribution of the photon index clearly tends towards higher values, typical for radio loud quasars at these redshifts (Vito et al. 2019a).

In summary, the eROSITA detected quasars J0836+0054 and J1429+5447 are both X-ray luminous and have steep radio spectra. They differ in two aspects:

(a) Their photon indices: J1429+5447 has a well-constrained, steep Γ, while J0836+0054 tends to a slightly flatter value, consistent with the population of X-ray-detected, radio-quiet $z > 6$ quasars presented in Vito et al. (2019a).

(b) The relative strengths of their optical and X-ray emissions: Unlike J1429+5447, J0836+0054 does not show any significant X-ray excess luminosity with respect to the bulk of the AGN population (Fig. 4.6).

These observations lead us to believe that the X-ray output of J0836+0054 is dominated by classical accretion processes and is not boosted by the relativistic bulk motion of a jet (e.g. Siemiginowska et al. 2008). Its radio core bears the typical spectral signature of confined jets, which nevertheless do not appear to contribute strongly to the overall X-ray output of the quasar. With its mass of $\sim 3 \times 10^9\,\mathrm{M}_\odot$, J0836+0054 lies at the high end of the $z > 5.8$ quasar mass distribution (e.g. Figure 7 in Onoue et al. 2019). Estimating the bolometric luminosity of J0836+0054 from $\mathrm{M}_{1450\,\mathrm{Å}}$ (Runnoe et al. 2012), we find that it accretes at $\sim 0.1\,\mathrm{L}_{\mathrm{Edd}}$. In summary, the emergent picture for J0836+0054 is that of a relatively massive and moderately accreting black hole powering a young and expanding radio core.

Investigating the properties of a large sample of radio-loud quasars, Zhu et al. (2020) showed that steep-spectrum radio quasars follow a similar $\alpha_{\mathrm{OX}} - L_{2500\,\mathrm{Å}}$ relation to that of radio-quiet quasars, indicating that the X-ray emission of these sources originates from the corona. Parametrising the corona–jet relation and performing model fitting, these latter authors find no evidence for a significant jet contribution to the X-ray output of steep spectrum quasars. Our findings for J0836+0054 fit well within this picture. The coronal origin of the X-ray emission confirms that by deriving constraints on the XLF from the detection of J0836+0054, we are truly tracing black hole accretion at high redshifts.

4.8.3 $z \sim 6$ quasar demographics from optical surveys

With an absolute UV magnitude of $\mathrm{M}_{1450\,\text{Å}} = -27.86$, J0836+0054 is the most luminous SDSS $z > 5.7$ quasar found to date (Jiang et al. 2016). As such, it belongs to a class of extremely rare objects, given the steep decline of the quasar luminosity function (QLF) beyond the break luminosity (e.g. Shen et al. 2020). The complete sample of $5.7 < z < 6.4$ quasars found in $11240\,\mathrm{deg}^2$ of the SDSS main survey Jiang et al. (2016) contains 29 extremely luminous sources ($M_{1450\,\text{Å}} < -26.22$). We investigate here if the detection of J0836+0054 is consistent with the space density of bright quasars inferred from optical surveys. We first note that the choice of the location of the eFEDS field was not driven by the presence of spectroscopically confirmed high-redshift quasars; the main motivation was the availability of a large array of complementary multi-wavelength surveys. A second observation is that J0836+0054 was initially discovered in the main single-epoch imaging survey and not in deeper fields such as overlap regions and SDSS Stripe 82. Therefore, eFEDS is not biased towards a higher density of bright SDSS $z \sim 6$ quasars.

The colour and magnitude incompleteness of high-redshift quasar surveys are encoded in well-defined selection functions (e.g. Fan et al. 2001; Jiang et al. 2016). These selection biases are accounted for in the QLF fit. In Fig. 4.6, we show contours of the eFEDS normalised sensitive area. The sensitive area is a function of the net count rate (e.g. Georgakakis et al. 2008). Its dependency on α_{OX}, i.e. $L_{2\,\mathrm{keV}}$, was computed by simulating X-ray spectra with a redshifted power law as baseline model: *clumin*tbabs*zpowerlw*. The convolutional model *clumin* enabled us to generate spectra for configurations of $L_{2-10\,\mathrm{keV}}$ and z. We fixed $\Gamma = 2$, $z = 6$ and assumed a Galactic absorption of $N_\mathrm{H} = 3 \times 10^{20}\mathrm{cm}^{-2}$. By converting the broad-band restframe luminosities $L_{2-10\,\mathrm{keV}}$ to the monochromatic $L_{2\,\mathrm{keV}}$ we were able to derive the eFEDS sensitivity for a grid of α_{OX} and $M_{1450\,\text{Å}}$ using Eq. 4.1 (a similar procedure was applied in 4.7.1). eROSITA is sensitive to $z \sim 6$ quasars which have an $\alpha_{OX}(L_{2500\,\text{Å}})$ within 1σ of the $\alpha_{ox} - L_{2500\,\text{Å}}$ relation of Lusso et al. (2010) beyond $M_{1450\,\text{Å}} < -24$.

We can obtain the expected number of sources beyond a certain UV luminosity threshold at a given redshift by computing:

$$N(< M_{1450}, z + \Delta z) = \int_{z}^{\Delta z} \int_{-\infty}^{M_{1450}} \phi_{UV}(M, z)\, \Omega \frac{\mathrm{d}V(z)}{\mathrm{d}z}\, \mathrm{d}M\, \mathrm{d}z, \tag{4.5}$$

where ϕ_{UV} is the quasar UV luminosity function, Ω is the solid angle subtended by the survey, and $\frac{\mathrm{d}V(z)}{\mathrm{d}z}$ is the differential comoving volume (e.g. Manti et al. 2017). From the best-fitting double power-law model for the UV luminosity function derived by Kulkarni et al. (2019), which was fitted to a sample including the high-redshift SDSS quasars of Jiang et al. (2016), we find that we expect $0.79^{+13.19}_{-0.76}$ sources at $M_{1450\text{Å}} < -26.22$ (i.e. the absolute magnitude of the faintest high-redshift quasar from the main survey) and $z \in [5.7, 6.4]$ in an eFEDS-sized field. These calculations account for the 1σ uncertainties on the best-fitting parameters of the broken power law. We note that, at $z \sim 6$, the scatter on the fit parameters is large. From the same calculation using the best-fitting nominal double power-law model of Jiang et al. (2016) we obtain approximately one expected source count in eFEDS. In order to compute the number of bright quasars that we expect to detect with eROSITA in eFEDS, we assume a fixed $\alpha_{ox} - L_{2500\text{Å}}$ (Lusso et al. 2010) and positive offsets therefrom (in fractions of its 1σ uncertainty). Fixing $\alpha_{ox}(L_{2500\text{Å}})$, we can derive the eFEDS sensitivity as a function of $M_{1450\text{Å}}$ and account for it in the integration of Eq. 4.5. A spectral slope of $\alpha \sim -0.3$ was assumed to convert $L_{2500\text{Å}}$ to $M_{1450\text{Å}}$. The cumulative expected integrated counts at $z = 6$ in eFEDS for the UV QLF fitted by Kulkarni et al. (2019) are shown in Fig. 4.12. The $\alpha_{ox} - L_{2500\text{Å}}$ relation from Lusso et al. (2010) has been assumed. At $M_{1450\text{Å}} < -26.22$, we can expect to detect one source assuming at least a $+0.2\sigma$ deviation from the Lusso et al. (2010) $\alpha_{ox} - L_{2500\text{Å}}$ scaling relation.

The detection of J0836+0054 is therefore consistent with the findings of optical surveys and does not require a significant deviation from the $\alpha_{ox} - L_{2500}$ relation. We note that Vito et al. (2019a) found no significant evolution of α_{OX} with redshift. A larger sample of X-ray-selected quasars will be needed to further characterise the corona-to-disc relation at high redshifts. The detection of sources such as J1429+5447 with additional jet-driven X-ray-emission components may point to a greater diversity in terms of optical to X-ray properties.

At high redshifts, cosmic variance is another important source of uncertainty in the space density measurement of quasars (Trenti & Stiavelli 2008; Robertson 2010; Moster et al. 2011; Bhowmick et al. 2020). This is particularly true at the bright end of the QLF, because the most luminous sources are expected to populate the most massive haloes. However, the relatively large sky area covered by the eFEDS significantly reduces the effect of clustering in the large-scale structure on the expected space density of high-z quasars. Using the method of Trenti & Stiavelli (2008), which is based on mock observations of dark-matter-only simulations, we compute the cosmic variance in the redshift selection window detailed in Jiang et al. (2016) over the $140\,\text{deg}^2$ of eFEDS. A duty cycle of 0.5 was assumed. We obtain a vanishing relative cosmic variance, that is, the effect of large-scale structure is negligible compared to the Poisson noise.

4.8.4 eRASS:8 count prediction

We investigate how the detection of an X-ray source associated with J0836+0054 can be used to predict number counts of high-redshift AGN in eRASS:8. Given that eFEDS was initially designed to reach the average final depth of the all-sky survey, a first-order approach would consist in re-scaling the single detection in eFEDS to the full-extragalactic sky (i.e. eROSITA-DE+eROITA-RU; $34100\,\text{deg}^2$, $\mid b \mid > 10°$). The exact sky-area of eFEDS is: 142deg^2. The area-scale factor is therefore: $s_{AREA} = 244$. Considering one detection in eFEDS at $z > 5.7$ (the lowest redshift of all spectroscopically confirmed high-redshift quasars in eFEDS), error bars on the number of expected detections in a field of eFEDS-like depth and area can be obtained by inverting the Poisson probability distribution. The lower (upper) bounds are 0.17 (1.8). Multiplying these by the scale factor, we obtain

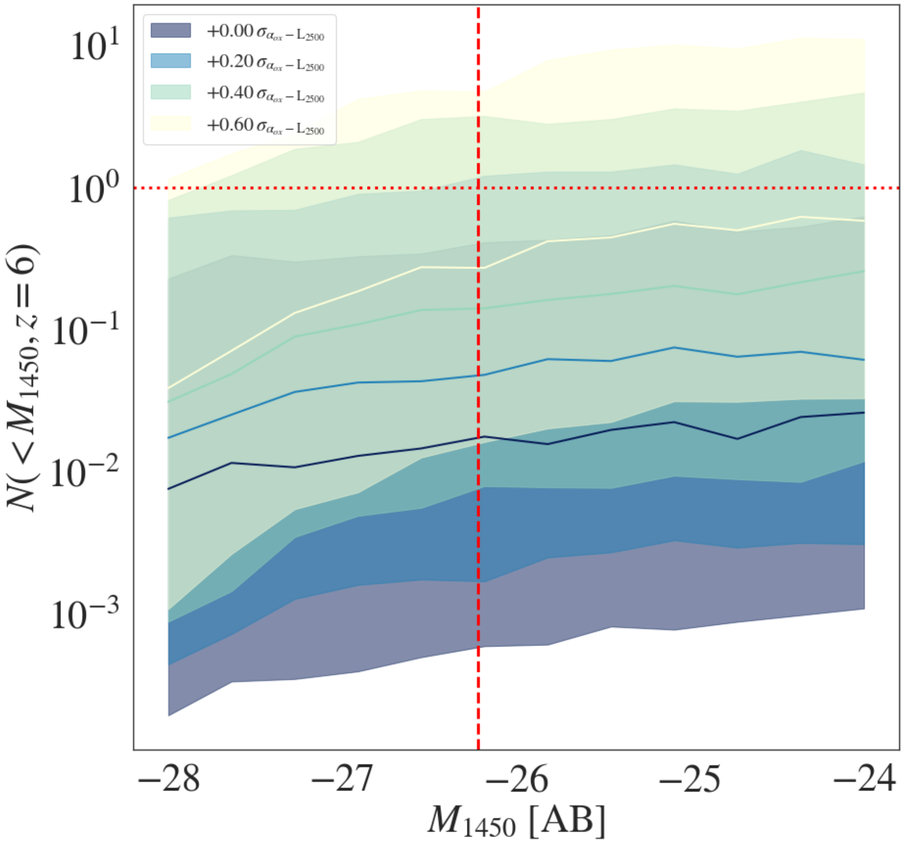


Figure 4.12: Confidence intervals for the expected number of $z \sim 6$ optically selected quasars detectable by eROSITA in eFEDS. These predictions are derived from the UV QLF presented in Kulkarni et al. (2019), for a fixed $\alpha_{ox} - L_{2500}$ relation (Lusso et al. 2010) and deviations therefrom. The eFEDS sensitivity is accounted for. The vertical line shows the magnitude limit of the SDSS main survey sample of Jiang et al. (2016). The horizontal line marks the single detection limit. Within 1σ of the typical α_{OX} the detection of J0836+0054 in eFEDS is consistent with predictions from optical surveys.

$N_{scaled} = 244^{+195}_{-202}$. We note that the prediction assumes that the J0836+0054 is the only source in eFEDS in this redshift regime. However, the average exposure of eFEDS is near-uniformly $\sim 2.3\,\mathrm{ks}$ while the eROSITA scan pattern makes the exposure of eRASS:8 non-uniform, with an average of $\sim$ 1.6ks in the equatorial region (Clerc et al. 2018). N_{scaled} therefore possibly over-predicts the actual number of detectable $z > 5.7$ sources in eRASS:8.

Alternatively, using Eq. 4.4, we can also make predictions for eRASS:8 using XLF extrapolations which are consistent with the eFEDS detection. Kolodzig et al. (2013) followed a similar approach to compute pre-mission estimates using the LDDE XLF parametrisation of Hasinger et al. (2005). We use the methodology detailed in Section 4.7.1, accounting this time for the predicted eRASS:8 sensitivity to point sources. Clerc et al. (2018) generated sensitivity curves for eRASS:8, by computing the selection function of point sources from a simulated eROSITA sky in three exposure modes: equatorial ($\sim 1.6\,\mathrm{ks}$), intermediate ($\sim 4\,\mathrm{ks}$), and deep ($\sim 9.7\,\mathrm{ks}$). For a more conservative estimate, we select the sensitivity curve computed from the shallower equatorial simulation. We integrate the XLFs over a redshift range $z = 5.7 - 6.4$ and luminosity range $\log\left(L_{2-10\,\mathrm{keV}}/(\mathrm{erg/s})\right) = 45.5 - 50$. Sampling from the uncertainties on the fit parameters, we integrate XLFs by M15, U14, and G15 and obtain distributions of expected source counts. We compute weighted percentiles of these distributions. The weight we ascribe to each count is the Poisson probability of detecting at least one source in the redshift-luminosity bin and in eFEDS given the sampled XLF model (see Section 4.7.2). The resulting weighted percentiles are reported in Table 4.5. From the extrapolated XLFs we obtain an average of 88 detections in the probed redshift–luminosity bin. The confidence intervals spanned by the 15.9th and 84.1th percentiles are large and right-skewed because of the poor constraints on the parameters governing the shape of XLF models. This prediction is higher than the one presented in Kolodzig et al. (2013) by a factor of about three. The expected value from simple area scaling is more optimistic with 244 counts, but the error-bars obtained from the inversion of the Poisson distribution are larger than the confidence intervals from the XLF predictions. We underline the conservative nature of the XLF estimates, which is due to the following: (1) eFEDS is currently 30% redshift-complete, (2) the integration of extrapolated XLFs favours a non-detection, with the eFEDS detection only being supported at the $+1\sigma$ limit, and (3) we have not accounted for regions of deeper exposure in the all-sky survey.

4.9 Conclusions

We report the blind detection of eFEDSU J083644.0+005459, an eROSITA X-ray source matched to the known quasar SDSS J083643.85+005453.3 (z=5.81). The detection is robust in terms of X-ray photon counts, astrometry, and multi-wavelength counterpart association. The eROSITA flux of the source is consistent with previous X-ray observations carried out with Chandra (Brandt et al. 2002).

From GHz radio surveys, SDSS J083643.85+005453.3 is known to host a steep spectrum radio core within its central 40 pc. With the LOFAR 145 MHz and ASKAP 888 MHz

Mod.	$N_{eRASS:8}^{(5.7<z<6.4)}$	$N_{eRASS:8}^{15.9-th}$	$N_{eRASS:8}^{84.1-th}$
G15	**99**	38	245
U14	**73**	38	136
M15	**92**	41	204
s_{AREA}	**244**	42	439

Table 4.5: The 50th, 15.9th, and 84.1th percentiles of the count predictions of $5.7 < z < 6.4$ and $\log(L_{2-10\,\mathrm{keV}}/(\mathrm{erg/s})) > 45.5$ AGN that will be detected with eROSITA in eRASS:8 (restricted to the extragalactic sky, 34100 deg^2). The counts are obtained by sampling from the XLF models of G15, U14, and M15. The distributions are weighted by the probability of the sampled XLFs of supporting at least one detection in eFEDS. The results for s_{AREA} are obtained from naive area-scaling. The edges of the 1σ confidence interval are estimated from the percentile method.

observations, we confirm a spectral flattening at frequencies below 1 GHz. The shape of its radio spectrum indicates that this quasar has (possibly young) jets confined in its central region. Alternatively, the observed flattening of the spectral slope could be the spectral signature of adiabatically expanding jets.

The multi-wavelength properties of the quasar are evidence against relativistic beaming or iC-CMB boosting of the X-ray emission, suggesting that it originates in the X-ray-emitting corona. We examined the constraints on the XLF implied by this detection which favour a relatively shallow slope of the XLF beyond the break luminosity at $z \sim 6$. The population of X-ray-luminous high-redshift quasars may therefore be larger than previously thought. From the parametric XLFs presented by G15, M15, and U14 we predict the detection of ~ 90 AGN at the bright end of the XLF ($z = 5.7 - 6.4$ and $\log(L_{2-10\,\mathrm{keV}}/(\mathrm{erg/s})) > 45.5$) in the eROSITA full-sky survey by the end of the final scan.

Chapter 5

An X-ray loud quasar at $z > 6$

The results presented in this chapter were published in "X-ray emission from a rapidly accreting narrow-line Seyfert 1 galaxy at z = 6.56" ***Wolf J.****, Nandra K., Salvato M., Buchner J. et al., (2023), Astronomy and Astrophysics, 669, A127.*

5.1 Summary

Here, we presented the eROSITA X-ray detection of a second high-redshift quasar in the eFEDS field: the $z = 6.56$ quasar J0921+0007, initially discovered in a dedicated survey based on the HSC Subaru Strategic Program (SSP; Aihara et al. 2022): the Subaru High-z Exploration of Low-luminosity Quasars (SHELLQs; Matsuoka et al. 2018a). The source is optically faint but X-ray bright. Its optical and near-infrared (NIR) spectral properties potentially make it a high-redshift NLS1. We present the eROSITA detection of this optically faint source with a 21 ks *Chandra* Advanced CCD Imaging Spectrometer (spectroscopy array, ACIS-S) follow-up observation and derive its X-ray properties in Section 5.3. We derive the black hole mass of the source and Eddington ratio with a new Ks-band spectrum that covers the MgII region in Section 5.4. We connect the global optical and X-ray emission in Section 5.6 and show that J0921+0007 is X-ray loud. It is the highest-redshift blindly detected X-ray source to date, and its detection further supports a flattening of the XLF beyond the break luminosity, L_*. We present its contribution to the global accretion density in Section 5.7.

Dr. Masafusa Onoue acquired and documented the NIR spectroscopy of the source presented in Section 5.4.1. Dr. Stefano Ciroi and Dr. Francesco Di Mille obtained the optical spectrum used in Section 5.5 to measure the size of the proximity zone.

As in the previous Chapter, we have assumed a standard Λ cold dark matter cosmology with parameters from Planck Collaboration et al. (2020). Throughout this Chapter, uncertainties are quoted at the 68% confidence level (1σ).

5.2 Introduction: Narrow-line Seyfert 1, a population of strongly accreting, young black holes at high redshifts ?

While the bulk of active galaxies discovered at $z > 5.7$ host black holes with masses of 1 to 10 billion solar masses (e.g. Onoue et al. 2019; Yang et al. 2021), there should be a population of less massive ($10^8\,M_\odot$), super-Eddington-accreting SMBHs caught in an earlier evolutionary state. Some of these less massive, strongly accreting black holes have been found at the centre of high-redshift quasars that display optical properties similar to local NLS1s (e.g., Koptelova et al. 2017, 2019; Onoue et al. 2019). NLS1s are a class of active galaxies that are defined by their narrow Hβ emission lines (FWHM $< 2000\,\mathrm{km\,s^{-1}}$) and the weakness of their [OIII] narrow-line emission relative to Hβ, [OIII]/H$\beta < 3$ (Osterbrock & Dahari 1983; Goodrich 1989). They show strong FeII emission (Osterbrock & Pogge 1985), typically host SMBHs with lower black hole masses ($M_\mathrm{BH} < 10^8\,M_\odot$), and accrete at a significant fraction of their Eddington limit ($10-100\%$; e.g. Pounds et al. 1995; Grupe et al. 2010; Rakshit et al. 2017; Waddell & Gallo 2020), as expected from young and strongly accreting black holes. Large amplitude, short timescale flaring behaviour in the UV continuum has been observed for this class of source (Collier et al. 2001). Rapid, high amplitude variability is also seen at shorter wavelengths, in X-rays (e.g. Turner et al. 2001; Romano et al. 2002). NLS1s usually have steeper X-ray spectra (i.e. larger power-law photon indices). than typical broad-line Seyfert 1 galaxies (Nandra & Pounds 1994; Boller et al. 1996).

Beyond the end of the epoch of re-ionisation, hard X-ray photons unhindered by dust and gas are collectable by sensitive soft X-ray telescopes at observer-frame energies, $\sim 2\,\mathrm{keV}$. To date, ~ 50 quasars at $z > 5.7$ have been observed in X-rays, mostly via pointed observations with *Chandra* and *XMM-Newton* (Brandt et al. 2002; Vignali et al. 2003; Nanni et al. 2017; Vito et al. 2019a) following their discovery with optical telescopes. Recently, Barlow-Hall et al. (2022) reported the blind detection of a quasar spectroscopically confirmed at $z = 6.31$ in the Extragalactic Serendipitous Swift Survey (ExSeSS). The first unbiased, blind X-ray detections of quasars at $z > 5.7$ in the performance verification fields and all-sky maps of the extended ROentgen Survey with an Imaging Telescope Array (eROSITA; Predehl et al. 2021; Sunyaev et al. 2021) were reported by Medvedev et al. (2020) and Chapter 4 (Wolf et al. 2021). Khorunzhev et al. (2021) present the discovery with eROSITA of the most X-ray-luminous quasar beyond $z \gtrsim 5.5$ with $\log L_{2-10\,\mathrm{keV}} = 3 \times 10^{46}\,\mathrm{erg\,s^{-1}}$. In addition to being among the most X-ray-luminous quasars at the end of re-ionisation, all of the eROSITA-detected high-redshift quasars are radio detected and radio loud (according to the radio-loudness definition $R = f_{\nu,5\,\mathrm{GHz}}/f_{\nu,4400\,\mathrm{\mathring{A}}} > 10$ defined by Kellermann et al. 1989). However, the $z = 5.81$ quasar detected in (eFEDS; Brunner et al. 2022) does not show evidence of any jet contribution in its X-ray output, making it a secure probe of coronal activity and hence black hole accretion history (Wolf et al. 2021, Chapter 4,). The direct X-ray selection of this spectroscopically confirmed quasar in a contiguous field of uniform exposure imposes constraints on the XLF just after the epoch of re-ionisation ($z \sim 6$). In Chapter 4 we show that an exponential decline at high redshift cannot be excluded. However, models that show a shallower slope at the bright end of the XLF are preferred in the probed high-redshift bin.

5.3 HSC J092120.56+000722.9: An X-ray-luminous quasar

5.3.1 eROSITA detection

In eFEDS, a detection likelihood threshold DET_LIKE $\geq$ 6 was applied. A supplementary catalogue[1] of 4774 eROSITA/eFEDS sources detected just below this threshold (5 < DET_LIKE < 6) was also made available. It is expected to contain a high fraction of spurious sources, but it also gives access to interesting faint objects. Using simulations, Brunner et al. (2022) show that reducing the source detection likelihood threshold from DET_LIKE = 6 to DET_LIKE = 5 results in an increase in the number of detections of the simulated point-sources. They report an increase (from 93 % to 94 %) in the fraction of detected simulated point-sources brighter than $F_{0.5-2\,\mathrm{keV}} = 1 \times 10^{-14}\,\mathrm{erg\,s^{-1}cm^{-2}}$ and an increase (from 59 % to 63 %) in the fraction of detected simulated sources brighter than $F_{0.5-2\,\mathrm{keV}} = 4 \times 10^{-15}\,\mathrm{erg\,s^{-1}cm^{-2}}$. Similarly, Liu et al. (2022b), performing a standard single-band (0.2-2.3 keV) detection run on a simulated eFEDS field, show that reducing the DET_LIKE threshold from 6 to 5 results in an overall increase in completeness from $\sim 79\%$ to $\sim 82\%$ and an increase in the spurious fraction from $\sim 7\%$ to $\sim 12\%$. Similarly to the procedure described in Chapter 4, we identified LS8 optical counterparts to sources in the supplementary catalogue using the Bayesian cross-matching algorithm `NWAY` (Salvato et al. 2018, 2022).

The resulting best optical counterpart solutions were positionally cross-matched to a complete list of spectroscopically confirmed $z > 5.5$ quasars in the eFEDS footprint compiled from literature (31 sources, Fan et al. 2001; Venemans et al. 2015; Matsuoka et al. 2016, 2018a,b, 2022) within $1''$. This exercise returned one match. The eFEDS source J092120.6+000725.9, hereafter J0921+0007, has a LS8 match within $3.11''$ of its centroid, which coincides spatially with the spectroscopically confirmed SHELLQs quasar J0921+0007 ($z = 6.56$, Matsuoka et al. 2018a) within $0.08''$. The spectroscopic redshift of this quasar was securely measured by Matsuoka et al. (2018a), using the Lyα emission line. It was subsequently confirmed by Onoue et al. (in prep.) using MgII and Yang et al. (2022) who found $z = 6.5646 \pm 0.0003$ using [CII]. There were 22 possible LS8 counterpart candidates within $30''$ of the X-ray source J0921+0007. The individual probability for the chosen LS8 counterpart of being the correct one amongst the candidates is unequivocally high `p_i` = 0.82 (other candidates have `p_i` < 0.13). For a definition of `p_i`, see Chapter 2. A summary of the `NWAY` match is presented in Table 5.1.

---

[1] https://erosita.mpe.mpg.de/edr/eROSITAObservations/Catalogues/

eROSITA ID	-	22224
eROSITA Name	-	eFEDS J0921+0007
$\mathrm{RA}_{\mathrm{eFEDS}}$	[deg]	140.3361
$\mathrm{DEC}_{\mathrm{eFEDS}}$	[deg]	0.1237
$\sigma_{\mathrm{RADEC,eFEDS}}$	[arcsec]	3.87
DET_LIKE (0.2-2.3 keV)	-	5.04
Net Counts (0.2-2.3 keV)	-	10.5 ± 4.7
LS8 objID/brickID	-	5281/330929
Sep.eRO/LS8	[arcsec]	2.77
QSO ID	-	HSC J0921+0007
QSO Redshift	-	6.56
Sep.QSO/LS8	[arcsec]	<0.1

Table 5.1: Basic source and counterpart information of eFEDS J0921+0007. The coordinates of the eFEDS source are equatorial, with $\sigma_{\mathrm{RADEC,\,eFEDS}}$ being the 1σ X-ray positional uncertainty. The net counts and errors are obtained via photon-mode PSF fitting (Brunner et al. 2022). Sep. eRO/LS8 measures the separation between the centroid of the eFEDS X-ray source and the position of the LS8 counterpart. Sep. QSO/LS8 corresponds to the separation between the quasar optical position and the LS8 counterpart.

While the detection likelihood is at a low level where a large spurious fraction is expected, the alignment with a high-redshift quasar strengthens the detection. At an X-ray detection likelihood of 5, 12% of sources in the eFEDS field are expected to be spurious (Liu et al. 2022b). We estimated the probability of a chance alignment of any spectroscopically confirmed high-redshift quasar in the eFEDS footprint with a spurious detection (i.e. a background fluctuation detected as catalogue source). For this we first observe that, in the eFEDS footprint excluding the borders with lower exposure, higher background and stronger vignetting (90% of the total area; see Liu et al. 2022a) there are 29482 sources detected in the 0.2-2.3 keV band above the detection likelihood of the quasar DET_LIKE > 5.04. Among these sources, 3277.6 are expected to be spurious from simulations. We thus obtained the probability of a chance alignment as a function of the separating distance R of a quasar and a spurious source as: $31 \times 3277.6\pi R^2/\mathrm{area}_{90}$. Here area_{90} is the area of the '90%' region: 1640219392 arcsec^2. The probability of a spurious chance alignment as a function of radius is shown in Fig. 5.1. Accounting for the eFEDS bi-variate positional error ($\sigma = \mathrm{RADEC\_ERR}/\sqrt{2} = 3.87''$), within the maximum separation between J0921+0007 and the eROSITA source, the probability of a chance alignment of the quasar with a spurious source is less than 1%.

In order to confirm the eROSITA detection, we obtained a 21ks *Chandra* follow-up observation (GTO proposal, cycle 22, ObsID 24738) pointed at the optical position of the quasar.

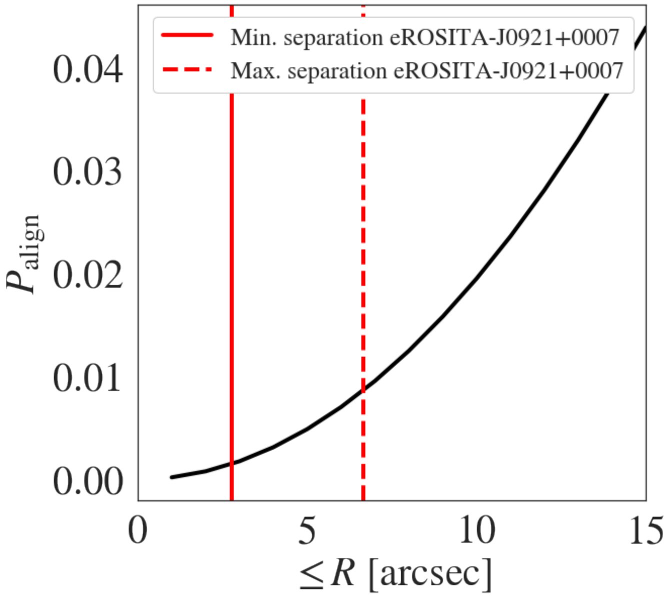


Figure 5.1: Probability of finding any of the 31 spectroscopically confirmed quasars in the eFEDS footprint within a distance R of a spurious X-ray source. The solid (dashed) red line shows the minimum (maximum) distance between the eROSITA source and J0921+0007, accounting for the positional uncertainty of the X-ray source.

5.3.2 Confirmation with a Chandra pointed observation

On October 26, 2021, the quasar was observed with *Chandra* ACIS-S with a total exposure time of 21.47 ks (PI:Predehl, Observer: Wolf). On the 0.5 - 7 keV band *Chandra* image, sources were detected with the mexican-hat wavelet algorithm *wavdetect* from the `CIAO` software package. The default detection parameters for the pixel radii (`scales`) as well as the significance thresholds (`sighthresh` and `bkgsigthresh`) have been used (respectively 2 and 4 pixels, 10^{-6} and 0.001). We confirm the significant detection of an X-ray source whose centroid lies within $0.78''$ of the optical coordinates of the quasar. The positional counts-weighted variances in pixels are X_ERR = 0.45 and Y_ERR = 0.28 (RA_ERR = $0.22''$ and DEC_ERR = $0.14''$), for a point-spread function (PSF) size of $0.46''$.

No other source was detected within $30''$ of the optical quasar coordinates in the 0.5 - 7 keV band, confirming the `NWAY` identification. A high-resolution broadband image and the *wavdetect* 3σ elliptical source detection region are presented in the right panel of Fig. 5.2. In parallel to this automated detection procedure, we performed forced photometry at the quasar position on the 0.5-7 keV *Chandra* images. We extracted source counts in a circular region of $2''$ radius centred on the coordinates of the quasar. Background counts were extracted in a ring with inner and outer radii of $4''$ and $20''$. As shown in the right panel of Fig. 5.2, no other bright source is present in the background region. We computed the binomial no-source probability (e.g. Weisskopf et al. 2007; Vito et al. 2019a) as

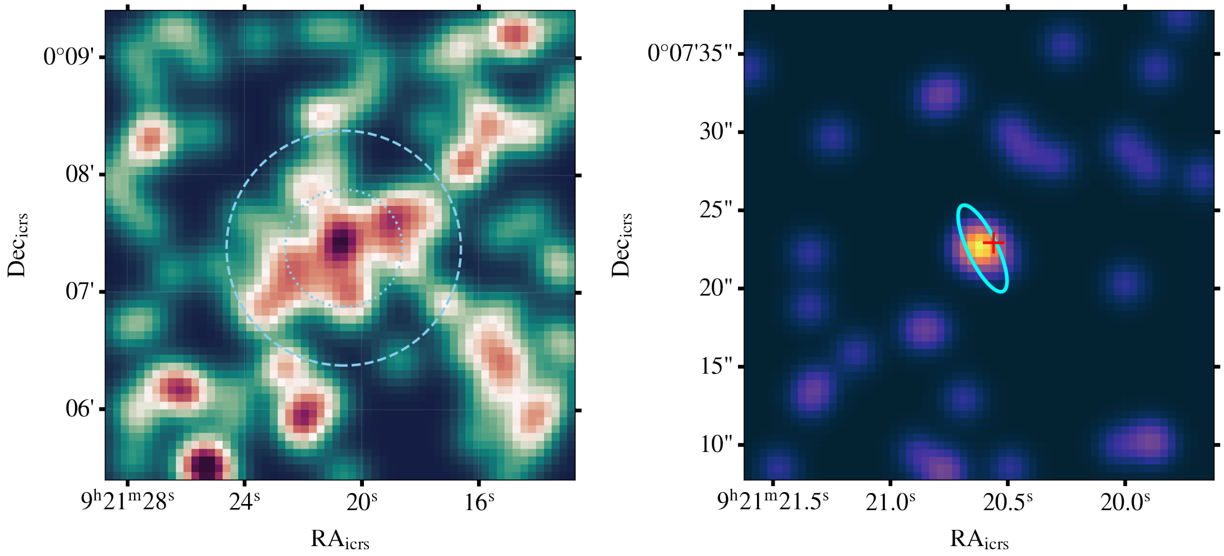


Figure 5.2: X-ray image cutouts in the region of J0921+0007. *Left:* $4' \times 4'$ eROSITA/eFEDS image in the $0.2 - 2.3\,\mathrm{keV}$ band, smoothed with a Gaussian kernel. The image is centred on the optical position of the quasar J0921+0007. The concentric circles have radii of $30''$ (dotted) and $60''$ (dashed). *Right:* $30'' \times 30''$ high-resolution *Chandra* image in the energy range 0.5 - 7 keV. The ellipse shows a *wavdetect*-detected source at the optical position of the quasar (marked by a red cross).

$$P_B(i \geq s) = \sum_{i=s}^{s+b} \frac{(s+b)!}{i!(i-s-b)!} \left(\frac{1}{1+r}\right)^i \left(\frac{2+r}{1+r}\right)^{s+b-i} . \tag{5.1}$$

Here s and b are counts in the source and background region, while r is given by the ratio of areas of these two extraction regions. In the 0.5-7 keV band, we extracted 7 counts in the source region and 52 counts in the background region. We obtained $P_{\mathrm{B},0.5-7\,\mathrm{keV}} \sim 3 \times 10^{-6}$; in other words, the source detection is highly significant. The binomial no source probabilities in the narrower energy bands 0.5 - 1.2 keV and 1.2 - 2 keV are $P_{\mathrm{B},\,0.5-1.2\,\mathrm{keV}} \sim 3 \times 10^{-4}$ and $P_{\mathrm{B},\,1.2-2\,\mathrm{keV}} \sim 6 \times 10^{-4}$.

5.3.3 X-ray properties

The forced PSF-fitting photometry results in the supplementary eFEDS catalogue show that J0921+0007 was significantly detected in the 0.5-1 keV band (DET_LIKE = 6.2). This band samples the rest-frame hard X-ray emission ($\sim 3.8 - 7.6\,\mathrm{keV}$). In this band, there are 4.74 ± 2.74 source net counts (as measured from the rate). The background at the source position is 2.42 counts/arcmin^2. The measured count rate is $r_{0.5-1\,\mathrm{keV}} = (3.86 \pm 2.23) \times 10^{-3} cts/s$. We converted the count-rate to an observed frame flux of $F_{0.5-1\,keV} = (2.8 \pm 1.6) \times 10^{-15}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$ assuming a power law with photon index $\Gamma = 2$ and Galactic

Obs.	Net rate	Flux abs.	$L_{2-10\,\mathrm{keV}}$	Γ
	10^{-4} cts/s	10^{-15} erg /cm^2/s	10^{45} erg/s	-
eROSITA 0.5-1 keV	38.6 ± 22.3	2.84 ± 1.64	2.96 ± 1.71	2
Chandra 0.5-2 keV	$2.79^{+2.47}_{-1.56}$	$3.26^{+2.87}_{-1.83}$	$2.10^{+1.86}_{-1.17}$	2
Chandra `BXA`	-	$4.43^{+2.78}_{-1.75}$	$3.72^{+3.97}_{-1.89}$	$3.2^{+0.7}_{-0.6}$

Table 5.2: X-ray properties of J0921+0007. The two first lines present the photometry derived from the eROSITA and *Chandra* observations of J0921+0007 assuming a fixed spectral model ($\Gamma = 2$). The flux and luminosities derived from the spectral fit to the *Chandra* data are labelled '*Chandra* `BXA`'.

foreground absorption $N_\mathrm{H} = 2.65 \times 10^{20}\,\mathrm{cm}^{-2}$ from HI4PI Collaboration et al. (2016)[2]. This corresponds to a rest-frame $2-10\,\mathrm{keV}$ luminosity of $L_{2-10\,\mathrm{keV}} = (2.96 \pm 1.71) \times 10^{45} \mathrm{erg\,s}^{-1}$. We also computed count rates and fluxes from the *Chandra* follow-up data using the `srcflux` script of the `CIAO` software package. We manually selected source and background regions centred at the optical coordinates of the quasar. The circular source region has a radius of $2''$, while the annulus describing the background region has radii $(4'', 13'')$. Counts were extracted in the 0.5-2 keV band. For the PSF model, we opted for the `arfcorr` method. We obtained the net count rate: $r_{0.5-2\,\mathrm{keV}} = (2.8^{+2.5}_{-1.6}) \times 10^{-4}$cts/s. Assuming a nominal absorbed power law with $\Gamma = 2$ and $N_\mathrm{H} = 2.65 \times 10^{20}\,\mathrm{cm}^{-2}$, we obtained the absorption-corrected energy flux $F_{0.5-2.0\,\mathrm{keV}} = (3.5^{+3.1}_{-2.0}) \times 10^{-15}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$ and the 2-10 keV luminosity $L_{2-10\,\mathrm{keV}} = (2.1^{+1.9}_{-1.2}) \times 10^{45} \mathrm{erg\,s}^{-1}$. The luminosity derived from the *Chandra* data is consistent with the eROSITA results within the 1σ measurement uncertainties. The large error bars are driven by the low-count statistics. Within these uncertainties, no flux variability is detected between the eROSITA observations and *Chandra* (i.e. $\Delta t \sim 95$ days) in the quasars rest frame.

We also performed a tentative spectral fit of the counts extracted from the manually defined source and background regions. We used the Bayesian analysis software `BXA` (Buchner et al. 2014) coupled to the X-ray fitting library `XSPEC` (Arnaud 1996) and the nested sampling algorithm `ultranest` (Buchner 2021). We restricted the fitting region to the 0.1-8.5 keV band. Given the low number of counts, we adopt an absorbed, redshifted power-law model: *tbabs*zpowerlw*, with photoelectric absorption fixed to the Galactic absorbing column density N_H. The normalisation and the photon index Γ were allowed to vary in the fit, assuming a wide, uniform prior of $\Gamma = 1.5 - 5$ for the photon index and a Jeffreys prior $A = 10^{-4} - 10^{-2}$ photons/keV/cm^2 on the norm. The redshift of *zpowerlw* was fixed to the redshift of the quasar, $z = 6.56$. We used the C-statistic for the spectral analysis

---

[2] At this stage the choice of the spectral shape is arbitrary. The aim here is a consistent comparison of the eFEDS and *Chandra* data.

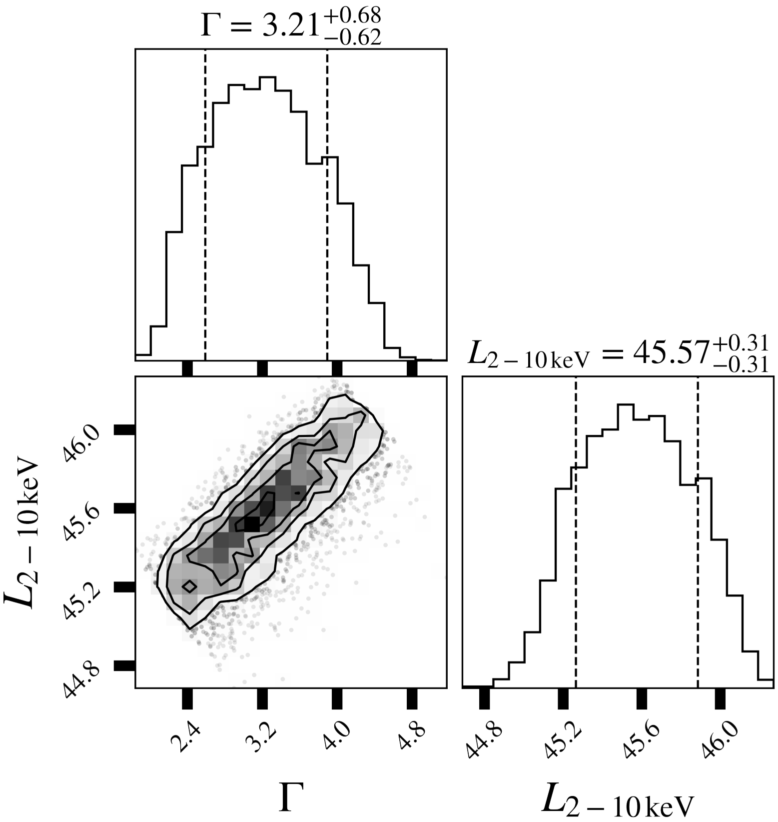


Figure 5.3: Marginal posterior distributions of the photon index, Γ, and the hard X-ray luminosity from the `BXA` fit to the *Chandra* spectrum.

(Cash 1979). The marginal posterior distributions of the photon index and $L_{2-10\,\mathrm{keV}}$ are presented in Fig. 5.3. We obtained a steep power-law photon index $\Gamma = 3.2^{+0.7}_{-0.6}$. Even assuming no intrinsic absorption, the photon index is higher than typical quasars in this redshift regime ($\Gamma = 2.20$, Vito et al. 2019a). A possible explanation for this steep soft spectrum is the presence of a soft-excess (Arnaud et al. 1985; Boller et al. 1996; Magdziarz et al. 1998; Gierliński & Done 2004; Ross & Fabian 2005; Crummy et al. 2006; Walton et al. 2013). We derived absorbed fluxes and luminosities, accounting for the posterior samples of Γ and the power-law normalisation (see Table 5.2). We adopted $\Gamma = 3.2$ for the spectral shape of J0921+0007 hereafter.

Table 5.2 summarises the X-ray photometry derived from the eROSITA and *Chandra* data. Fig. 5.4 presents the 2-10 keV luminosity-redshift plane for an up-to-date sample of X-ray-detected quasars (Nanni et al. 2017; Vito et al. 2019a; Pons et al. 2020; Belladitta et al. 2020; Medvedev et al. 2021; Khorunzhev et al. 2021; Wolf et al. 2021). It also displays the eFEDS normalised sensitive area. The area sensitivity curves are obtained with the eSASS task `apetool`. `apetool` returns the sensitive area in square degrees as a function of source count-rates. We converted count-rates to $L_{2-10\,\mathrm{keV}}$ assuming an absorbed power-law model with $\Gamma = 3.2$ and $N_{\rm H} = 3 \times 10^{20}\,{\rm cm}^{-2}$. The luminosity of J0921+0007 derived from the spectral analysis is shown. The quasar lies at the sensitivity limit of the survey.

5.4 Physical properties and active galactic nucleus type from a Ks-band spectrum

J0921+0007 was discovered by Matsuoka et al. (2018a) in the dedicated HSC high-redshift quasar survey SHELLQs. It was selected as quasar candidate based on its red $i - z$ colour and retained by a Bayesian selection method detailed by Matsuoka et al. (2016). The photometry of this quasar resembles that of Galactic brown dwarfs and could only be disentangled from stellar contaminants due to an unusually strong Lyα + NV complex, which steepens the $z - Y$ slope (Matsuoka et al. 2018a). Indeed, J0921+0007 possesses the second most luminous Lyman α line of the entire SHELLQs $z > 5.8$ quasar sample with $\log\left(L_{{\rm Ly}\alpha}/({\rm erg\,s^{-1}})\right) = 45.04 \pm 0.01$ (Matsuoka et al. 2016, 2018a). Matsuoka et al. (2018a) further report a relatively narrow line width ${\rm FWHM}_{{\rm Ly}\alpha} = 1400 \pm 100\,{\rm km\,s^{-1}}$ compared to the bulk of the sources in the SHELLQs sample. We present a new NIR observation of this source, derive its black hole properties and determine its AGN type.

5.4.1 Black hole mass and accretion rate

We estimated the black hole mass from infrared spectroscopy. The MgII λ2798 emission line, which can be used as a virial black hole mass estimator (Vestergaard & Osmer 2009) is redshifted outside the optical discovery spectrum, which ends at 1.02um (1330 $\mathring{A}$ in the rest frame), requiring NIR spectroscopy.

The Ks-band spectrum of J0921+0007 was obtained on April 22, 2019 by MOIRCS (Ichikawa et al. 2006; Suzuki et al. 2008), a Cassegrain instrument mounted on the Subaru Telescope. The observation was performed in the multi-object spectroscopy mode for secure target acquisition. The VPH-K grism (Ebizuka et al. 2011) was used to cover the 2.0 – 2.4 μm range with a spectral resolution of $R = 1700$ for a $0.8''$ slit width. J0921+0007 was observed for 72 minutes with mean airmass 1.1 and K-band seeing size $0.8''$. More details of the observations and data analysis will be presented in Onoue et al. (in prep.).

The raw data were reduced and 1D-extracted in the standard manner based on the software system Image Reduction and Analysis Facility (IRAF). To correct for the telluric absorption an A0-type star was observed just before the exposures of J0921+0007. The telluric-corrected 1D spectrum was then scaled to the Ks-band magnitude of J0921+0007 ($= 20.691 \pm 0.052$ AB mag) that was obtained by the same run with MOIRCS. Observations in the Ks band are preferable since it covers MgII, the emission line of interest. This 10 minutes imaging observation enables us to flux-calibrate the observed K-band spectrum accurately without being affected by potential variability of the quasar. The spectrum was scaled to correct for the Galactic extinction in the K band.

Figure 5.5 shows the obtained Ks-band spectrum, where a strong MgII emission line is clearly detected. We model the spectrum with three components: power-law continuum,

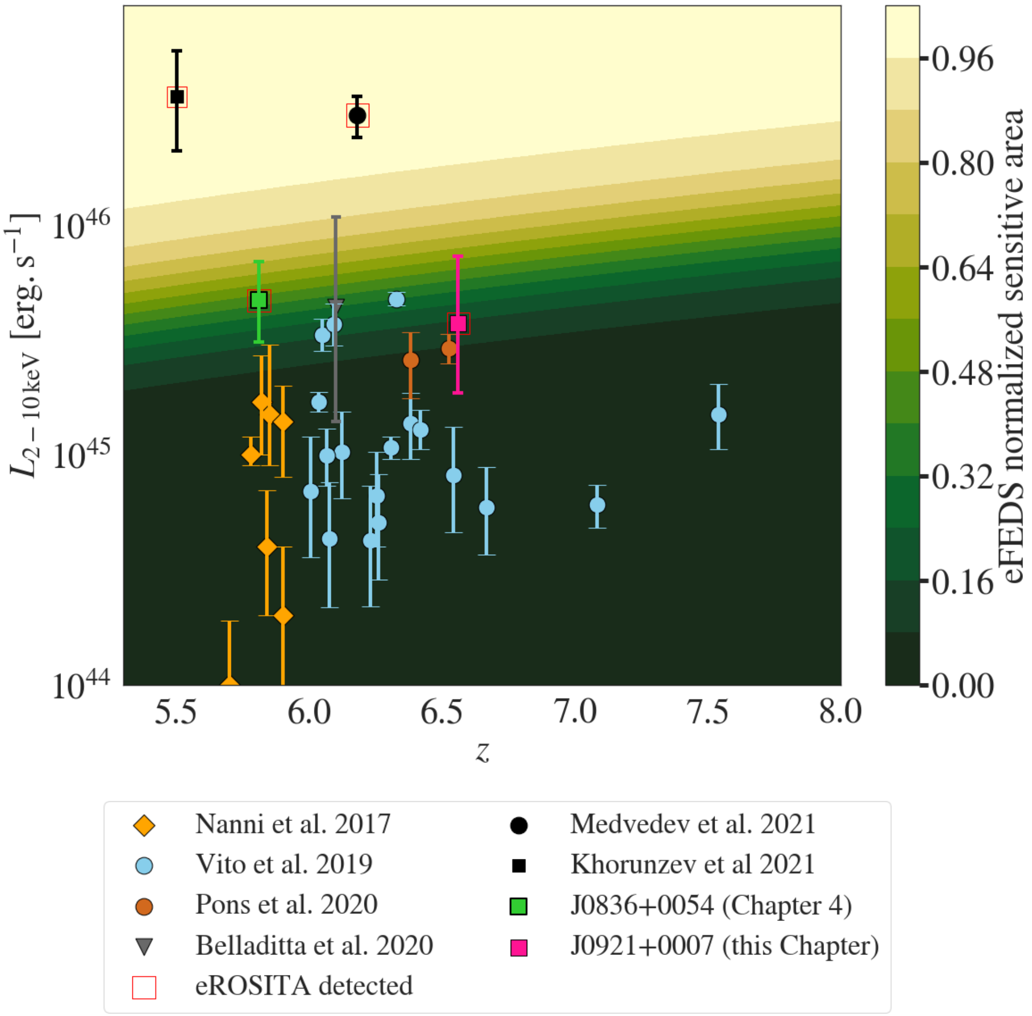


Figure 5.4: Hard X-ray luminosities of X-ray-detected high-redshift quasars. All eROSITA-detected quasars are marked by a red square. The normalised area of eFEDS that is sensitive to sources modelled by a fiducial absorbed power law ($\Gamma = 3.2$ and $N_{\mathrm{H}} = 3 \times 10^{20}\,\mathrm{cm}^{-2}$) is represented by the background colour map. J0921+0007, shown here as a pink square, lies at the detection limit of the eFEDS survey, as expected from the low eROSITA detection likelihood.

FeII pseudo-continuum, and MgII. Since the MOIRCS spectrum covers a narrow wavelength range, the continuum power-law slope ($\alpha_\lambda = -1.052$; $F_\lambda \propto \lambda^{\alpha_\lambda}$) was estimated by the photometric colour of the optical HSC y band and MOIRCS Ks band. The monochromatic luminosity at rest frame 3000 $\mathring{A}$ $L_{3000} = (4.8 \pm 0.2) \times 10^{42}$ erg s^{-1} s^{-1} $\mathring{A}^{-1}$ was derived by the scaled power-law continuum model with $\alpha_\lambda = -1.021$. The 3000 $\mathring{A}$ luminosity was then converted to the bolometric luminosity $L_{\rm bol} = (7.4 \pm 0.3) \times 10^{46}$ erg s^{-1} assuming a bolometric factor of 5.5 (Richards et al. 2006). For FeII emission lines, the empirical template of a local narrow-line Seyfert galaxy, 1 Zw 1 (Tsuzuki et al. 2006) was convolved with a Gaussian kernel and fitted to the observed continuum together with the power-law continuum. A single Gaussian profile was fitted to the residual to measure the MgII line shape. The derived MgII redshift of $6.5634^{+0.0013}_{-0.0012}$ is consistent with the Lyα redshift reported by Matsuoka et al. (2018a, $z = 6.56$). The redshift measurement is also consistent with the recent [CII] redshift reported by Yang et al. (2021) for this object ($z = 6.5646 \pm 0.0003$).

The BH mass and Eddington ratio were derived based on the MgII single-epoch method (Vestergaard & Osmer 2009). From the MgII line FWHM = 1699^{+99}_{-110} km s^{-1} and the 3000 $\mathring{A}$ luminosity, we measured the virial black hole mass $M_{\rm BH} = (2.48^{+0.31}_{-0.29}) \times 10^8 M_\odot$ and an Eddington ratio $\lambda = L_{\rm bol}/L_{\rm Edd} = L_{\rm bol}/(4\pi\, c\, G\, M_{\rm BH} m_{\rm p}/\sigma_{\rm T}) = 2.29^{+0.30}_{-0.29}$. J0921+0007 shows a super-Eddington ratio, as expected from its steep photon index.

We note that J0921+0007 was also observed with Gemini/GNIRS by Yang et al. (2021). Their continuum and MgII measurements are mostly consistent with ours, while they report a slightly fainter 3000 angstrom luminosity ($L_{3000} = (3.9 \pm 0.4) \times 10^{42}$ erg s^{-1}). This difference is likely attributed to the different absolute flux calibration between the GNIRS and MOIRCS spectra. For the calibration of their spectrum, Yang et al. (2021) used a J-band magnitude measurement with a relatively large error ($J = 21.21 \pm 0.28$) for their calibration.

5.4.2 NLS1 classification

Following Osterbrock & Dahari (1983) and Goodrich (1989), an AGN is required to show a narrow Hβ line (FWHM$_{\rm H\beta} < 2000\,{\rm km\,s^{-1}}$) and a small narrow-line to broad-line flux ratio ([OIII]/H$\beta < 3$) in order to be classified as NLS1s. At $z > 6$, however, Hβ is shifted out to mid-infrared wavelengths, making the direct classification of high-z quasars as NLS1s by this definition impossible. Using the correlation between MgII and Hβ widths in a sample of Sloan Digital Sky Survey DR14 quasars, Rakshit et al. (2021) proposed selecting NLS1s using FWHM$_{\rm MgII} < 2000\,{\rm km\,s^{-1}}$. With an MgII width FWHM$_{\rm MgII} = 1699^{+99}_{-110}$ km s^{-1}, J0921+0007 can be classified according to this selection criterion. Rakshit et al. (2021) propose a UV proxy measurement of the ratio of FeII and Hβ equivalent widths $r_{\rm FeII}$ (Boroson & Green 1992) as $r_{\rm FeII,UV} = {\rm EW(FeII)/EW(MgII)}$. Together with the FWHM of Hβ, $r_{\rm FeII}$ is one of the parameters defining the quasar main sequence (e.g. see Marziani et al. 2018, and references therein). $r_{\rm FeII}$ has been shown to correlate tightly with the Eddington ratio (e.g. Rakshit et al. 2017). We computed the equivalent width of the fitted FeII template in the rest-frame wavelength interval $2200 - 3090\,\mathring{A}$. We note that this measurement is based on an extrapolation of the iron template, since the spectrum only

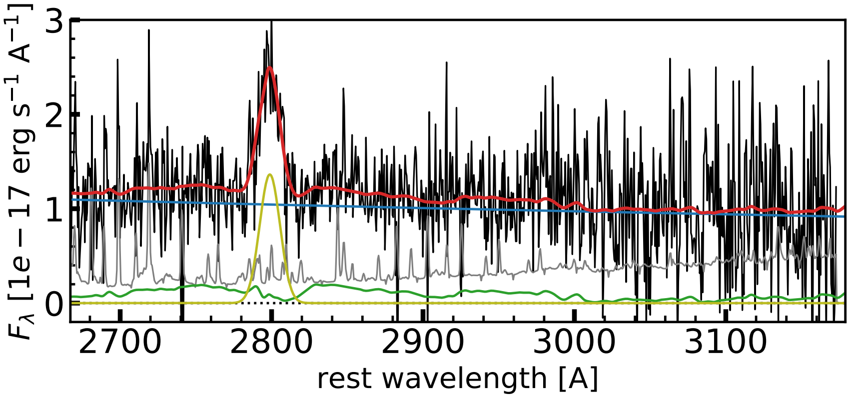


Figure 5.5: *Ks*-band MOIRCS spectrum of J0921+0007. The combined fit to the continuum and MgII line is shown by the solid red curve. The decomposition in continuum (blue), the FeII pseudo continuum (green), and the Gaussian line fit (yellow) are also shown.

covers rest frame $\sim 2700 - 3150\,\mathring{A}$. We obtained $r_{\mathrm{FeII,UV}} \sim 3.77$, a value locating the source at the strongly accreting end of the quasar main sequence (Population A see Marziani et al. 2018; Marziani et al. 2018).

5.5 Measuring the size of the proximity zone with an optical spectrum

Lyα proximity zones are ionised regions along the line of sight that are transparent to the quasar flux bluewards of the Lyα line. The surrounding inter-galactic medium (IGM) is thought to have been ionised by the UV radiation emitted by the quasar at the centre. Considering a discrete ionised HII region expanding in a neutral and uniform IGM, Haiman & Cen (2001) related the radius of the quasar proximity zone (or Stromgren spheres) to the emission rate of ionising photons, $\dot{N}_{ion}$, the mean neutral hydrogen density in the IGM, n_{HI}, and the lifetime of the quasar, t_{q}:

$$R_{\mathrm{HII}} = \left(\frac{3\dot{N}_{\mathrm{ion}}\, t_{\mathrm{q}}}{4\pi n_{\mathrm{HI}}} \right)^{1/3} . \tag{5.2}$$

The radius of the proximity zone of J0921+0007 has already been measured to be $R_{\mathrm{p}} = 3.05 \pm 0.45\,\mathrm{pMpc}$ by Ishimoto et al. (2020) in the context of the SHELLQs survey (proper distance in Mpc). They used the low signal-to-noise ratio (S/N) discovery spectrum of the quasar taken with the Optical System for Imaging and low-Intermediate Resolution Integrated Spectroscopy (OSIRIS) at the 10.4 m Gran Telescopio Canarias (GTC; 0.9 ks Matsuoka et al. 2018a). With the availability of precise [CII] redshift measurements from Yang et al. (2021), $z = 6.5646 \pm 0.0003$, we could improve upon this measurement. We re-observed J0921+0007 on February 28, 2022, with the LDSS3-C spectrograph mounted at the Magellan-Clay Telescope of the Las Campanas Observatory (Chile). A long-slit spectrum was obtained with the VPH-Red grism, which covers a range of about 6000-10500 Å with a dispersion of about 1.16 Å/px. A $1''$ slit was used, which allowed a spectral resolution of about 4.7 Å to be reached. An exposure time of 3 $\times 1200\,\mathrm{s}$ was applied to effectively remove cosmic rays. The seeing was around $0.6\text{-}0.7''$. We reduced the spectra with IRAF following the classic procedure of overscan subtraction, flat-field correction, and wavelength calibration. The standard star LTT3864 was observed with the same aperture slit to perform the flux calibration. Finally, the three exposures were sky-subtracted and averaged.

The radius of the proximity zone was measured following the methodology of Fan et al. (2006b). The spectrum was normalised by a model for the continuum and smoothed by convolving a boxcar function of size $20\,\mathring{A}$ with the signal. The edge of the proximity zone was then set to be the wavelength at which the continuum-normalised flux bluewards of Lyman α first drops below 10% of the extrapolated model. In practice, the wavelength λ_{edge} is found as the wavelength at which the first of three consecutive pixels of the smoothed spectrum are below this threshold (Eilers et al. 2017).

The continuum bluewards of the Lyman α line is strongly affected by absorption and needs to be reconstructed. This can be achieved by performing a principal component analysis on the continua of lower-redshift quasars (Francis et al. 1992; Suzuki et al. 2005; Pâris et al. 2011). Following this method, quasar spectra q_{mod} were modelled as the sum of an average spectrum and projections along principal components as

$$q_{\mathrm{mod}}(\lambda) \sim \mu(\lambda) + \sum_{j=1}^{m} c_j \xi_j(\lambda), \tag{5.3}$$

where μ is an average quasar spectrum , ξ_j is the j-th principal component and c_j a weight specific to this quasar. In order to reconstruct the blue side of the spectrum ($\lambda < 1216\mathring{A}$) from the red side ($\lambda > 1216\mathring{A}$) for high-redshift quasars, two sets of principal components and associated weights were derived from lower-redshift training samples, one for the full probed wavelength range (e.g. $1020\,\mathring{A} - 2000\,\mathring{A}$) and one only for the red range. Projections from red-range weights to full-range weights were then derived.

While Eilers et al. (2017) use a mean quasar spectrum and principal component projections derived by Pâris et al. (2011), Ishimoto et al. (2020) used results by Suzuki et al. (2005), since their principal components are derived from fainter quasars, which better represent the SHELLQs quasars. We refer to these work for more details of the continuum reconstruction technique. For our quasar spectrum, we measured the proximity zone using projections from Pâris et al. (2011). The radius of the proximity zone is then: $R_{\mathrm{p}} = (d_{\mathrm{q}} - d_{\mathrm{edge}})/(1+z_{\mathrm{q}})$, where z_{q} is the quasar redshift, d_{q} and d_{edge} are the comoving distances derived from z_{q} and the redshift of the edge of the proximity zone: $z_{\mathrm{edge}} = \lambda_{\mathrm{edge}}/1215.67\mathring{A} - 1$. Accounting solely for the uncertainty in the [CII] redshift reported by Yang et al. (2021), we obtained $R_{\mathrm{p}} = 3.05 \pm 0.01\,\mathrm{Mpc}$, a result consistent with the measurement of Ishimoto et al. (2020). In addition, we can account for systematic uncertainties on the [CII] redshift by applying a conservative offset $\delta_v = 100\,\mathrm{km\,s^{-1}}$ in quadrature (Eilers et al. 2020). At $z = 6.56$, such a velocity offset results in an additional redshift systematic of $\sigma_{z,\mathrm{syst}} = 0.0025$. We obtained $R_\mathrm{p} = 3.05 \pm 0.13\,\mathrm{pMpc}$.

Since the size of the proximity zone is expected to depend on the quasar luminosity, tracing the redshift evolution of R_p usually requires its measurements to be corrected to a common scale, that is, normalised to the same absolute magnitude at 1450 Å, M_{1450}. Following the relation by Eilers et al. (2017),

$$R_\mathrm{p,corr} = R_\mathrm{p} \times 10^{0.4(27+M_{1450})/2.35}, \tag{5.4}$$

we obtained $R_\mathrm{p,corr} = 5.48 \pm 0.72\mathrm{pMpc}$. Here we have used $M_{1450} = -25.55 \pm 0.23$, as derived from the NIR spectral slope (see Section 5.4.1)[3]. The evolution of the size of quasar proximity zones with redshift has been extensively investigated (e.g. Carilli et al. 2010; Venemans et al. 2015; Mazzucchelli et al. 2017; Eilers et al. 2017; Davies et al. 2020). From a sample of $z < 6.6$ quasars, Eilers et al. (2017) recover a relatively shallow redshift evolution:

$$R_\mathrm{p,corr} \approx 4.86\,\mathrm{pMpc} \times \left(\frac{1+z}{7}\right)^{-1.44}. \tag{5.5}$$

At $z = 6.56$, the average luminosity-corrected proximity zone radius is $R_\mathrm{p,corr} \approx 4.34\mathrm{pMpc}$. Similarly, using the luminosity scaling of Ishimoto et al. (2020),

$$R_\mathrm{p,corr,-25} = R_\mathrm{p} \times 10^{0.4(25+M_{1450})/1.80}, \tag{5.6}$$

we obtained a corrected proximity zone radius $R_\mathrm{p,corr,-25} = 2.36 \pm 0.40\,\mathrm{pMpc}$, which is larger than their prediction of the best-fit $R_\mathrm{p,corr,-25} - z$ relation: $R_\mathrm{p,corr,-25} = 1.37 \pm 0.23\,\mathrm{pMpc}$.

The luminosity-scaled proximity zone radius of J0921+0007 is thus relatively large with respect to the bulk of the high-redshift quasar population. Indeed, as can be seen from the unsmoothed spectrum presented in Fig. 5.6, the profile of the strong Lyman α line appears double-peaked with a sharp absorption feature at the exact wavelength of Lyman α. The strong transmission bluewards of the absorption edge indicates that the quasar is embedded in a large and completely ionised region of the IGM. The unabsorbed blue wing of the Lyman α drives the overall line luminosity reported by Matsuoka et al. (2018a).

Uncertainty on the absorbed quasar continuum can significantly affect the measurement of the proximity zone. Performing Monte Carlo simulations, Ishimoto et al. (2020) show that the low S/N causes an uncertainty (0.71 pMpc) in their R_p measurement that is larger than the uncertainty due to the redshift error. We estimated the uncertainty in R_p imprinted by the uncertainty on the flux measurement in the red part of the spectrum. For this, we generated 1000 spectra by perturbing the red part of the LDSS3 spectrum, at wavelengths in the observed frame at $< 1216\mathring{A} \times (1+z)$. We applied Gaussian random noise accounting for the RMS of the spectrum at 9000 $\mathring{A}$, $\sigma = 2.285 \times 10^{-18}\,\mathrm{erg\,s^{-1}cm^{-2}}$.

---

[3] Assuming a steeper power-law slope of $\alpha = -1.5$, Ishimoto et al. (2020) adopted a brighter value $M_{1450} = -26.16 \pm 0.29$.

This is a conservative estimate of the spectral noise since the higher transmission close to the Lyman break is expected to yield lower flux uncertainties. For each of these simulated spectra, we computed the radius of the proximity zone and obtained $R_{\rm p} = (3.00^{+0.49}_{-0.58})$ pMpc. The error related to the uncertainties bluewards of the Lyman α line therefore dominates uncertainties due to the redshift estimate. We stress the necessity of a higher S/N spectrum to confirm the large size of the proximity zone. We note that the LDSS3 spectrum shows no evolution in the spectral shape of Lyα with respect to the OSIRIS discovery spectrum presented by Matsuoka et al. (2018a). We repeated the measurement of the proximity zone radius of J0921+0007, using the OSIRIS discovery spectrum, accounting for the flux uncertainties redwards of the Lyα emission line. We obtained $R_{\rm p} = 3.63^{+1.00}_{-0.44}$ pMpc. While this result shows larger uncertainties, it is consistent within 1σ with the measurement on our new LDSS3 spectrum.

The relatively large proximity zone can have two origins. The first hypothesis is that the active phase of the quasar exceeds the typical lifetime of quasars $t_{\rm q} \sim 10^6$yr (e.g. Khrykin et al. 2016; Davies et al. 2019; Khrykin et al. 2021; Eilers et al. 2021). This interpretation is inconsistent with the idea that the low black hole mass of J0921+0007 is indicative of a young SMBH. The luminosity-scaled proximity zone radius $R_{\rm p,corr,-25} = 2.36 \pm 0.40$ pMpc is relatively large for the black hole mass measured in this Chapter (Ishimoto et al. 2020, see their Fig. 8). Alternatively the quasar has alternated between highly luminous phases and quiescent phases. The proximity zone could have grown during extremely luminous phases of the AGN. Such large amplitude variability in the ionising continuum emission is indeed expected in NLS1s (Collier et al. 2001; Romano et al. 2002). The SED of NLS1s is also potentially quite different to that of broad-line quasars, which can further affect the ionisation balance in the proximity zone. J0921+0007 is the second most Lyman α luminous quasar of the SHELLQs survey (Matsuoka et al. 2016, 2018a; Matsuoka et al. 2019; Matsuoka et al. 2022), which implies that the broad-line region gas is exposed to strong ionising radiation. This is also confirmed by its unusually high X-ray luminosity (Fig. 5.4 and 5.7). We therefore propose that the quasar is currently undergoing a super-Eddington accretion phase that generates powerful UV/X-ray radiation, which in turn ionises the surrounding IGM efficiently. It is in such phases that the proximity zone grows to relatively large radii. The ionising continuum radiation of NLS1s is, however, expected to show large amplitude short-term variability. The response of non-equilibrium blinking light-bulb quasar models has been studied by Davies et al. (2020); however, the timescales on which NLS1s are expected to show X-ray/UV variability (hours to days) are much smaller than the shortest timescales explored here (100 yr). Capturing X-ray variability on timescales of days for this quasar would require a long monitoring campaign ($>$ 180 ks; e.g. Nanni et al. 2018; Yang et al. 2022). The confirmation of short timescale X-ray variability could explain the relatively large proximity zone of J0921+0007. This could imply phases of even more extreme X-ray loudness. In addition, a long exposure would put more stringent constraints on the spectral shape of the source.

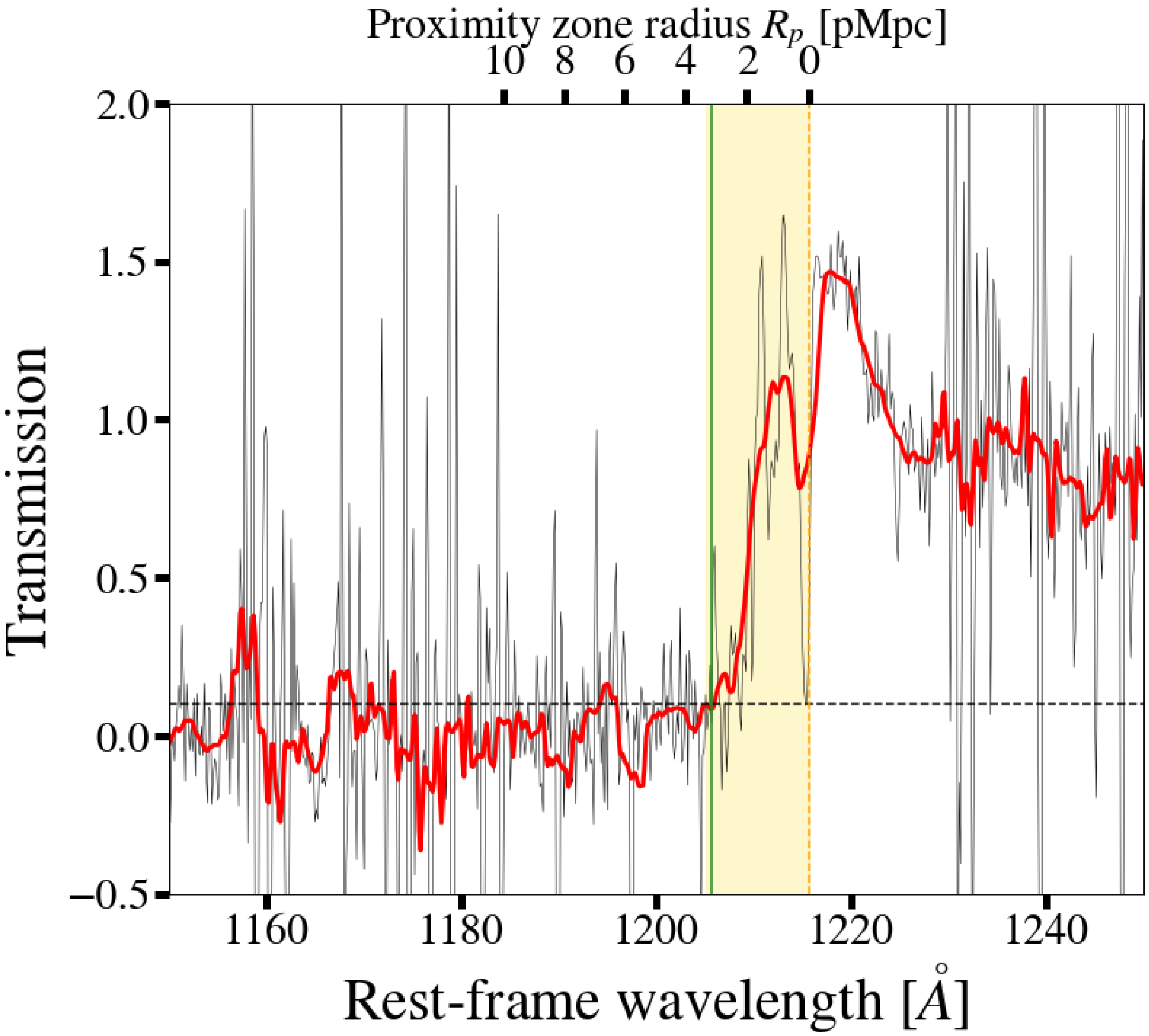


Figure 5.6: Continuum-normalised LDSS3 spectrum of J0921+0007. The size of the proximity zone is defined here as the proper distance between the rest-frame wavelength of $Ly\alpha$ (dashed orange line) and the wavelength where the smoothed continuum normalised flux (solid red line) first drops below 0.1% of its extrapolated value, marked here by a solid green line. The unsmoothed continuum-normalised spectrum is shown in black.

5.6 Relative X-ray and optical/UV output

5.6.1 X-ray loudness

As previously shown in Eq. 4.1, the relative output of the hot corona and the UV disc emission is characterised by the ratio of monochromatic luminosities at 2 keV and 2500 $\AA$, $\alpha_{\rm OX}$.

There is a well-known anti-correlation between $\alpha_{\rm OX}$ and the 2500 $\AA$ monochromatic luminosity (e.g. Steffen et al. 2006; Just et al. 2007; Lusso & Risaliti 2016), which signifies that more UV luminous quasars tend to show a stronger UV contribution to their total emission (with respect to the coronal emission). This relation does not show evolution with redshift (e.g. Steffen et al. 2006; Just et al. 2007; Lusso & Risaliti 2016; Nanni et al. 2017; Vito et al. 2019a). We computed the $\alpha_{\rm OX}$ by deriving L_{2500} from the monochromatic luminosity of 3000 $\AA$ and the spectral slope reported in Section 5.4.1. The UV luminosity is $L_{2500} = (1.74 \pm 0.07) \times 10^{31}\,{\rm erg\,s^{-1}Hz^{-1}}$. We further obtained an estimate of $L_{2\,{\rm keV}}$ from the 2-10 keV luminosity obtained from the spectral analysis of the *Chandra* data, taking into account the posterior distribution of Γ. We obtained $\alpha_{\rm OX} = -1.21 \pm 0.09$. In Fig. 5.7, we show how this X-ray to UV power-law slope compares to other X-ray-detected high-redshift quasars from Nanni et al. (2017); Vito et al. (2019a); Pons et al. (2020); Medvedev et al. (2021) and Wolf et al. (2021) (see Chapter 4). The $\alpha_{\rm OX}$ relation derived for $z > 5.7$ sources by Nanni et al. (2017) is shown. The eFEDS normalised sensitive area to a fiducial source modelled by an absorbed power law with $\Gamma = 3.2$ and $N_{\rm H} = 3 \times 10^{20}\,{\rm cm^{-2}}$, is represented as background colour gradient. J0921+0007 shows a significantly flatter $\alpha_{\rm OX}$ slope than other high-redshift quasars with comparable UV luminosities. Accounting for the 1σ confidence interval for the $\alpha_{\rm OX} - L_{2500}$ relation of Nanni et al. (2017), the nominal $\alpha_{\rm OX}$ of J0921+0007 is a $> 3\sigma$ outlier. With respect to the more conservative relation of Just et al. (2007), J0921+0007 is a $> 1.5\sigma$ outlier. This indicates that the X-ray contribution to the total emission in this quasar is higher. The eFEDS sensitivity map corroborates the outlier nature in terms of relative X-ray to UV output of J0921+0007: at L_{2500}, we could not have detected this source in eFEDS if it followed the typical $\alpha_{\rm OX} - L_{2500}$ relation. J0921+0007 is X-ray loud with respect to the bulk of X-ray-detected high-redshift quasars. We note that this result relies on the assumption that the steep photon index measured in the rest-frame hard band of J0921+0007, $\Gamma = 3.2$, is not due to any non-coronal components (e.g. soft excess). The object of Medvedev et al. (2020) shows a similarly strong deviation from the $\alpha_{\rm OX} - L_{2500}$ relation, which is due to non-coronal X-ray emission from the jet. J0921+0007 in contrast is radio-quiet. The significant α_{OX} outlier from the Vito et al. (2019a) sample at $\alpha_{OX} = -1.28$ and log$L_{2500} \sim 31$ is the radio-quiet quasar CFHQS J1641+3755 at $z = 6.04$. This source shows remarkable similarities to J0921+0007 in terms of black hole properties and its X-ray emission (see discussion).

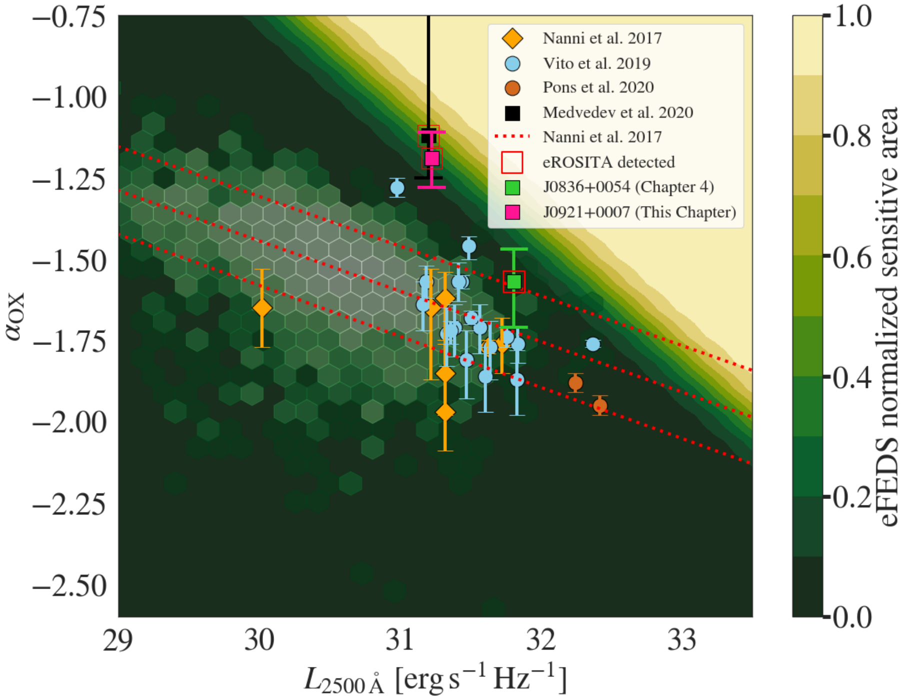


Figure 5.7: X-ray to UV slope $\alpha_{\rm OX}$ as a function of L_{2500} for quasars at $z > 5.7$. Because of its strong X-ray emission, J0921+0007 (pink square) deviates significantly from the $\alpha_{\rm OX}$-L_{2500} relation, the 1σ confidence interval of which is delimited by the dotted red lines (Nanni et al. 2017). The sensitivity of eFEDS to sources modelled by a steep absorbed power law ($\Gamma = 3.2$ and $N_{\rm H} = 3 \times 10^{20}\,{\rm cm}^{-2}$) is traced by the colour map. The hexagonal pattern represents a sample of XMM-detected SDSS quasars presented by Lusso & Risaliti (2016).

5.6.2 An increased coronal contribution to the bolometric luminosity

In the previous section we demonstrate that the quasar shows a stronger X-ray emission than typical quasars of similar rest-frame UV luminosity L_{2500}. This implies that the total AGN emission may be affected by the increased coronal contribution at shorter wavelengths. To account for this, we estimated a corrected AGN bolometric luminosity $L_{\mathrm{bol,corr}}$ by performing an SED fit with the tool Code Investigating GALaxy Emission (`CIGALE`; Boquien et al. 2019) and its X-ray module `X-CIGALE`, (Yang et al. 2020). We obtained photometry from the third public data release of the HSC-SSP, LS8, the VISTA Kilo-degree Infrared Galaxy Survey (VIKING; Arnaboldi et al. 2007), the UKIRT Infrared Deep Sky Survey (UKIDSS; Lawrence et al. 2007), and CatWISE2020 (Marocco et al. 2021), by cross-matching the optical position of the quasar to these catalogues and selecting the nearest detection within $1''$. Each of the above surveys yielded a detection within $1''$. We did not make use of photometry in the absorbed region of the spectrum at $\lambda < 1215\mathring{A}$, since transmitted flux, for example from the quasar proximity zone, can negatively affect the SED fit (see the red squares Fig. 5.8). In addition we used the 0.5-2 keV flux from the spectral fit to the *Chandra* data ($\Gamma \sim 3.2$; see Chandra `BXA` in Table 5.2).

We fitted an AGN disc model as defined in the AGN `SKIRTOR` module, which uses a disc spectral energy distribution (SED) adapted from Feltre et al. (2012, see also Duras et al. 2017; Yang et al. 2020). We fixed the viewing angle with respect to the AGN axis to $i = 30°$ and probed a grid of E(B-V) values for extinction in the polar direction. The AGN fraction was set to 0.999, as the host galaxy emission is completely dominated by the AGN emission in luminous, distant quasars. We further allowed a large dispersion in the α_{OX}-L_{2500} relation: $\Delta\alpha_{\mathrm{OX}} = |\alpha_{\mathrm{OX}} - \alpha_{\mathrm{OX}}(L_{2500})| = 1.0$, where $\alpha_{\mathrm{OX}}(L_{2500})$ was determined using the relation measured by Just et al. (2007). We fixed the power-law photon index of the *xray* module to $\Gamma = 3.2$, a value supported by our X-ray spectral fitting results. We used fiducial galaxy population parameters set in the modules *sfhdelayed* (delayed star-formation history) and *bc03* (single stellar populations, Bruzual & Charlot 2003). Galactic dust attenuation was accounted for via *dustatt_calzleit* (Calzetti et al. 2000; Leitherer et al. 2002). The redshift was fixed to $z = 6.56$. The best-fitting model has a reduced χ^2 of 0.72. The AGN disc dominates the rest-frame optical/UV part of the SED. The AGN dust emission remains unconstrained. The SED as well as the total model are presented in Fig. 5.8.

We performed sanity checks by comparing the 2-10 X-ray luminosity $L_{2-10\,\mathrm{keV,XCIG}}$ and the $\alpha_{\mathrm{OX,XCIG}}$ from the X-CIGALE output to the measurements from Sects. 5.3.3 and 5.4.1. The values are : $L_{2-10\,\mathrm{keV,XCIG}} = (3.04 \pm 2.72) \times 10^{45}\,\mathrm{erg\,s^{-1}}$ and $\alpha_{\mathrm{OX,XCIG}} = -1.2$. These results are consistent within 1σ with the more precise measurements from the X-ray and infrared spectral analysis.

The corrected bolometric AGN luminosity was obtained by summing the intrinsic AGN disc luminosity averaged over all directions (*agn.accretion_power*) and the total X-ray luminosity (0.2-100 keV): $L_{\mathrm{bol,corr}} = L_{\mathrm{disc}} + L_{\mathrm{X,tot}} = (9.9 \pm 4.0) \times 10^{46}\,\mathrm{erg\,s^{-1}}$. We did not

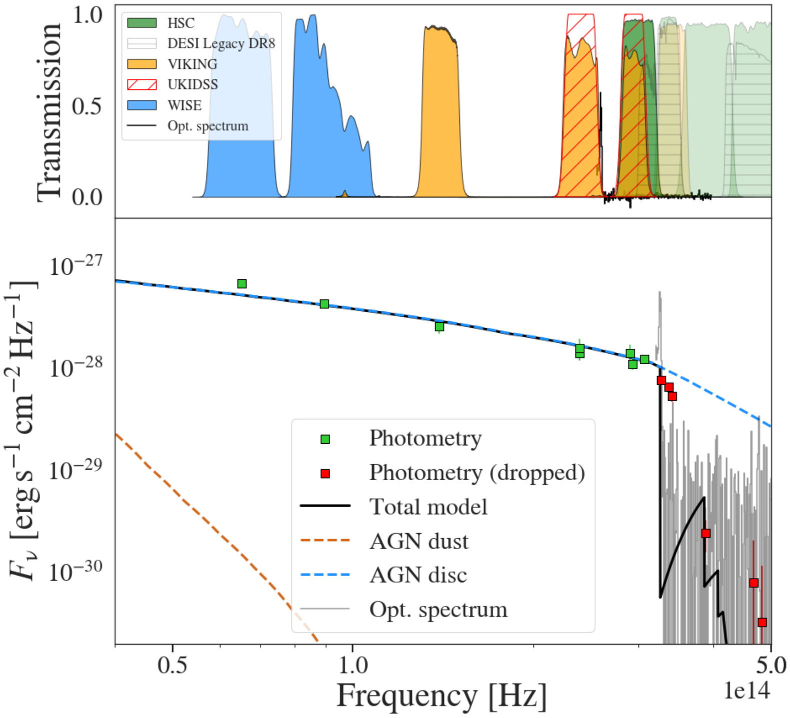


Figure 5.8: X-CIGALE fit to the quasar optical, infrared, and X-ray photometry. Photometric points from HSC, LS8, UKIDSS, VIKING, and unWISE are marked by squares. The photometry that has been ignored for the SED fit is shown as red squares. The solid black line is the best-fitting model. The dashed blue line shows the fitted disc model and the dashed brown line the unconstrained torus emission. We also show the LDSS3 spectrum presented in Appendix 5.5. The upper panel shows corresponding photometric filters.

include the dust emission in this calculation since it arises from re-processed nuclear UV and X-ray photons (e.g. Lusso et al. 2012; Duras et al. 2020).

We compared the resulting $2 - 10\,\mathrm{keV}$ bolometric correction, $K_{\mathrm{bol}} = L_{\mathrm{bol}}/L_{2-10\,\mathrm{keV}}$, to typical type 1 AGN values from literature. J0921+0007 follows a bolometric correction $K_{\mathrm{bol}} = 23^{+29}_{-13}$, about a factor of 4 smaller than the prediction from the general $K_{\mathrm{bol}} - L_{\mathrm{bol}}$ relation of Duras et al. (2020) and the one for type 1 AGN by Lusso et al. (2012), as can be seen in Fig. 5.9. These findings further confirm the unusual X-ray loudness of the quasar. Using the same method, we also derived a bolometric correction for the second high-redshift quasar detected in eFEDS, SDSS J0836+0054, from the photometry presented in Chapter 4. This source also shows X-ray emission excess.

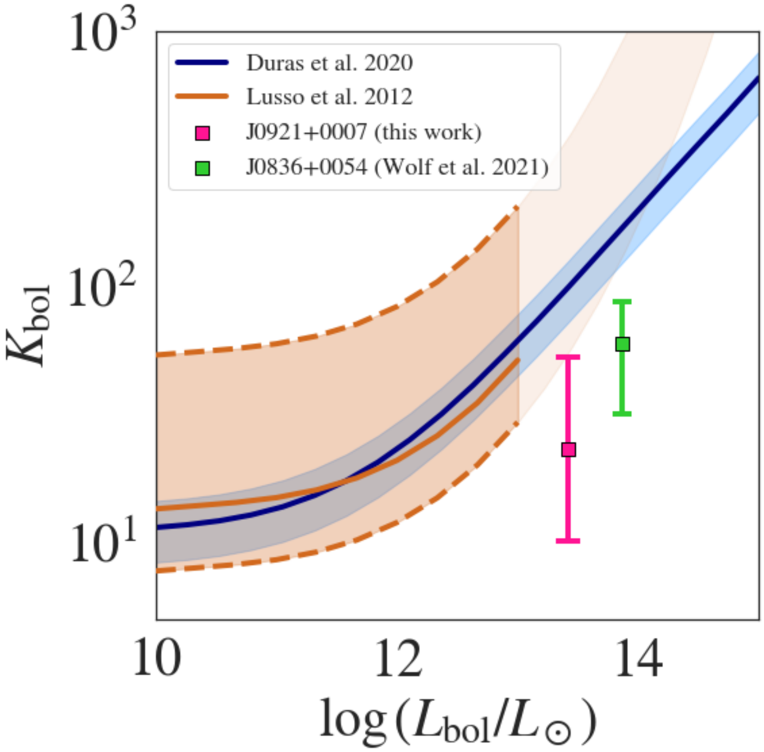


Figure 5.9: Bolometric correction from the 2-10 keV band for J0921+0007 (pink error bars) compared to predictions from Duras et al. (2020) for the full AGN population (solid blue line) and Lusso et al. (2012) for type 1 AGN (solid brown line). The dispersions are quoted at the 68% level (blue shaded area and brown shaded area). The strong X-ray emission of J0921+0007 with respect to its disc emission causes it to deviate at the $> 1\sigma$ level from the typical AGN $K_{\rm bol} - L_{\rm bol}$. The green error bars show the bolometric correction for the second eFEDS high-z quasar presented in Chapter 4. We note that the bolometric correction of Lusso et al. (2012) is only valid up to $\log L_{\rm bol} = 13$; the extrapolation (light-shaded area) is shown for visualisation purposes only.

5.7 AGN demographics in the first gigayear of the Universe

5.7.1 Comparison to XLF models

X-ray-selected quasars can be used to trace black hole accretion through cosmic time via the XLF. The AGN XLF has been studied in detail over a variety of populations in deep and/or wide surveys (Hasinger et al. 2005; Barger et al. 2005; Vito et al. 2014; Ueda et al. 2014; Buchner et al. 2015; Aird et al. 2015; Georgakakis et al. 2015; Fotopoulou et al. 2016; Ananna et al. 2019). In these works, the space density of X-ray-detected AGN has been investigated up to $z = 5$ and has been shown to decline exponentially with redshift beyond the luminosity at the knee of the XLF $\log L_* \sim 44$ (Brusa et al. 2009; Civano et al. 2011). Selecting high-redshift AGN in the *Chandra* Deep Fields, Vito et al. (2018) extended the analysis out to $z = 6$, with a sample of sources with photometric redshifts beyond $z \geq 5.5$. In Chapter 4, we showed that the eROSITA detection of a spectroscopically confirmed $z = 5.81$ quasar SDSS J0836+0054 in the eFEDS field imposes new constraints on the AGN XLF at its bright end. XLF models with a milder decline in AGN space density in the highest X-ray luminosity bins are favoured by this detection. These findings are corroborated by an analysis of a sample of X-ray-detected high-redshift AGN in the ExSeSS by Barlow-Hall et al. (2022). They also report the detection of ATLAS J025.6821-33.4627, a spectroscopically confirmed quasar at $z = 6.31$, which had been, until the detection of J0921+0007 in eFEDS, the highest-redshift, blindly X-ray-detected AGN.

In Chapter 4, we show evidence against the steepest declines of the space density of luminous X-ray-detected AGN with increasing redshift. Here we go beyond this claim and show that the detection of J0921+0007 in the field is not supported by predictions of current XLF models over all luminosities extrapolated to higher redshifts. We compare the number count predictions from the best-fitting, extrapolated XLF models presented by (Ueda et al. 2014, luminosity-dependent density evolution), Vito et al. (2014, pure density evolution), Georgakakis et al. (2015, pure density evolution), Miyaji et al. (2015, luminosity-dependent density evolution) and Aird et al. (2015, flexible double power law), to the eFEDS detections. We stress that these models were evaluated on AGN samples at $z < 5$ and that our comparison assumes that the parametric form of the XLF derived by these authors does not strongly evolve from $z = 5$ to $z \sim 6$. Following Eq.5.7, we obtained number counts beyond a given redshift and luminosity threshold, $z_{\min}$ and $L_{\min}$, from the XLF models by computing

$$N = \int_{\log L_{\min}}^{\infty} \int_{z_{\min}}^{z=10} A_{\Gamma}(log\, L_{\mathrm{X}}, z) \frac{\mathrm{d}V}{\mathrm{d}z} \phi_m(\theta)\, \mathrm{d}\, z\, \mathrm{d}\log L_{\mathrm{X}}. \tag{5.7}$$

$A_{\Gamma}(log\, L_{\mathrm{X}}, z)$ is the normalised sensitive area of the survey to a source of luminosity L_X and redshift z, $\frac{\mathrm{d}V}{\mathrm{d}z}$ the differential comoving volume and $\phi_m(\theta)$ the model XLF (in units of Mpc^{-3}). As in Chapter 4, the sensitive area was obtained by converting the `apetool` (Georgakakis et al. 2008) count-rate based area curve to a luminosity-based area curve

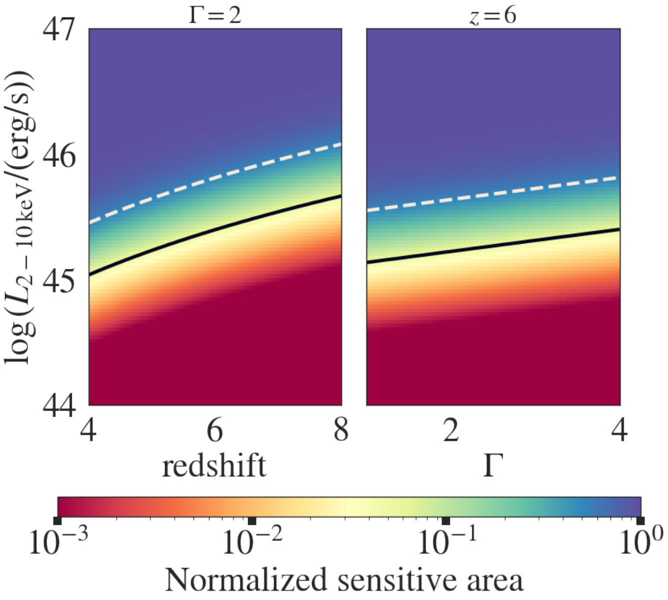


Figure 5.10: Slices of the simulated eFEDS sensitive area cube. *Left:* Sensitive area for fixed $\Gamma = 2$ as a function of $L_{2-10\,\mathrm{keV}}$ and z. *Right:* Sensitive area as a function of Γ and $L_{2-10\,\mathrm{keV}}$ for fixed redshift $z = 6$. The dotted white (solid black) line shows the parameter configurations at 50% (5%) of the normalised sensitive area.

with an X-ray spectral model. Here, we assumed a redshifted power law under Galactic absorption, (*tbabs*zpowerlw*). Following this methodology, redshift-luminosity configurations as a function of the photon index were converted into soft band counts (and thereby to a normalised sensitive area) by generating X-ray spectra using the convolutional `XSPEC` component *clumin*. This conversion was fitted with logistic regression. A novelty is that we fitted the luminosity-counts relation over a grid of Γ. Slices of the resulting fitted surface are shown in Fig. 5.10. In Eq. 5.7, we can now evaluate the model-dependent sensitivity using this function. We integrated over all luminosities by setting $\log L_{\mathrm{min}} = 42$.

For each of the XLF models, we evaluated Eq. 5.7 over the range $z_{\mathrm{min}} = 5.5 - 7$. We accounted for two spectral models by estimating the sensitive survey area $A(\log L_{\mathrm{X}}, z)$ for $\Gamma = 2.2$ (Vito et al. 2019a) and $\Gamma = 3.2$, the median value derived from the *Chandra* observation of J0921+0007. This effectively yields the expected number counts over all luminosities beyond an increasing redshift threshold z_{min}. For each model and each z_{min} we computed the 15.9-th and 84.1-th percentile of the count expectations to estimate their 1σ confidence intervals. The resulting inverse cumulative distribution of predicted counts is presented in Fig. 5.11. We obtained 1σ confidence intervals by sampling from the parameter uncertainties of the models. As the covariance matrix of these parameters was not accounted for, the uncertainty may have been over-estimated. The model predictions are compared to the source counts detected in eFEDS (Fig. 5.11). We can conclude that over all luminosities, none of the extrapolated XLF models supports two detections in the eFEDS field at high redshift, regardless of the X-ray spectral model. At $z > 5.81$, the models from Vito et al. (2014), Ueda et al. (2014), Miyaji et al. (2015) and Georgakakis et al. (2015) are consistent with one detection to within 1σ. No model supports a detection at $z = 6.56$. This is further shown in the 2 lower panels of Fig. 5.11, where we present for each lower-redshift interval edge, $z_{\rm min}$, the Poisson probability for each XLF model of supporting at least one ($P_\lambda(\geq 1, z)$) or two counts ($P_\lambda(\geq 2, z)$)in the eFEDS field. Even for the model with the highest number count prediction at $z \geq 5.81$ (Vito et al. 2014) the Poisson probability of detecting two sources in eFEDS is less than 0.4. For all models, the probability of generating 1 count at $z \geq 6.56$ is less than 0.5. As expected for sources with steeper photon indices, the count expectations are lower. In particular, a detection at $z \geq 6.56$ is unlikely with $P_\lambda(\geq 1, z \geq 6.56) < 0.25$. As discussed in Chapter 4, the volume probed by the eFEDS survey at $z = 5.81 - 6.56$ is such that the uncertainty due to cosmic variance in the expected source counts is negligible with respect to the Poisson error. We conclude that the current extrapolated XLF models under-predict the number of high-redshift quasars that we detect in eFEDS.

5.7.2 Contribution of X-ray-luminous quasars to the accretion density at $z \sim 6$

The total black hole mass accretion rate per unit volume can be traced through cosmic time via the black hole accretion rate density (BHAD), $\Psi_{\rm bhar}(z)$, which is related to the AGN bolometric luminosity function $\phi(L_{\rm bol}, z)$, as

$$\Psi_{\rm bhar}(z) = \int \frac{1-\epsilon}{\epsilon c^2} L_{\rm bol}\, \phi(L_{\rm bol}, z) {\rm d}\log L_{\rm bol}, \tag{5.8}$$

where $L_{\rm bol}$ is the bolometric luminosity and ϵ the radiative efficiency. The radiative efficiency ϵ is related to accretion efficiency η and the physics governing the accretion flow. We assume that the X-ray-luminous eFEDS quasars accrete above the critical rate, which delimits a radiatively efficient accretion disc from an inefficient one. We adopted a standard thin disc estimate $\epsilon = \eta = 0.1$ (e.g. Soltan 1982; Fabian & Iwasawa 1999; Merloni & Heinz 2008; Delvecchio et al. 2014), that is, the radiative efficiency is set equal to the accretion efficiency. We note that the value of ϵ only affects the normalisation of $\Psi_{\rm bhar}$. At higher redshifts, the evolution of the BHAD has been derived from X-ray-detected (Aird et al. 2015; Vito et al. 2018) and X-ray-undetected, stacked AGN (Vito et al. 2016). To estimate the total contribution of the eFEDS-detected quasars, we first note that

$$\Psi_{\rm bhar}(z) = \frac{1-\epsilon K_{\rm bol}}{\epsilon c^2} \int L_{2-10\,{\rm keV}}\, \phi_X(L_{2-10\,{\rm keV}}, z) {\rm d}\log L_{2-10\,{\rm keV}}, \tag{5.9}$$

where $L_{2-10\,{\rm keV}}$ is the 2-10 keV luminosity, $\phi_X(L_{2-10\,{\rm keV}}, z)$ the hard XLF and $K_{\rm bol}$ the bolometric correction from the 2-10 keV band. The integral is the total AGN emissivity per unit volume. Therefore, we can rewrite Eq. 5 as

$$\Psi_{\rm bhar}(z) \approx \frac{1-\epsilon}{\epsilon c^2} \times \sum_{\Delta z} \frac{L_{\rm bol,AGN}}{V_{\rm eFEDS,\Gamma}}, \tag{5.10}$$

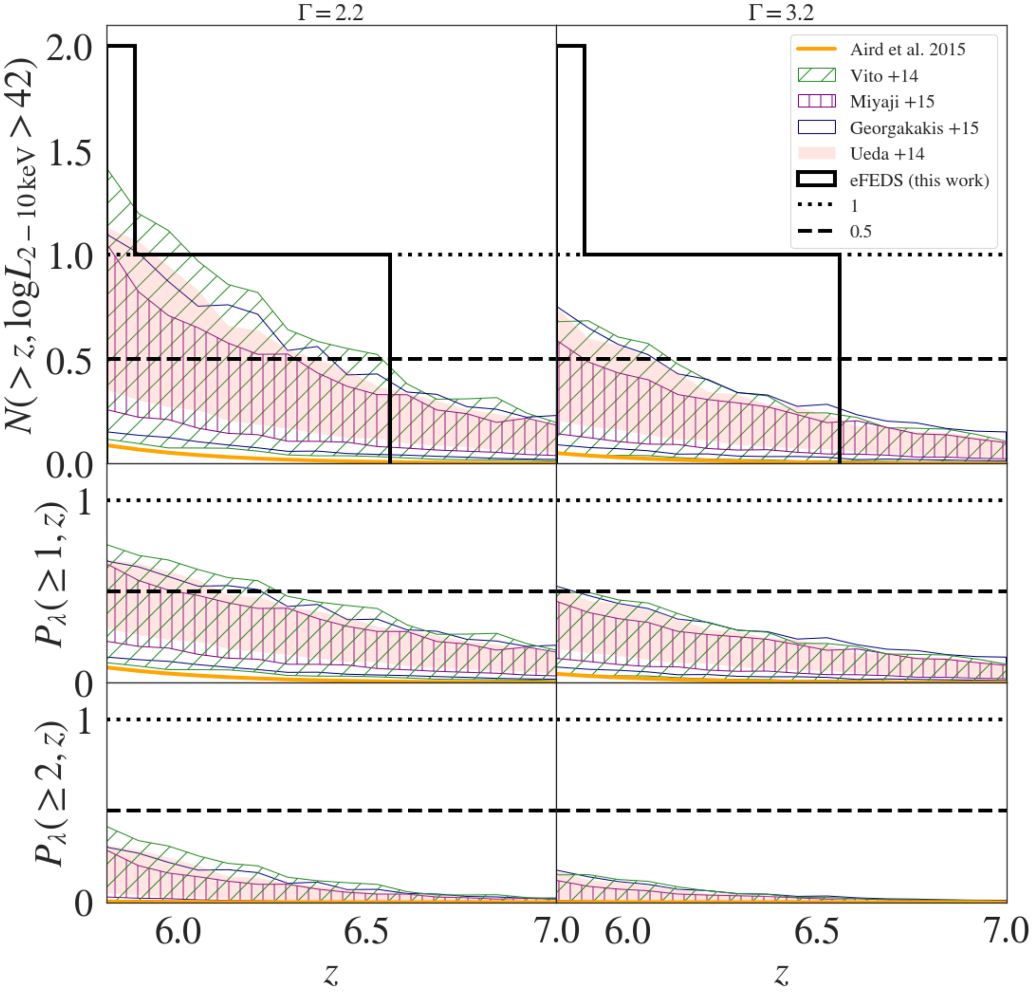


Figure 5.11: Inverse cumulative source count predictions in eFEDS from extrapolated XLF models from Vito et al. (2014), Ueda et al. (2014), Aird et al. (2015), Miyaji et al. (2015) and Georgakakis et al. (2015) integrated over all luminosities. The shaded and hashed areas show the 1σ confidence intervals derived from the model parameter uncertainties. These count predictions depend on the sensitive eFEDS area (Eq. 5.7) and therefore on the assumed spectral model for the AGN. We show the predictions for two different photon indices, $\Gamma = 2.2$ and $\Gamma = 3.2$, respectively in the left and right panels. The black line shows the inverse cumulative distribution of detection in eFEDS. No model supports two detections in eFEDS for the chosen photon indices. At $\Gamma = 2.2$, the Vito et al. (2014), Ueda et al. (2014), Miyaji et al. (2015) and Georgakakis et al. (2015) models support the unique detection at $z \geq 5.81$. The central and lower panel present the Poisson probabilities of the XLF models supporting respectively one and two detections in eFEDS beyond a given redshift threshold. These probabilities are overall low and demonstrate the discrepancy between the eFEDS counts and the model predictions. With $\Gamma = 3.2$ the detection probabilities are lower than with $\Gamma = 2.2$.

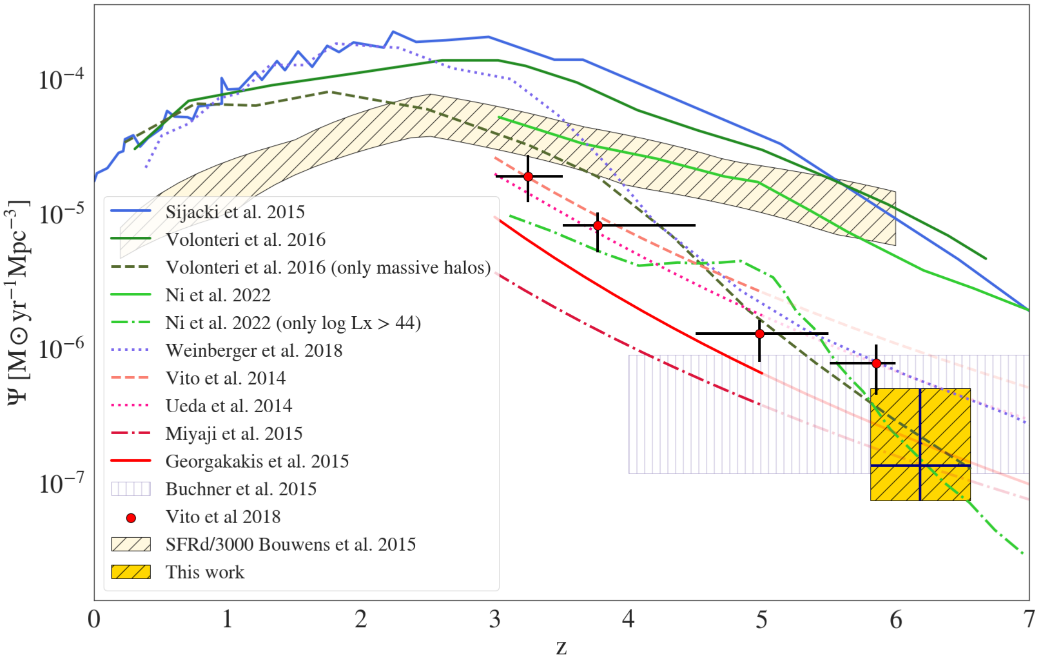


Figure 5.12: Black hole accretion rate density (BHAD) for various XLF models and growth simulations. XLF models from Vito et al. (2014), Ueda et al. (2014), Miyaji et al. (2015) and Georgakakis et al. (2015) are extrapolated beyond $z = 5$ (pale continuation of the red BHAD curves), and the prediction from Buchner et al. (2015) is given over the full $z = 4 - 7$ range. We include observational results from Vito et al. (2018) from the *Chandra* Deep Fields. The measurement derived from the high-redshift quasar detections in eFEDS is shown as a yellow square. Our result is consistent with theoretical predictions restricted to the highest halo (and black hole) masses (Volonteri et al. 2016; Ni et al. 2022). For comparison, a scaled version of the star formation rate density from Bouwens et al. (2015) is shown as beige-shaded and dashed area.

where $L_{\mathrm{bol,AGN}}$ the bolometric AGN luminosity of sources detected in the Δz bin and $V_{\mathrm{eFEDS},\Gamma}$ the sensitive comoving volume of the eFEDS survey. The sum is taken over all detections in Δz. We computed this estimator for the redshift interval $z = [5.81-6.56]$, the redshift interval spanned by the two quasars. The contribution of the eFEDS high-z quasars to black hole accretion in this redshift bin can be obtained by summing up the ratio of accretion luminosities obtained in the SED fits presented in Section 5.6.2 (*agn.accretion_power*) to the corresponding eFEDS sensitive volume (Vito et al. 2016). We obtained this volume by accounting for the sensitivity to sources that have 2-10 keV luminosities of the quasars detected in eFEDS (see Fig. 5.10). We assumed $\Gamma = 2.2$ for the sensitive survey area of J0836+0054 (Chapter 4) and $\Gamma = 3.2$ for J0921+0007 (as derived from the spectral fit in Sec 5.3.3). The resulting accretion density is $\Psi_{z=5.81-6.56} = (1.36^{+3.58}_{-0.62}) \times 10^{-7}\mathrm{M}_\odot\,\mathrm{yr}^{-1}\mathrm{Mpc}^{-3}$. We note that this total AGN emissivity per unit comoving volume only accounts for the un-extincted disc luminosity and not the disc photons reprocessed by the torus and the corona.

We compare this result to lower-redshift measurements and theoretical predictions in Fig. 5.12. The theoretical predictions shown in this figure assume different seeding masses and growth modes. We show results across the entire halo mass scale (Sijacki et al. 2015; Volonteri et al. 2016; Weinberger et al. 2018; Ni et al. 2022) and results restricted to subsamples at the high-mass and high-luminosity end (Volonteri et al. 2016; Ni et al. 2022). The departure from co-evolution of black hole accretion rate and the star formation rate at $z > 3$ is in general difficult to achieve with a cosmological model.

It has been suggested that good agreement between observations and simulations is only warranted when only including large black hole or halo masses, while using all masses causes simulations to over-predict the black hole accretion density (Sijacki et al. 2015; Volonteri et al. 2016). For example, the results from the Horizon-AGN simulations (Volonteri et al. 2016) are presented for the total mass range (solid dark green) and for halos with a halo hole mass of $> 5\times10^{11}\mathrm{M}_\odot$ (dashed dark green). Volonteri et al. (2016) proposed that supernova feedback could be the reason why observational results only agree with simulations when applying a high-mass threshold. Indeed supernovae feedback is expected to deplete the AGN core, effectively stopping black hole growth in low-mass galaxies (Dubois et al. 2015; Habouzit et al. 2017). As the galaxy grows in mass its deeper gravitational potential allows it to more efficiently confine the gas in the nucleus. The black hole accretion density derived from the X-ray-luminous eFEDS detections is in good agreement with the predictions from Volonteri et al. (2016) at $M_{\mathrm{halo}} > 5\times10^{11}\mathrm{M}_\odot$. Similarly, in their ASTRID simulations Ni et al. (2022) present their BHAD for various X-ray luminosity thresholds. At high luminosities ($\log L_X > 44$) the predictions of the steeper falling BHAD curves and the value derived from eFEDS are in excellent agreement. However, we stress again that the eFEDS survey is not sensitive to $44 < \log L_X < 45$ in the probed redshift regime. We are therefore missing contributions to accretion rate density from quasars accounted for in the Ni et al. (2022) BHAD curve. Because of its sensitivity limit, eFEDS becomes highly incomplete at $\log L_X < 45$ at $z \sim 6$ (see Fig. 5.10). The sample of Vito et al. (2018) is extracted from the *Chandra* Deep Fields. These surveys are smaller in area but deeper than eFEDS and can sample AGN efficiently down to luminosities $\log L_X > 42.5$. This difference in sensitivity explains the discrepancy seen at $z \sim 6$ between the results of Vito et al. (2018) and the lower boundary obtained in this Chapter: eFEDS misses quasars in the range $42.5 < L_X < 45$, which still significantly contribute to the overal BHAD. We also point out that the main and supplementary eFEDS catalogues are not spectroscopically complete. In this regard, the data point we derived should be considered a lower limit on the BHAD. Another source of discrepancy between the BHAD derived from various X-ray surveys is the use of photometric redshifts, which can potentially populate the $z > 5$ bin with interlopers. For our study, we only used X-ray sources with clear multi-wavelength identifications and spectroscopic redshifts. Following Volonteri et al. (2016), our results suggest that, assuming supernova-feedback-regulated black hole growth, most black hole accretion is dominated by extremely luminous AGN. Alternatively, the agreement with the prediction of Ni et al. (2022) indicates that at $\log L_X > 44$, black hole accretion is truly dominated by the most X-ray-luminous quasars at $\log L_X > 45$.

We have shown that the extrapolated XLF models by Vito et al. (2014), Ueda et al. (2014), Aird et al. (2015), Miyaji et al. (2015) and Georgakakis et al. (2015) underestimate the number of high-z quasar detections in eFEDS (see Fig. 5.11); however, it can be seen in Fig. 5.12 that the black hole accretion density derived from these models appears to be consistent with the one resulting from the eFEDS detections. This can be explained by the high X-ray to optical flux ratio for both eFEDS quasars, which results in significantly smaller bolometric corrections (see Fig. 5.9). The bolometric correction assumed for the conversion of XLFs to BHAD (Eq. 5.9) from Duras et al. (2020) causes a higher extrapolated BHAD, despite the underprediction of actual luminous high-z sources in the field. In addition, the black hole accretion density is calculated in Eq. 5.10 as the efficiency-scaled total emissivity of the quasars detected in the $z = 5.81 - 6.56$ interval and is therefore inversely proportional to the sensitive volume probed by eFEDS at the luminosities and redshifts of these quasars. The X-ray luminosity-redshift configurations of the quasars detected in eFEDS, in particular that of SDSS J0836+0054, result in a larger sensitive volume (see Fig. 5.4) and therefore a lower contribution to the black hole accretion density.

5.8 Discussion and conclusions

We have characterised a $z > 6$ super-Eddington-accreting NLS1 with low black hole mass based on eFEDS X-ray data, archival photometry and a new NIR spectrum. We discuss how our findings support the idea that $z > 6$ NLS1s potentially show physical properties that resemble those of their lower-redshift counterparts. At $z = 6.56$, J0921+0007 is the most distant X-ray-selected AGN to date and can therefore be used to impose constraints on the high-z XLF.

We derived a low black hole mass compared to a sample of high-redshift quasars with comparable optical/UV luminosity (see e.g. Onoue et al. 2019), which implies that the source is accreting at a super-Eddington rate. The values found ($M_{\rm BH} = (2.5 \pm 0.3) \times 10^8 M_\odot$ and $\lambda = 2.3^{+0.4}_{-0.3}$) are consistent with the typical properties of local NLS1s (e.g. Sulentic et al. 2000c; Collin & Kawaguchi 2004; Rakshit et al. 2017). We obtained a relatively steep power-law fit to the X-ray spectrum of the source: $\Gamma = 3.2$. Such a high value is usually found in the rest-frame soft band of archetypal low-z NLS1s (e.g. Boller et al. 1996; Brandt et al. 1997; Ojha et al. 2020). In the rest-frame hard band, NLS1s typically show photon indices below this value (~ 2; e.g. Zhou & Zhang 2010). The steeper photon index found here can be driven by either the large accretion rate (Shemmer et al. 2006) or the presence of unresolved non-coronal components. Similar sources, in terms of rest-frame optical properties, have been discovered by Koptelova et al. (2017) and Bañados et al. (2021). The quasar CFHQS J1641+3755 at $z = 6.04$ was initially discovered by Willott et al. (2007). Willott et al. (2010) obtained NIR spectroscopy for this source with the NIRI instrument on the Gemini-North Telescope. It shows an MgII profile ($\mathrm{FWHM}_{\mathrm{MgII}} = 1740 \pm 190\,\mathrm{km\,s^{-1}}$) that is very similar to the one observed in the MOIRCS spectrum of J0921+0007 presented here. According to the Rakshit et al. (2021) classification criterion, this makes it a high-z NLS1. The derived black hole mass and Eddington ratio are $M_{\mathrm{bh}} = 2.4 \times 10^8\,M_\odot$ and $\lambda = 2.3$, indicating that CFHQS J1641+3755 may be powered by a low-mass, strongly accreting black hole. Vito et al. (2019a) report the X-ray observation of this quasar with *Chandra*. While it has a relatively modest bolometric luminosity, it is the second-most X-ray-luminous source in their sample, making it deviate from the $\alpha_{OX} - L_{\mathrm{UV}}$ by 1.8σ with respect to the best-fitting relation of Steffen et al. (2006). Vito et al. (2019a) also derive a steep photon index for this source ($\Gamma = 2.56$). We conclude that CFHQS J1641+3755 is another archetypal NLS1 at high redshift. To further test the NLS1 classification of J0921+0007, we measured the extent of the quasar proximity zone and present these results in 5.5. A small proximity zone radius would have supported the hypothesis that J0921+0007 is a young quasar. However, we found a surprisingly large, luminosity-corrected proximity zone radius which challenges this interpretation. In Section 5.7.1 we show that the number of high-z source detections in the eFEDS field, combining the findings in this Chapter with the results from Chapter 4, is significantly higher than predictions from a large range of XLF models in the literature extrapolated out to $z \sim 6$. eFEDS is the largest contiguous public X-ray survey to date with sufficient depth to investigate $z \sim 6$ AGN demographics. It probes a cosmological volume that is sufficiently large to contain rare $\log L_{\mathrm{X}} > 45$ quasars at high redshift, including the unexpected class of high-z NLS1s discussed in this Chapter. The discrepancy between previous XLF models obtained using smaller pencil-beam or non-contiguous surveys underlines the necessity for wide surveys to obtain a realistic census of the rare, powerful sources at the bright end of the XLF (see e.g. Barlow-Hall et al. 2022).

Stacking the *Chandra* Deep Field South data from a sample of $3.5 < z < 6.5$ galaxies, Vito et al. (2016) show that the contribution of detected luminous quasars at $z \sim 6$ to the black hole accretion density is higher than the one from stacked undetected sources by an order of magnitude. These findings corroborate the results of Volonteri et al. (2016), who concluded that most of the black hole growth is contributed by luminous quasars ($L_{\mathrm{bol}} > 10^{43}\,\mathrm{erg\,s^{-1}}$) in massive halos ($> 5 \times 10^{11}\mathrm{M}_\odot$). The accretion density derived from the two detected quasars in eFEDS is consistent with these previous results. At the flux limit of eFEDS, it is only possible to sample the luminous population with $L_{2-10\,\mathrm{keV}} > 10^{45}\,\mathrm{erg\,s^{-1}}$. Despite this sensitivity limit, our results are already consistent with the predictions from

Volonteri et al. (2016) and Ni et al. (2022), indicating that most of the black hole growth is in fact driven by X-ray-ultra-luminous quasars, above the eROSITA sensitivity limit. J0921+0007 is an unexpected member of this category of extreme quasars. Its X-ray luminosity is significantly higher than the value extrapolated from the $\alpha_{\rm OX} - L_{\rm UV}$ relation. In order to quantify how much of the accretion density is in fact driven by young, super-Eddington black holes, a wider survey area will be required at this depth to obtain a more informative sample. This will be made possible in the cumulative eROSITA All-Sky Survey (Merloni et al. 2012, see also Seppi et al. 2022).

Chapter 6

Conclusions and ongoing discovery programme

6.1 Summary

The seeding and growth channels of SMBHs are, to date, an unresolved puzzle. How could $> 10^9\ M_\odot$ black holes have formed within ~ 600 Myr? In this thesis, I have investigated this issue via the multi-wavelength characterisation and population study of X-ray luminous AGN, which are thought to pinpoint strong accretion onto black holes on a large redshift scale. The focus is on two important AGN population tracers: the masses of the black holes at their centre and their space density through cosmic times. These observables essentially provide a handle on the growth history of SMBHs and their overall evolution.

The primary datasets explored in this thesis were obtained from X-ray missions. In the year the presented thesis project was started, the X-ray telescope eROSITA was launched aboard the SRG observatory. It has already delivered groundbreaking surveys such as the deep and wide eFEDS field (Brunner et al. 2022; Salvato et al. 2022) and the first iterations of the all-sky survey (Merloni et al., in prep.). The depth reached by the surveys on the entire sky enables the detection of scarce populations, such as the highest redshift quasars.

In order to identify the AGN population in these surveys, a multi-wavelength characterisation of the detected X-ray sources is needed. However, the identification of optical, IR and UV counterparts to X-ray sources is a challenging task because of the positional uncertainties usually associated with their detections. In Chapter 2, I have presented a state-of-the-art Bayesian catalogue cross-matching technique, accounting for both astrometry and independent auxiliary information (Salvato et al. 2018). To this effect, random forest classifiers were trained for various optical and IR surveys (e.g. DESI Legacy DR9, DR10) in order to predict the probability of sources of emitting X-ray radiation based purely on photometry, morphology and proper motion. For any source in these surveys, this quantity can be stored.

In Chapter 3, I presented a statistical analysis of the SDSS-IV/SPIDERS optical spectra of X-ray luminous Type 1 AGN (Wolf et al. 2020). The initial goal of the study was to determine the drivers of spectral diversity in Type 1 AGN. The correlation space of spectral parameters, such as the strength and the shape of emission lines, provided by Coffey et al. (2019) were investigated with a PCA and put in the context of the quasar main sequence. The distribution of broad Balmer asymmetries along the main quasar sequence confirmed the presence of outflows in high Eddington-ratio AGN. When estimating $M_{\rm BH}$ via the single-epoch method, non-Keplerian kinematics, as traced by shifted broad components, can affect the width of the lines and, therefore, the mass computation.

Shortly after the launch of eROSITA and its performance verification observations, the focus was shifted to larger scales and earlier times as we investigated AGN that emitted their light within the first billion years after the Big Bang. These objects hold the most stringent constraints on black hole evolution. The studies are presented in Chapters 4 and 5. We searched for high-redshift quasars in the eFEDS field by cross-matching X-ray detections to the sample of previously discovered, spectroscopically confirmed sample of $z > 5.7$ quasars in the footprint. We could report the eROSITA detection of two quasars SDSS J0836+0054 ($z = 5.81$, Fan et al. 2001) and HSC J0921+0007 ($z = 6.56$, Matsuoka et al. 2018a). The latter is the most distant serendipitous X-ray detection known. While both sources are X-ray luminous by eROSITA selection, they display very different restframe optical/UV properties. SDSS J0836+0054 is an optically luminous quasar in the typical $\alpha_{\rm OX} - L_{\rm UV}$ range, while HSC J0921+0007 is of low optical luminosity. From a newly obtained NIR spectrum, we could confirm that HSC J0921+0007 is powered by a relatively low-mass SMBH of $\sim 2 \times 10^8\ M_\odot$, accreting at super-Eddington rate $\lambda_{\rm Edd} \sim 2$; a high-z analogue to local NLS1 galaxies. With the detection of the sources, we could impose first spectroscopic constraints on the XLF at $z \sim 6$ and confirm that it follows a flatter slope on its luminous end than predicted by redshift-extrapolated models from the literature. A larger number of powerful quasars is therefore expected to be discovered at cosmic dawn, which will, in the near future, result in new constraints on viable SMBH seed models. We further could show that the detections translate to a black hole accretion rate density at $z \sim 6$, which is consistent with models restricting black holes to the most massive dark matter halos (e.g. Volonteri et al. 2016, $> 5 \times 10^{11}\,{\rm M}_\odot$). The deep gravitational potential of such halos suppresses supernovae feedback-induced gas depletion.

Extrapolating from these early eROSITA results, we predicted the detection of hundreds of $z > 5.7$ quasars in eRASS:8 (Section 4.8.4), which will ultimately yield very stringent constraints on X-ray AGN demographics and thus SMBH growth. However, less than 50 quasars have been detected in the X-rays. A complementary approach to the a posteriori identification of already spectroscopically confirmed sources (Chapters 4 and 5) is the search for new yet unidentified $z > 5.7$ quasars in eRASS.

6.2 Discovery of rare $z > 5.7$ X-ray luminous quasars in eRASS

In this outlook section, I present a novel high-redshift selection pipeline currently under development and whose results will be published after the submission of this thesis. (Wolf et al. in prep.).

To identify yet unknown rare quasars in eRASS, I have recently developed a multi-wavelength pilot study. The photometric pipeline pre-selects optical/NIR candidates, filters them and finally investigates their positions in the eRASS data. The examined footprint consists of the intersections of the eRASS, VHS and DES data release 2 (DES DR2) footprint (essentially corresponding to the ~ 5000 deg^2 of the DES DR2 footprint). The consecutive steps of the eRASS $z > 5.7$ quasar selection pipeline are presented in the following sub-sections.

6.2.1 Candidate selection

DES colour pre-selection in NIR detection

The DES DR2 main catalogue of sources is down-selected following a very loose set of colour and magnitude selection criteria similar to those applied in previous high-redshift quasar surveys (e.g., Bañados et al. 2016; Matsuoka et al. 2018a). These aim to reduce the number of filtered sources in the next steps. They select i-band drop-out objects and limit the slope in the power-law redwards of the Lyman α line. In addition, the photometric pre-selection avoids typical colours of L-, M- and T-dwarfs, the primary contaminants in distant quasar searches. The exact set of cuts will be detailed in Wolf et al. (in prep.). The left panel in Fig. 6.1 shows the selection region in the $i - z$ vs $z - Y$ colour-colour plane. Corresponding candidates are then cross-matched positionally within ($< 1''$) to the VHS data release 5 (VHS DR5) and the CatWISE2020 (Marocco et al. 2021) source catalogues. This step yields a pre-selected sample of optical/IR quasar candidates.

Filter I: SED template fitting

The first filter is the photometric redshift estimation of the pre-selected optical/NIR sample using SED template fitting. I made use of the photometric code Le Phare (Arnouts et al. 1999; Ilbert et al. 2006), which is based on a χ^2-minimization method. Photometry in the following bands was evaluated in the procedure: g, r, i, Z, Y from DES DR2, Y, J, H, Ks from VHS DR5 and $W1, W2$ from CatWISE2020. For each source, AGN templates and stellar templates are compared to the photometry. The extended stellar and AGN template libraries used for this procedure will be presented in Wolf et al. (in prep.). Sources are kept as high-redshift quasar candidates if they have a photometric redshift $z_{\rm phot} > 5.7$. Additionally, the AGN template is required to outperform the stellar template

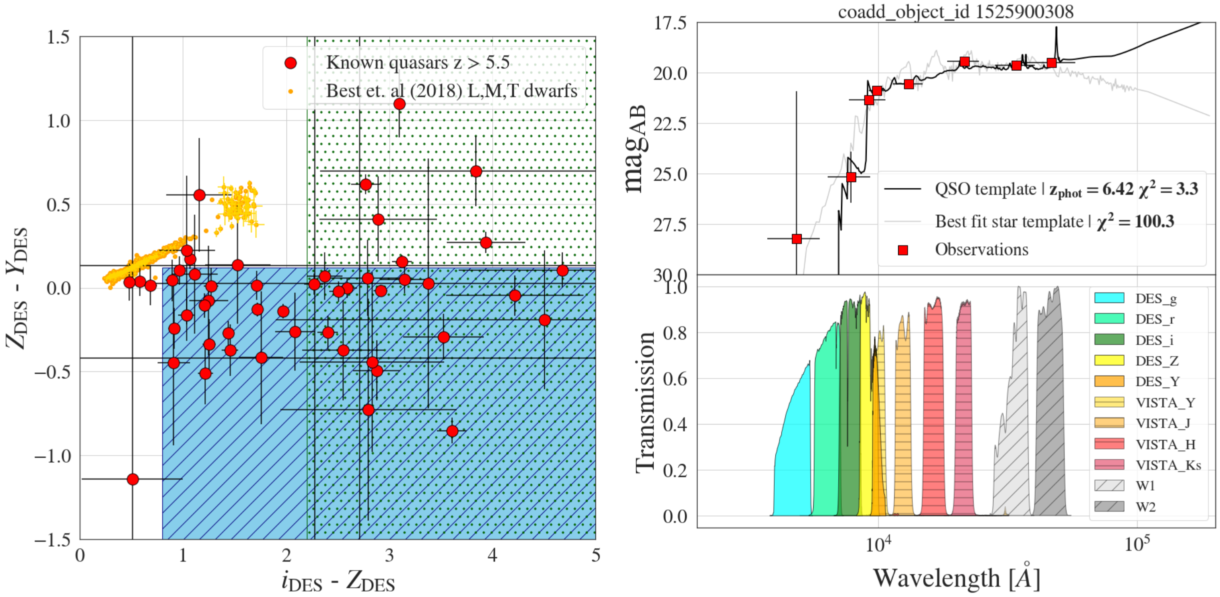


Figure 6.1: **Left panel:** Selection region of quasar candidates in the DES $i - Z$ vs $Z - Y$ colour-colour plane. Red points correspond to previously known, spectroscopically confirmed quasars detected in DES DR2. The yellow points are L-, M- and T-dwarfs from Best et al. (2018). Such cool stars are the main source of contaminants in high-z quasar searches. The two colour-colour cuts represented in this plane (green-dotted and blue-dashed areas) loosely avoid the region of contamination. **Right panel:** An example Le Phare fit to the optical/IR photometry of a high-redshift quasar candidate is shown. The photometric data points and their associated errors are displayed as red squares. The best fit AGN model ($\chi_{\text{AGN}} = 3.3$) is shown as the solid black line, while the best stellar model is displayed in grey ($\chi_{\text{stellar}} = 100.3$). The photometric redshift of this source is $z_{\text{phot}} = 6.42$. The lower panel of this figure shows the transmission curves of the various filters used for the template fitting.

significantly(i.e., $\chi^2_{\text{stellar}} > 5 \times \chi^2_{\text{AGN}}$). An example Le Phare fit to a candidate quasar is shown in Fig. 6.1 (right panel).

Filter II: Random Forest photometric classification

A second machine learning filter was developed to account simultaneously for optical/IR photometry and morphology. A random forest was trained to classify a training sample of spectroscopically confirmed quasars from literature, L- M- and T-dwarfs (Best et al. 2018) and lower redshift quasars from the SDSS DR16Q catalogue (Lyke et al. 2020). All the training samples are detected in DES DR2 and CatWISE2020. Morphology is accounted for by including DES DR2 aperture photometry as a training feature. Proper motion and basic photometric features from CatWISE2020 are additional features included in the training. The complete list of features and training sample definition will be released in Wolf et al. (in prep.). The classes of the pre-selected optical/IR counterparts are inferred with the trained random forest. Sources classified as high-redshift quasars are retained.

eRASS detection

Following the above down-selection steps, I obtain a sample of 319 $z_{\text{phot}} > 5.7$ optical/IR candidates, for which I now investigate the eROSITA data. A candidate is considered eROSITA detected if:

- It is associated with an X-ray detection from any eRASS point-source catalogue.[1] Preliminary optical counterparts from the DESI Legacy Surveys were determined using `NWAY` (see 2.2 for more details). The sample of high-redshift quasar candidates is matched positionally to the list of reliable DESI Legacy counterparts of the eRASS catalogue sources (match radius: $1''$).

- It is detected on the cumulative eROSITA soft-band images (currently eRASS:4, $0.2 - 2.3\,\text{keV}$). For this, I performed aperture photometry on the soft X-ray images within the manually defined source and background regions. The aperture is centred on the optical coordinates of the quasar candidates. As in Chapter 5, I computed the binomial no-source probability (Weisskopf et al. 2007) as

  $$P_B(i \geq s) = \sum_{i=s}^{s+b} \frac{(s+b)!}{i!(i-s-b)!} \left(\frac{1}{1+r}\right)^i \left(\frac{2+r}{1+r}\right)^{s+b-i}, \tag{6.1}$$

  where s and b are counts in the source and background region and r is the ratio of areas of the two extraction regions.

  I currently use a detection threshold $P_B(i \geq s) < 0.01$ following Vito et al. (2019a). This value will be calibrated with simulations in Wolf et al. (in prep.).

---

[1] As of January 2023, these catalogues are not yet public and will be gradually released to the astronomical community in a sequence of data releases.

6.2.2 5 new X-ray luminous high-redshift quasars

In the context of this pilot survey, 9 quasar candidates were submitted for spectroscopic follow-up at the Magellan Clay telescope in Las Campanas, Chile. The main objective of these observations was the redshift confirmation of the quasars using the Lyman α emission line. The candidates were observed with the Low Dispersion Survey Spectrograph (LDSS-3, red grism: $6000 - 10000\,\text{Å}$).[2]. Nine targets were submitted and observed from July 2021 to October 2021, with exposure times between 2.4 and 3.6 ks. Out of these nine candidates, five could be confirmed as $z > 5.6$ quasars. The redshifts were determined by the shift of the Lyman α line and other emission features. Fig. 6.2 presents their discovery spectra and associated redshifts. The non-confirmed sources have stellar SEDs and will be further discussed in Wolf et al. (in prep.). The success rate of 55% is competitive but will need to be confirmed in the context of a larger spectroscopic programme. The discovered quasars are located, by selection, at the sparsely populated bright end of the XLF, which makes them powerful probes of black hole accretion density. An extensive multi-wavelength follow-up campaign has been initiated to characterise these sources further. It comprises deeper Chandra ACIS-S follow-up exposures, radio follow-up with the Karl G. Jansky Very Large Array (JVLA) and NIR spectroscopy with the Folded Port Infrared Echelette spectrograph (FIRE). These observations will be carried out over the course of 2023 and respectively increase the X-ray count statistics, investigate the blazar nature of the sources and provide single-epoch black hole masses. The pilot survey demonstrates the large discovery potential of combining deep optical/IR photometry with sensitive and wide X-ray surveys such as eRASS.

6.3 Concluding remarks

eROSITA high-redshift quasars are, by selection, a population at the tip-of-the-iceberg in terms of accretion power, as the instrument's sensitivity limits the exploitable parameter space to the most X-ray luminous AGN at $z > 5.7$. Through detections and discoveries of distant quasars on the entire sky, eROSITA will, by the end of its main survey cycle, provide the most robust constraints on the bright end of the AGN XLF and, thus, on rapid SMBH accretion in the early universe. For the identification of these extremely rare objects, a full-sky coverage of deep ancillary optical/IR photometric surveys is of critical importance. Currently, surveys such as DES, DESI Legacy, Pan-STARR, HSC, VIKING, VHS and CatWISE2020 provide a rich baseline that can be combined to sample the restframe optical/UV AGN SED over large ranges of redshifts. Soon, a paradigm change will be set in motion by the Euclid space telescope (Euclid Collaboration et al. 2019) and the Vera C. Rubin Observatory (Ivezić et al. 2019). These missions will supply high-redshift quasar searches with pristine IR and optical photometry of yet unmatched depth over large areas of the sky. Simultaneously, the recently launched JWST and its

---

[2] The observations and data reductions were carried out by my collaborators Dr Francesco Di Mille and Dr S. Ciroi.

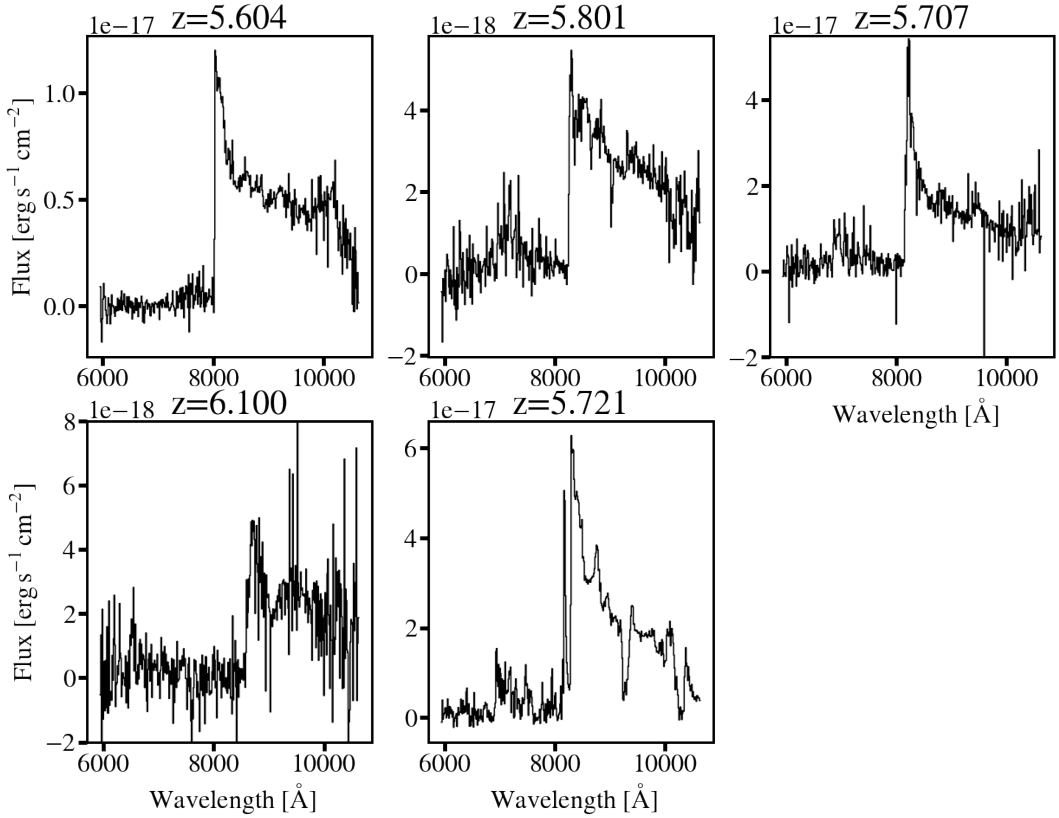


Figure 6.2: Discovery spectra of high-redshift quasars in eRASS. The manually estimated associated redshifts are displayed on the top of each panel. The sources will be presented in an upcoming paper by Wolf et al. (in prep.).

now activated advanced NIR spectroscopic and imaging instrumentation will serve as a magnifying glass to resolve the structure, chemical content and kinematics within quasars uncovered deep in the epoch of reionization.

The fleet of space observatories and the ground-based telescope park of the 2020s will likely enable a significant push on the redshift frontier of SMBH studies. It is not unlikely that in the wake of this progression, astronomers will also catch the progenitors of massive $z \sim 7$ quasars in their infancy. Studying the growth channels of these seeds will provide significant insights into the evolution of black holes and the co-evolution between host and nucleus. It is beyond doubt that the coming years will deliver the direly needed answers on the fundamental question of modern astronomy and astrophysics at the heart of this thesis: *how did the massive black holes in the centre of nearly all galaxies form ?*

Bibliography

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Phys. Rev. Lett., 116, 061102

Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, ApJS, 235, 42

Abramowicz, M. A., Chen, X., Kato, S., Lasota, J.-P., & Regev, O. 1995, ApJ, 438, L37

Adhikari, T. P., Różańska, A., Czerny, B., Hryniewicz, K., & Ferland, G. J. 2016, ApJ, 831, 68

Ahumada, R., Prieto, C. A., Almeida, A., et al. 2020, ApJS, 249, 3

Ai, Y., Dou, L., Fan, X., et al. 2016, ApJ, 823, L37

Aihara, H., AlSayyad, Y., Ando, M., et al. 2022, PASJ, 74, 247

Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, PASJ, 70, S4

Aird, J., Coil, A. L., Georgakakis, A., et al. 2015, MNRAS, 451, 1892

Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, ARA&A, 49, 409

Alonso-Herrero, A., Quillen, A. C., Rieke, G. H., Ivanov, V. D., & Efstathiou, A. 2003, AJ, 126, 81

Alvarez, M. A., Wise, J. H., & Abel, T. 2009, ApJ, 701, L133

Ameijeiras-Alonso, J., Crujeiras, R. M., & Rodríguez-Casal, A. 2016, arXiv e-prints, arXiv:1609.05188

Ananna, T. T., Treister, E., Urry, C. M., et al. 2019, ApJ, 871, 240

Antonucci, R. 1993, ARA&A, 31, 473

Antonucci, R. R. J. & Miller, J. S. 1985, ApJ, 297, 621

Arnaboldi, M., Neeser, M. J., Parker, L. C., et al. 2007, The Messenger, 127, 28

Arnaud, K. A. 1996, in Astronomical Society of the Pacific Conference Series, Vol. 101, Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby & J. Barnes, 17

Arnaud, K. A., Branduardi-Raymont, G., Culhane, J. L., et al. 1985, MNRAS, 217, 105

Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540

Ashtekar, A. & Bojowald, M. 2005, Classical and Quantum Gravity, 22, 3349

Asmus, D., Hönig, S. F., & Gandhi, P. 2016, ApJ, 822, 109

Assef, R. J., Denney, K. D., Kochanek, C. S., et al. 2011, ApJ, 742, 93

Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, ApJ, 772, 26

Avni, Y. & Tananbaum, H. 1986, ApJ, 305, 83

Bañados, E., Mazzucchelli, C., Momjian, E., et al. 2021, ApJ, 909, 80

Bañados, E., Venemans, B. P., Decarli, R., et al. 2016, ApJS, 227, 11

Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018, Nature, 553, 473

Bañados, E., Venemans, B. P., Morganson, E., et al. 2015, ApJ, 804, 118

Baade, W. & Minkowski, R. 1954, ApJ, 119, 206

Baba, K., Shibata, R., & Sibuya, M. 2004, Australian & New Zealand Journal of Statistics, 46, 657

Baldwin, J., Ferland, G., Korista, K., & Verner, D. 1995, ApJ, 455, L119

Baldwin, J. A. 1977, ApJ, 214, 679

Baldwin, J. A., Burke, W. L., Gaskell, C. M., & Wampler, E. J. 1978, Nature, 273, 431

Baldwin, J. A., Ferland, G. J., Korista, K. T., Hamann, F., & LaCluyzé, A. 2004, ApJ, 615, 610

Bambi, C. 2020, in Multifrequency Behaviour of High Energy Cosmic Sources - XIII. 3-8 June 2019. Palermo, 28

Banados, E., Schindler, J.-T., Venemans, B. P., et al. 2022, arXiv e-prints, arXiv:2212.04452

Barger, A. J., Cowie, L. L., Capak, P., et al. 2003, ApJ, 584, L61

Barger, A. J., Cowie, L. L., Mushotzky, R. F., et al. 2005, AJ, 129, 578

Barkana, R. & Loeb, A. 2001, Phys. Rep., 349, 125

Barlow-Hall, C. L., Delaney, J., Aird, J., et al. 2022, arXiv e-prints, arXiv:2201.11139

Baron, D. 2019, arXiv e-prints [`arXiv:1904.07248`]

Baron, D. & Ménard, B. 2019, MNRAS, 487, 3404

Barth, A. J., Martini, P., Nelson, C. H., & Ho, L. C. 2003, ApJ, 594, L95

Barth, A. J., Pancoast, A., Bennert, V. N., et al. 2013, ApJ, 769, 128

Barthelmy, S. D., Barbier, L. M., Cummings, J. R., et al. 2005, Space Sci. Rev., 120, 143

Barvainis, R. 1987, ApJ, 320, 537

Baskin, A. & Laor, A. 2018, MNRAS, 474, 1970

Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559

Begelman, M. C. 1979, MNRAS, 187, 237

Begelman, M. C. & Rees, M. J. 1978, MNRAS, 185, 847

Begelman, M. C., Volonteri, M., & Rees, M. J. 2006, MNRAS, 370, 289

Belladitta, S., Moretti, A., Caccianiga, A., et al. 2020, A&A, 635, L7

Bennett, A. S. 1962, MNRAS, 125, 75

Bentz, M. C. & Katz, S. 2015, PASP, 127, 67

Bentz, M. C., Walsh, J. L., Barth, A. J., et al. 2009, ApJ, 705, 199

Best, W. M. J., Magnier, E. A., Liu, M. C., et al. 2018, in American Astronomical Society Meeting Abstracts, Vol. 231, American Astronomical Society Meeting Abstracts #231, 137.01

Bhowmick, A. K., Somerville, R. S., Di Matteo, T., et al. 2020, MNRAS, 496, 754

Bian, W.-H. & Zhao, Y.-H. 2003, PASJ, 55, 599

Bianchi, S., Guainazzi, M., Matt, G., Fonseca Bonilla, N., & Ponti, G. 2009, A&A, 495, 421

Blandford, R. D. & Königl, A. 1979, ApJ, 232, 34

Blandford, R. D. & McKee, C. F. 1982, ApJ, 255, 419

Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28

Bochkarev, N. G. & Antokhin, I. I. 1982, Astronomicheskij Tsirkulyar, 1238, 1

Boksenberg, A., Snijders, M. A. J., Wilson, R., et al. 1978, Nature, 275, 404

Boller, T., Brandt, W. N., & Fink, H. 1996, A&A, 305, 53

Boller, T., Freyberg, M. J., Trümper, J., et al. 2016, A&A, 588, A103

Bolton, C. T. 1972, Nature, 235, 271

Bon, N., Bon, E., & Marziani, P. 2018, Frontiers in Astronomy and Space Sciences, 5, 3

Bonanno, A. & Reuter, M. 2000, Phys. Rev. D, 62, 043008

Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103

Boroson, T. A. & Green, R. F. 1992, ApJS, 80, 109

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34

Bowyer, C. S., Lampton, M., Mack, J., & de Mendonca, F. 1970, ApJ, 161, L1

Bowyer, S., Byram, E. T., Chubb, T. A., & Friedman, H. 1965, Science, 147, 394

Brandt, W. N., Alexander, D. M., Hornschemeier, A. E., et al. 2001, AJ, 122, 2810

Brandt, W. N., Mathur, S., & Elvis, M. 1997, MNRAS, 285, L25

Brandt, W. N., Schneider, D. P., Fan, X., et al. 2002, ApJ, 569, L5

Brotherton, M. S. 1996, ApJS, 102, 1

Brotherton, M. S., Singh, V., & Runnoe, J. 2015, MNRAS, 454, 3864

Brotherton, M. S., Wills, B. J., Francis, P. J., & Steidel, C. C. 1994, ApJ, 430, 495

Brunner, H., Liu, T., Lamer, G., et al. 2022, A&A, 661, A1

Brusa, M., Civano, F., Comastri, A., et al. 2010, ApJ, 716, 348

Brusa, M., Comastri, A., Gilli, R., et al. 2009, ApJ, 693, 8

Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

Buchner, J. 2021, The Journal of Open Source Software, 6, 3001

Buchner, J., Georgakakis, A., Nandra, K., et al. 2015, ApJ, 802, 89

Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125

Budavári, T. & Szalay, A. S. 2008, ApJ, 679, 301

Calderone, G., Nicastro, L., Ghisellini, G., et al. 2017, MNRAS, 472, 4051

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682

Cao, X. & Li, F. 2008, MNRAS, 390, 561

Capak, P. L., Riechers, D., Scoville, N. Z., et al. 2011, Nature, 470, 233

Carilli, C. L., Wang, R., Fan, X., et al. 2010, ApJ, 714, 834

Carter, B. 1966, Phys. Rev., 141, 1242

Carter, B. 1971, Phys. Rev. Lett., 26, 331

Cash, W. 1979, ApJ, 228, 939

Celotti, A., Ghisellini, G., & Chiaberge, M. 2001, MNRAS, 321, L1

Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv e-prints, arXiv:1612.05560

Chen, H. & Gnedin, N. Y. 2018, ApJ, 868, 126

Chen, K. & Halpern, J. P. 1989, ApJ, 344, 115

Chen, K., Halpern, J. P., & Filippenko, A. V. 1989, ApJ, 339, 742

Cherepashchuk, A. M. & Lyutyi, V. M. 1973, Astrophys. Lett., 13, 165

Choi, J.-H., Shlosman, I., & Begelman, M. C. 2015, MNRAS, 450, 4411

Civano, F., Brusa, M., Comastri, A., et al. 2011, ApJ, 741, 91

Clerc, N. & Finoguenov, A. 2022, arXiv e-prints, arXiv:2203.11906

Clerc, N., Merloni, A., Zhang, Y. Y., et al. 2016, MNRAS, 463, 4490

Clerc, N., Ramos-Ceja, M. E., Ridl, J., et al. 2018, A&A, 617, A92

Coffey, D., Salvato, M., Merloni, A., et al. 2019, A&A, 625, A123

Cohen, M. H., Cannon, W., Purcell, G. H., et al. 1971, ApJ, 170, 207

Collier, S., Crenshaw, D. M., Peterson, B. M., et al. 2001, ApJ, 561, 146

Collin, S. & Joly, M. 2000, New Astronomy Reviews, 44

Collin, S. & Kawaguchi, T. 2004, A&A, 426, 797

Comastri, A., Setti, G., Zamorani, G., & Hasinger, G. 1995, A&A, 296, 1

Combes, F., García-Burillo, S., Audibert, A., et al. 2019, A&A, 623, A79

Comparat, J., Merloni, A., Dwelly, T., et al. 2020, A&A, 636, A97

Comparat, J., Merloni, A., Salvato, M., et al. 2019, MNRAS, 487, 2005

Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693

Coppejans, R., Frey, S., Cseh, D., et al. 2016, MNRAS, 463, 3260

Coppejans, R., van Velzen, S., Intema, H. T., et al. 2017, MNRAS, stx215

Corbin, M. R. 1995, ApJ, 447, 496

Crummy, J., Fabian, A. C., Gallo, L., & Ross, R. R. 2006, MNRAS, 365, 1067

Curtis, H. D. 1918, Publications of Lick Observatory, 13, 9

Cutri, R. M., Wright, E. L., Conrow, T., et al. 2013, Explanatory Supplement to the AllWISE Data Release Products, Explanatory Supplement to the AllWISE Data Release Products

Czerny, B. & Hryniewicz, K. 2010, A&A, 525

Czerny, B., Karas, V., Li, Y.-R., et al. 2018, in XXXVIII Polish Astronomical Society Meeting, ed. A. Ró&żańska, Vol. 7, 264–266

Czerny, B., Li, Y.-R., Hryniewicz, K., et al. 2017, ApJ, 846

Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al. 2016, MNRAS, 460, 1270

Davies, F. B., Hennawi, J. F., & Eilers, A.-C. 2019, ApJ, 884, L19

Davies, F. B., Hennawi, J. F., & Eilers, A.-C. 2020, MNRAS, 493, 1330

Davis, S. W. & Laor, A. 2011, ApJ, 728, 98

Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, AJ, 151, 44

de Gasperin, F., Dijkema, T. J., Drabent, A., et al. 2019, A&A, 622, A5

De Marco, B., Ponti, G., Cappi, M., et al. 2013, MNRAS, 431, 2441

De Rosa, G., Venemans, B. P., Decarli, R., et al. 2014, ApJ, 790, 145

Delaney, J. N., Aird, J., Evans, P. A., et al. 2022, arXiv e-prints, arXiv:2212.07464

Delvecchio, I., Gruppioni, C., Pozzi, F., et al. 2014, MNRAS, 439, 2736

Dent, W. A. 1965, Science, 148, 1458

Devecchi, B. & Volonteri, M. 2009, ApJ, 694, 302

Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168

Dietrich, M., Hamann, F., Shields, J. C., et al. 2002, ApJ, 581, 912

Dijkstra, M., Haiman, Z., Mesinger, A., & Wyithe, J. S. B. 2008, MNRAS, 391, 1961

Drinkwater, M. J., Jurek, R. J., Blake, C., et al. 2010, MNRAS, 401, 1429

Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, Astronomy and Geophysics, 50, 5.12

Du, P., Wang, J.-M., Hu, C., et al. 2016, ApJ, 818, L14

Dubois, Y., Volonteri, M., Silk, J., et al. 2015, MNRAS, 452, 1502

Dunn, G., Bellovary, J., Holley-Bockelmann, K., Christensen, C., & Quinn, T. 2018, ApJ, 861, 39

Duras, F., Bongiorno, A., Piconcelli, E., et al. 2017, A&A, 604, A67

Duras, F., Bongiorno, A., Ricci, F., et al. 2020, A&A, 636, A73

Dwelly, T. et al. 2017, Mon. Not. Roy. Astron. Soc., 469, 1065

Ebizuka, N., Ichiyama, K., Yamada, T., et al. 2011, PASJ, 63, 605

Edge, A., Sutherland, W., Kuijken, K., et al. 2013, The Messenger, 154, 32

Edge, D. O., Shakeshaft, J. R., McAdam, W. B., Baldwin, J. E., & Archer, S. 1959, MmRAS, 68, 37

Eilers, A.-C., Davies, F. B., Hennawi, J. F., et al. 2017, ApJ, 840, 24

Eilers, A.-C., Hennawi, J. F., Davies, F. B., & Simcoe, R. A. 2021, ApJ, 917, 38

Eilers, A.-C., Hennawi, J. F., Decarli, R., et al. 2020, ApJ, 900, 37

Einstein, A. 1916, Annalen der Physik, 354, 769

Elitzur, M. 2012, ApJ, 747, L33

Elvis, M. 2000, ApJ, 545, 63

Elvis, M., Civano, F., Vignali, C., et al. 2009, ApJS, 184, 158

Elvis, M., Maccacaro, T., Wilson, A. S., et al. 1978, MNRAS, 183, 129

Elvis, M., Wilkes, B. J., McDowell, J. C., et al. 1994, ApJS, 95, 1

Emmering, R. T., Blandford, R. D., & Shlosman, I. 1992, ApJ, 385, 460

Eracleous, M. & Halpern, J. P. 1994, ApJS, 90, 1

Eracleous, M. & Halpern, J. P. 2003, ApJ, 599, 886

Euclid Collaboration, Barnett, R., Warren, S. J., et al. 2019, A&A, 631, A85

Evans, I. N., Primini, F. A., Miller, J. B., et al. 2020, in American Astronomical Society Meeting Abstracts, American Astronomical Society Meeting Abstracts, 154.05

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2022, ApJ, 930, L12

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019a, ApJ, 875, L1

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019b, ApJ, 875, L6

Everett, J. E. 2005, ApJ, 631, 689

Fabian, A. C. 2012, ARA&A, 50, 455

Fabian, A. C. & Iwasawa, K. 1999, MNRAS, 303, L34

Fabian, A. C., Rees, M. J., Stella, L., & White, N. E. 1989, MNRAS, 238, 729

Fan, X., Banados, E., & Simcoe, R. A. 2022, arXiv e-prints, arXiv:2212.06907

Fan, X., Carilli, C. L., & Keating, B. 2006a, ARA&A, 44, 415

Fan, X., Hennawi, J. F., Richards, G. T., et al. 2004, AJ, 128, 515

Fan, X., Narayanan, V. K., Lupton, R. H., et al. 2001, AJ, 122, 2833

Fan, X., Strauss, M. A., Becker, R. H., et al. 2006b, AJ, 132, 117

Fanaroff, B. L. & Riley, J. M. 1974, MNRAS, 167, 31P

Fanidakis, N., Baugh, C. M., Benson, A. J., et al. 2012, MNRAS, 419, 2797

Feltre, A., Hatziminaoglou, E., Fritz, J., & Franceschini, A. 2012, MNRAS, 426, 120

Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, Rev. Mexicana Astron. Astrofis., 49, 137

Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601

Ferrarese, L. & Merritt, D. 2000, ApJ, 539, L9

Finkelstein, D. 1958, Physical Review, 110, 965

Fischer, N., Mammen, E., & Marron, J. 1994, Computational Statistics & Data Analysis, 18, 499

Fotopoulou, S., Buchner, J., Georgantopoulos, I., et al. 2016, A&A, 587, A142

Francis, P. J., Hewett, P. C., Foltz, C. B., & Chaffee, F. H. 1992, ApJ, 398, 476

Frey, S., Mosoni, L., Paragi, Z., & Gurvits, L. I. 2003, MNRAS, 343, 20

Frey, S., Paragi, Z., Gurvits, L. I., Gabányi, K. É., & Cseh, D. 2011, A&A, 531, L5

Frey, S., Paragi, Z., Mosoni, L., & Gurvits, L. I. 2005, A&A, 436, L13

Friedman, H. & Byram, E. T. 1967, Science, 158, 257

Frolov, V. P. 2014, Journal of High Energy Physics, 2014, 49

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1

Gambini, R. & Pullin, J. 2014, arXiv e-prints, arXiv:1408.3050

Ganci, V., Marziani, P., D'Onofrio, M., et al. 2019, A&A, 630, A110

Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, Space Sci. Rev., 123, 485

Gaskell, C. M. 2009, New Astronomy Reviews, 53, 140 , proceedings of the VII Serbian Conference on Spectral Line Shapes (VII SCSLSA) held in Zrenjanin, Serbia June 15th-19th 2009

Gaskell, C. M. & Harrington, P. Z. 2018, MNRAS, 478, 1660

Gebhardt, K., Kormendy, J., Ho, L. C., et al. 2000a, ApJ, 543, L5

Gebhardt, K., Richstone, D., Kormendy, J., et al. 2000b, AJ, 119, 1157

Gebhardt, K., Richstone, D., Tremaine, S., et al. 2003, ApJ, 583, 92

Gehrels, N. 1986, ApJ, 303, 336

Genzel, R., Eckart, A., Ott, T., & Eisenhauer, F. 1997, MNRAS, 291, 219

Genzel, R., Pichon, C., Eckart, A., Gerhard, O. E., & Ott, T. 2000, MNRAS, 317, 348

Georgakakis, A., Aird, J., Buchner, J., et al. 2015, MNRAS, 453, 1946

Georgakakis, A. & Nandra, K. 2011, MNRAS, 414, 992

Georgakakis, A., Nandra, K., Laird, E. S., Aird, J., & Trichas, M. 2008, MNRAS, 388, 1205

George, I. M. & Fabian, A. C. 1991, MNRAS, 249, 352

Ghez, A. M., Klein, B. L., Morris, M., & Becklin, E. E. 1998, ApJ, 509, 678

Ghez, A. M., Morris, M., Becklin, E. E., Tanner, A., & Kremenek, T. 2000, Nature, 407, 349

Ghirardini, V., Bulbul, E., Hoang, D. N., et al. 2021, A&A, 647, A4

Giacconi, R., Gursky, H., Paolini, F. R., & Rossi, B. B. 1962, Phys. Rev. Lett., 9, 439

Gierliński, M. & Done, C. 2004, MNRAS, 349, L7

Gilli, R., Comastri, A., & Hasinger, G. 2007, A&A, 463, 79

Gilli, R., Risaliti, G., & Salvati, M. 1999, A&A, 347, 424

Giroletti, M. & Polatidis, A. 2009, Astronomische Nachrichten, 330, 193

Godfrey, L. E. H., Bicknell, G. V., Lovell, J. E. J., et al. 2009, ApJ, 695, 707

Goodrich, R. W. 1989, ApJ, 342, 224

Grandi, P. & Palumbo, G. G. C. 2004, Science, 306, 998

GRAVITY Collaboration, Dexter, J., Shangguan, J., et al. 2020, A&A, 635, A92

Greene, J. E. & Ho, L. C. 2005, ApJ, 630, 122

Greenstein, J. L. & Schmidt, M. 1964, ApJ, 140, 1

Greiner, J., Bornemann, W., Clemens, C., et al. 2008, PASP, 120, 405

Grier, C. J., Trump, J. R., Shen, Y., et al. 2017, ApJ, 851, 21

Grogin, N. A., Conselice, C. J., Chatzichristou, E., et al. 2005, ApJ, 627, L97

Grupe, D. 2004, Astron. J., 127, 1799

Grupe, D. 2011, in Narrow-Line Seyfert 1 Galaxies and their Place in the Universe, 4

Grupe, D., Beuermann, K., Mannheim, K., & Thomas, H. C. 1999, A&A, 350, 805

Grupe, D., Komossa, S., Leighly, K. M., & Page, K. L. 2010, ApJS, 187, 64

Guerras, E., Dai, X., Steele, S., et al. 2017, ApJ, 836, 206

Gunn, J. E. & Peterson, B. A. 1965, ApJ, 142, 1633

Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332

Gursky, H., Kellogg, E. M., Leong, C., Tananbaum, H., & Giacconi, R. 1971, ApJ, 165, L43

Guzman, J., Whiting, M., Voronkov, M., et al. 2019, ASKAPsoft: ASKAP science data processor software, Astrophysics Source Code Library, record ascl:1912.003

Haardt, F. & Maraschi, L. 1991, ApJ, 380, L51

Habouzit, M., Onoue, M., Bañados, E., et al. 2022, MNRAS, 511, 3751

Habouzit, M., Volonteri, M., & Dubois, Y. 2017, MNRAS, 468, 3935

Haggard, H. M. & Rovelli, C. 2015, Phys. Rev. D, 92, 104020

Haiman, Z. 2013, in Astrophysics and Space Science Library, Vol. 396, The First Galaxies, ed. T. Wiklind, B. Mobasher, & V. Bromm, 293

Haiman, Z. & Cen, R. 2001, in Astronomical Society of the Pacific Conference Series, Vol. 222, The Physics of Galaxy Formation, ed. M. Umemura & H. Susa, 101

Haiman, Z., Rees, M. J., & Loeb, A. 1997, ApJ, 476, 458

Haiman, Z., Thoul, A. A., & Loeb, A. 1996, ApJ, 464, 523

Hall, P. & York, M. 2001, Statistica Sinica, 11, 515

Häring, N. & Rix, H.-W. 2004, ApJ, 604, L89

Harris, D. E., Moldón, J., Oonk, J. R. R., et al. 2019, ApJ, 873, 21

Harrison, C. 2014, PhD thesis, Durham University, UK

Hasinger, G., Cappelluti, N., Brunner, H., et al. 2007, ApJS, 172, 29

Hasinger, G., Miyaji, T., & Schmidt, M. 2005, A&A, 441, 417

Haynes, W. 2013, Bonferroni Correction, ed. W. Dubitzky, O. Wolkenhauer, K.-H. Cho, & H. Yokota (New York, NY: Springer New York), 154–154

Hazard, C., Mackey, M. B., & Shimmins, A. J. 1963, Nature, 197, 1037

Heckman, T. M., Miley, G. K., van Breugel, W. J. M., & Butcher, H. R. 1981, ApJ, 247, 403

HI4PI Collaboration, Ben Bekhti, N., Flöer, L., et al. 2016, A&A, 594, A116

Hickox, R. C. & Alexander, D. M. 2018, ARA&A, 56, 625

Hirano, S., Hosokawa, T., Yoshida, N., & Kuiper, R. 2017, Science, 357, 1375

Hirano, S., Hosokawa, T., Yoshida, N., et al. 2014, ApJ, 781, 60

Hiroi, K., Ueda, Y., Akiyama, M., & Watson, M. G. 2012, ApJ, 758, 49

Ho, L. C. & Kim, M. 2014, ApJ, 789, 17

Hoang, D. N., Brüggen, M., Botteon, A., et al. 2022, A&A, 665, A60

Hönig, S. F. 2019, ApJ, 884, 171

Hönig, S. F. & Kishimoto, M. 2017, ApJ, 838, L20

Hönig, S. F., Kishimoto, M., Antonucci, R., et al. 2012, ApJ, 755, 149

Hönig, S. F., Kishimoto, M., Tristram, K. R. W., et al. 2013, ApJ, 771, 87

Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, ApJS, 163, 1

Hubble, E. P. 1926, ApJ, 64, 321

Ichikawa, T., Suzuki, R., Tokoku, C., et al. 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6269, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. I. S. McLean & M. Iye, 626916

Ichimaru, S. 1977, ApJ, 214, 840

Ighina, L., Moretti, A., Tavecchio, F., et al. 2022, A&A, 659, A93

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841

Inayoshi, K., Visbal, E., & Haiman, Z. 2020, ARA&A, 58, 27

Intema, H. T., Jagannathan, P., Mooley, K. P., & Frail, D. A. 2017, A&A, 598, A78

Ishimoto, R., Kashikawa, N., Onoue, M., et al. 2020, ApJ, 903, 60

Israel, W. 1967, Physical Review, 164, 1776

Ives, J. C., Sanford, P. W., & Penston, M. V. 1976, ApJ, 207, L159

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111

Iwamuro, F., Kimura, M., Eto, S., et al. 2004, ApJ, 614, 69

Jaffe, W., Meisenheimer, K., Röttgering, H. J. A., et al. 2004, Nature, 429, 47

Jahnke, K. & Macciò, A. V. 2011, ApJ, 734, 92

Jansen, F., Lumb, D., Altieri, B., et al. 2001, A&A, 365, L1

Jarvis, M. J. & McLure, R. J. 2006, MNRAS, 369, 182

Jiang, L., Fan, X., Vestergaard, M., et al. 2007, AJ, 134, 1150

Jiang, L., McGreer, I. D., Fan, X., et al. 2016, ApJ, 833, 222

Johnson, J. L. & Haardt, F. 2016, Publications of the Astronomical Society of Australia, 33, e007

Johnston, S., Taylor, R., Bailes, M., et al. 2008, Experimental Astronomy, 22, 151

Jolliffe, I. T. & Cadima, J. 2016, Philosophical Transactions of the Royal Society of London Series A, 374, 20150202

Joly, M., Véron-Cetty, M., & Véron, P. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 373, The Central Engine of Active Galactic Nuclei, ed. L. C. Ho & J.-W. Wang, 376

Juarez, Y., Maiolino, R., Mujica, R., et al. 2009, A&A, 494, L25

Just, D. W., Brandt, W. N., Shemmer, O., et al. 2007, ApJ, 665, 1004

Kaiser, N., Burgett, W., Chambers, K., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7733, Ground-based and Airborne Telescopes III, ed. L. M. Stepp, R. Gilmozzi, & H. J. Hall, 77330E

Kara, E., Alston, W. N., Fabian, A. C., et al. 2016, MNRAS, 462, 511

Kaspi, S., Smith, P. S., Netzer, H., et al. 2000, ApJ, 533, 631

Keating, S. K., Everett, J. E., Gallagher, S. C., & Deo, R. P. 2012, ApJ, 749, 32

Kellermann, K. I., Sramek, R., Schmidt, M., Shaffer, D. B., & Green, R. 1989, AJ, 98, 1195

Kelly, B. C. 2007, ApJ, 665, 1489

Kerr, R. P. 1963, Phys. Rev. Lett., 11, 237

Khachikian, E. Y. & Weedman, D. W. 1974, ApJ, 192, 581

Khorunzhev, G. A., Meshcheryakov, A. V., Medvedev, P. S., et al. 2021, Astronomy Letters, 47, 123

Khorunzhev, G. A., Sazonov, S. Y., & Burenin, R. A. 2018, Astronomy Letters, 44, 500

Khrykin, I. S., Hennawi, J. F., McQuinn, M., & Worseck, G. 2016, ApJ, 824, 133

Khrykin, I. S., Hennawi, J. F., Worseck, G., & Davies, F. B. 2021, MNRAS, 505, 649

Kim, D. W., Wilkes, B. J., Green, P. J., et al. 2004, ApJ, 600, 59

King, A. 2016, in Active Galactic Nuclei: What's in a Name?, 3

Kishimoto, M., Hönig, S. F., Antonucci, R., et al. 2011, A&A, 527, A121

Knight, C. A., Robertson, D. S., Rogers, A. E. E., et al. 1971, Science, 172, 52

Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv e-prints, arXiv:1711.03234

Kolodzig, A., Gilfanov, M., Sunyaev, R., Sazonov, S., & Brusa, M. 2013, A&A, 558, A89

Koptelova, E., Hwang, C.-Y., Malkan, M. A., & Yu, P.-C. 2019, ApJ, 882, 144

Koptelova, E., Hwang, C.-Y., Yu, P.-C., Chen, W.-P., & Guo, J.-K. 2017, Scientific Reports, 7, 41617

Koshida, S., Minezaki, T., Yoshii, Y., et al. 2014, ApJ, 788, 159

Kraft, R. P., Burrows, D. N., & Nousek, J. A. 1991, ApJ, 374, 344

Kuijken, K., Heymans, C., Dvornik, A., et al. 2019, A&A, 625, A2

Kulkarni, G., Worseck, G., & Hennawi, J. F. 2019, MNRAS, 488, 1035

Kuraszkiewicz, J., Wilkes, B. J., Schmidt, G., et al. 2009, ApJ, 692, 1180

Kurk, J. D., Walter, F., Fan, X., et al. 2009, ApJ, 702, 833

Kurk, J. D., Walter, F., Fan, X., et al. 2007, ApJ, 669, 32

Lai, S., Bian, F., Onken, C. A., et al. 2022, MNRAS, 513, 1801

Laird, E. S. et al. 2009, ApJS, 180, 102

LaMassa, S. M., Urry, C. M., Cappelluti, N., et al. 2013, MNRAS, 436, 3581

Laor, A. 1991, ApJ, 376, 90

Latif, M. A. & Ferrara, A. 2016, PASA, 33, e051

Latif, M. A., Omukai, K., Habouzit, M., Schleicher, D. R. G., & Volonteri, M. 2016, ApJ, 823, 40

Latif, M. A., Schleicher, D. R. G., Schmidt, W., & Niemeyer, J. C. 2013, MNRAS, 436, 2989

Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599

Leitherer, C., Li, I. H., Calzetti, D., & Heckman, T. M. 2002, ApJS, 140, 303

Lewis, K. T., Eracleous, M., & Storchi-Bergmann, T. 2010, ApJS, 187, 416

Li, Y.-R., Wang, J.-M., & Ho, L. C. 2012, ApJ, 749, 187

Liu, T., Buchner, J., Nandra, K., et al. 2022a, A&A, 661, A5

Liu, T., Merloni, A., Comparat, J., et al. 2022b, A&A, 661, A27

Lodato, G. & Natarajan, P. 2006, MNRAS, 371, 1813

Loeb, A. & Rasio, F. A. 1994, ApJ, 432, 52

López-Gonzaga, N., Burtscher, L., Tristram, K. R. W., Meisenheimer, K., & Schartmann, M. 2016, A&A, 591, A47

Luminet, J. P. 1979, A&A, 75, 228

Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, ApJS, 228, 2

Lusso, E., Comastri, A., Simmons, B. D., et al. 2012, MNRAS, 425, 623

Lusso, E., Comastri, A., Vignali, C., et al. 2010, A&A, 512, A34

Lusso, E. & Risaliti, G. 2016, ApJ, 819, 154

Lyke, B. W., Higley, A. N., McLane, J. N., et al. 2020, ApJS, 250, 8

Lynden-Bell, D. 1969, Nature, 223, 690

Macchetto, F., Marconi, A., Axon, D. J., et al. 1997, ApJ, 489, 579

Madau, P. & Rees, M. J. 2000, ApJ, 542, L69

Madau, P. & Rees, M. J. 2001, ApJ, 551, L27

Madsen, K. K., Fürst, F., Walton, D. J., et al. 2015, ApJ, 812, 14

Magaudda, E., Stelzer, B., Raetz, S., et al. 2022, A&A, 661, A29

Magdziarz, P., Blaes, O. M., Zdziarski, A. A., Johnson, W. N., & Smith, D. A. 1998, MNRAS, 301, 179

Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285

Malizia, A., Molina, M., Bassani, L., et al. 2014, ApJ, 782, L25

Manti, S., Gallerani, S., Ferrara, A., Greig, B., & Feruglio, C. 2017, MNRAS, 466, 1160

Mao, Y.-F., Wang, J., & Wei, J.-Y. 2009, Research in Astronomy and Astrophysics, 9, 529

Marchesi, S., Lanzuisi, G., Civano, F., et al. 2016, ApJ, 830, 100

Marconi, A. & Hunt, L. K. 2003, ApJ, 589, L21

Marinello, M., Rodríguez-Ardila, A., Garcia-Rissmann, A., Sigut, T. A. A., & Pradhan, A. K. 2016, ApJ, 820, 116

Marocco, F., Eisenhardt, P. R. M., Fowler, J. W., et al. 2021, ApJS, 253, 8

Marshall, N., Warwick, R. S., & Pounds, K. A. 1981, MNRAS, 194, 987

Marziani, P., del Olmo, A., D'Onofrio, M., et al. 2018, in Revisiting Narrow-Line Seyfert 1 Galaxies and their Place in the Universe, 2

Marziani, P., Dultzin, D., Sulentic, J. W., et al. 2018, Frontiers in Astronomy and Space Sciences, 5, 6

Marziani, P., Negrete, C. A., Dultzin, D., et al. 2017, Frontiers in Astronomy and Space Sciences, 4, 16

Marziani, P., Sulentic, J. W., Negrete, C. A., et al. 2010, MNRAS, 409, 1033

Marziani, P., Sulentic, J. W., Plauchu-Frayn, I., & del Olmo, A. 2013a, A&A, 555, A89

Marziani, P., Sulentic, J. W., Plauchu-Frayn, I., & del Olmo, A. 2013b, ApJ, 764, 150

Marziani, P., Sulentic, J. W., Stirpe, G. M., Zamfir, S., & Calvani, M. 2009, A&A, 495, 83

Marziani, P., Sulentic, J. W., Zamanov, R., et al. 2003a, ApJS, 145, 199

Marziani, P., Sulentic, J. W., Zwitter, T., Dultzin-Hacyan, D., & Calvani, M. 2001, ApJ, 558, 553

Marziani, P., Zamanov, R. K., Sulentic, J. W., & Calvani, M. 2003b, MNRAS, 345, 1133

Matsuoka, Y., Iwasawa, K., Onoue, M., et al. 2022, ApJS, 259, 18

Matsuoka, Y., Iwasawa, K., Onoue, M., et al. 2019, ApJ, 883, 183

Matsuoka, Y., Iwasawa, K., Onoue, M., et al. 2018a, ApJS, 237, 5

Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2018b, PASJ, 70, S35

Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2018, Publications of the Astronomical Society of Japan, 70, S35

Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2016, ApJ, 828, 26

Matt, G., Guainazzi, M., Frontera, F., et al. 1997, A&A, 325, L13

Matt, G., Perola, G. C., & Piro, L. 1991, A&A, 247, 25

Matthews, T. A. & Sandage, A. R. 1963, ApJ, 138, 30

Mazzucchelli, C., Bañados, E., Venemans, B. P., et al. 2017, ApJ, 849, 91

McGreer, I. D., Mesinger, A., & D'Odorico, V. 2015, MNRAS, 447, 499

McLure, R. J. & Jarvis, M. J. 2002, MNRAS, 337, 109

McMahon, R. G., Banerji, M., Gonzalez, E., et al. 2013, The Messenger, 154, 35

Medvedev, P., Gilfanov, M., Sazonov, S., Schartel, N., & Sunyaev, R. 2021, MNRAS, 504, 576

Medvedev, P., Sazonov, S., Gilfanov, M., et al. 2020, MNRAS, 497, 1842

Meisenheimer, K., Tristram, K. R. W., Jaffe, W., et al. 2007, A&A, 471, 453

Mejia-Restrepo, J., Trakhtenbrot, B., Lira, P., Netzer, H., & Capellupo, D. 2016, in Active Galactic Nuclei: What's in a Name?, 7

Mejía-Restrepo, J. E., Trakhtenbrot, B., Lira, P., & Netzer, H. 2018, MNRAS, 478, 1929

Menzel, M.-L., Merloni, A., Georgakakis, A., et al. 2016, MNRAS, 457, 110

Merloni, A. & Heinz, S. 2008, MNRAS, 388, 1011

Merloni, A., Predehl, P., Becker, W., et al. 2012, arXiv e-prints, arXiv:1209.3114

Micic, M., Holley-Bockelmann, K., Sigurdsson, S., & Abel, T. 2007, MNRAS, 380, 1533

Middei, R., Bianchi, S., Marinucci, A., et al. 2019, A&A, 630, A131

Miller, J. M. 2007, ARA&A, 45, 441

Miller, J. S., Goodrich, R. W., & Mathews, W. G. 1991, ApJ, 378, 47

Mingaliev, M. G., Sotnikova, Y. V., Torniainen, I., Tornikoski, M., & Udovitskiy, R. Y. 2012, A&A, 544, A25

Miyaji, T., Hasinger, G., Salvato, M., et al. 2015, ApJ, 804, 104

Miyazaki, S., Komiyama, Y., Kawanomoto, S., et al. 2018, PASJ, 70, S1

Mortlock, D. J., Patel, M., Warren, S. J., et al. 2012, MNRAS, 419, 390, publisher: Oxford Academic

Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, Nature, 474, 616

Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, ApJ, 731, 113

Murray, N., Chiang, J., Grossman, S. A., & Voit, G. M. 1995, ApJ, 451, 498

Nandra, K., George, I. M., Mushotzky, R. F., Turner, T. J., & Yaqoob, T. 1997, ApJ, 476, 70

Nandra, K., Laird, E. S., Aird, J. A., et al. 2015, ApJS, 220, 10

Nandra, K., O'Neill, P. M., George, I. M., & Reeves, J. N. 2007, MNRAS, 382, 194

Nandra, K. & Pounds, K. A. 1994, MNRAS, 268, 405

Nanni, R., Gilli, R., Vignali, C., et al. 2018, A&A, 614, A121

Nanni, R., Vignali, C., Gilli, R., Moretti, A., & Brandt, W. N. 2017, A&A, 603, A128

Narayan, R. & Yi, I. 1994, ApJ, 428, L13

Narayan, R. & Yi, I. 1995a, ApJ, 444, 231

Narayan, R. & Yi, I. 1995b, ApJ, 452, 710

Negrete, C. A., Dultzin, D., Marziani, P., et al. 2018, A&A, 620, A118

Nelder, J. A. & Wedderburn, R. W. M. 1972, Journal of the Royal Statistical Society: Series A (General), 135, 370

Nemmen, R. S., Storchi-Bergmann, T., & Eracleous, M. 2014, MNRAS, 438, 2804

Nenkova, M., Ivezić, Ž., & Elitzur, M. 2002, ApJ, 570, L9

Nenkova, M., Sirocky, M. M., Ivezić, Ž., & Elitzur, M. 2008a, ApJ, 685, 147

Nenkova, M., Sirocky, M. M., Nikutta, R., Ivezić, Ž., & Elitzur, M. 2008b, ApJ, 685, 160

Netzer, H. 2008, New Astronomy Reviews, 52, 257

Netzer, H. 2015, ARA&A, 53, 365

Netzer, H. & Trakhtenbrot, B. 2007, ApJ, 654, 754

Newman, E. T., Couch, E., Chinnapared, K., et al. 1965, Journal of Mathematical Physics, 6, 918

Ni, Y., Di Matteo, T., Bird, S., et al. 2022, MNRAS, 513, 670

Nordström, G. 1918, Koninklijke Nederlandse Akademie van Wetenschappen Proceedings Series B Physical Sciences, 20, 1238

O'Dea, C. P. 1998, PASP, 110, 493

O'Dea, C. P. & Saikia, D. J. 2020, arXiv e-prints, arXiv:2009.02750

Ohsuga, K., Mori, M., Nakamoto, T., & Mineshige, S. 2005, ApJ, 628, 368

Ojha, V., Chand, H., Dewangan, G. C., & Rakshit, S. 2020, ApJ, 896, 95

Oke, J. B. 1963, Nature, 197, 1040

Oke, J. B. 1987, in Superluminal Radio Sources, ed. J. A. Zensus & T. J. Pearson, 267–272

Omukai, K. 2001, ApJ, 546, 635

Omukai, K., Schneider, R., & Haiman, Z. 2008, ApJ, 686, 801

Onken, C. A., Ferrarese, L., Merritt, D., et al. 2004, ApJ, 615, 645

Onoue, M., Kashikawa, N., Matsuoka, Y., et al. 2019, ApJ, 880, 77

Orienti, M. 2016, Astronomische Nachrichten, 337, 9

Orienti, M. & Dallacasa, D. 2008, A&A, 477, 807

Orienti, M., Dallacasa, D., & Stanghellini, C. 2007, A&A, 475, 813

Osterbrock, D. E. & Dahari, O. 1983, ApJ, 273, 478

Osterbrock, D. E. & Pogge, R. W. 1985, ApJ, 297, 166

Packham, C., Radomski, J. T., Roche, P. F., et al. 2005, ApJ, 618, L17

Pacucci, F. & Loeb, A. 2022, MNRAS, 509, 1885

Padovani, P. & Urry, C. M. 1992, ApJ, 387, 449

Padovani, P. et al. 2017, A&A Rev., 25, 2

Page, M. J. & Carrera, F. J. 2000, MNRAS, 311, 433, publisher: Oxford Academic

Panda, S., Czerny, B., Adhikari, T. P., et al. 2018, ApJ, 866, 115

Panda, S., Czerny, B., & Wildy, C. 2017, Frontiers in Astronomy and Space Sciences, 4, 33

Panda, S., Marziani, P., & Czerny, B. 2019a, arXiv e-prints, arXiv:1912.03119

Panda, S., Marziani, P., & Czerny, B. 2019b, arXiv e-prints, arXiv:1912.03118

Panda, S., Marziani, P., & Czerny, B. 2019c, arXiv e-prints, arXiv:1905.01729

Pâris, I., Petitjean, P., Rollinde, E., et al. 2011, A&A, 530, A50

Pavlinsky, M., Tkachenko, A., Levin, V., et al. 2021, A&A, 650, A42

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Penrose, R. 1963, Phys. Rev. Lett., 10, 66

Perola, G. C., Matt, G., Cappi, M., et al. 2002, A&A, 389, 802

Perrotta, S., Hamann, F., Zakamska, N. L., et al. 2019, arXiv e-prints, arXiv:1906.00980

Peterson, B. 2006, The Broad-Line Region in Active Galactic Nuclei, ed. D. Alloin, R. Johnson, & P. Lira (Berlin, Heidelberg: Springer Berlin Heidelberg), 77–100

Peterson, B. M. 1993, PASP, 105, 247

Peterson, B. M. 2010, in Co-Evolution of Central Black Holes and Galaxies, ed. B. M. Peterson, R. S. Somerville, & T. Storchi-Bergmann, Vol. 267, 151–160

Peterson, B. M. & Ferland, G. J. 1986, Nature, 324, 345

Peterson, B. M. & Wandel, A. 1999, ApJ, 521, L95

Petric, A. O., Carilli, C. L., Bertoldi, F., et al. 2003, The Astronomical Journal, 126, 15

Pierre, M., Pacaud, F., Adami, C., et al. 2016, A&A, 592, A1

Pineau, F. X., Derriere, S., Motch, C., et al. 2017, A&A, 597, A89

Pipien, S., Cuby, J. G., Basa, S., et al. 2018, A&A, 617, A127

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6

Pons, E., McMahon, R. G., Banerji, M., & Reed, S. L. 2020, MNRAS, 491, 3884

Popovic, L., G. Mediavilla, E., A, K., & Jovanović, P. 2002, A&A, 390, 473

Popović, L. Č., Mediavilla, E., Bon, E., & Ilić, D. 2004, A&A, 423, 909

Pott, J.-U., Malkan, M. A., Elitzur, M., et al. 2010, ApJ, 715, 736

Potter, W. J. & Cotter, G. 2012, MNRAS, 423, 756

Pounds, K. A., Done, C., & Osborne, J. P. 1995, MNRAS, 277, L5

Predehl, P., Andritschke, R., Arefiev, V., et al. 2021, A&A, 647, A1

Prieto, M. A., Reunanen, J., Tristram, K. R. W., et al. 2010, MNRAS, 402, 724

Raban, D., Jaffe, W., Röttgering, H., Meisenheimer, K., & Tristram, K. R. W. 2009, MNRAS, 394, 1325

Radomski, J. T., Packham, C., Levenson, N. A., et al. 2008, ApJ, 681, 141

Rakshit, S., Stalin, C. S., Chand, H., & Zhang, X.-G. 2017, ApJS, 229, 39

Rakshit, S., Stalin, C. S., Kotilainen, J., & Shin, J. 2021, ApJS, 253, 28

Rakshit, S. & Woo, J.-H. 2018, ApJ, 865, 5

Ramos Almeida, C., Levenson, N. A., Rodríguez Espinosa, J. M., et al. 2009, ApJ, 702, 1127

Ramos Almeida, C. & Ricci, C. 2017, Nature Astronomy, 1, 679

Ranalli, P., Comastri, A., Vignali, C., et al. 2013, A&A, 555, A42

Reed, S. L., Banerji, M., Becker, G. D., et al. 2019, MNRAS, 487, 1874

Reed, S. L., McMahon, R. G., Banerji, M., et al. 2015, MNRAS, 454, 3952

Rees, M. J., Begelman, M. C., Blandford, R. D., & Phinney, E. S. 1982, Nature, 295, 17

Rees, M. J., Netzer, H., & Ferland, G. J. 1989, ApJ, 347, 640

Reeves, J. N., Turner, M. J. L., Ohashi, T., & Kii, T. 1997, MNRAS, 292, 468

Regan, J. A., Johansson, P. H., & Wise, J. H. 2014, ApJ, 795, 137

Reissner, H. 1916, Annalen der Physik, 355, 106

Reynolds, C. S. & Fabian, A. C. 2008, ApJ, 675, 1048

Ricci, C., Ho, L. C., Fabian, A. C., et al. 2018, MNRAS, 480, 1819

Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006, ApJS, 166, 470

Richards, G. T., Vanden Berk, D. E., Reichard, T. A., et al. 2002, AJ, 124, 1

Robertson, B. E. 2010, ApJ, 716, L229

Robinson, D. C. 1975, Phys. Rev. Lett., 34, 905

Roche, P. F., Aitken, D. K., Smith, C. H., & Ward, M. J. 1991, MNRAS, 248, 606

Rodríguez, M. 1999, A&A, 348, 222

Rodríguez-Ardila, A., Viegas, S. M., Pastoriza, M. G., & Prato, L. 2002, ApJ, 565, 140

Romano, P., Turner, T. J., Mathur, S., & George, I. M. 2002, ApJ, 564, 162

Rosen, S. R., Webb, N. A., Watson, M. G., et al. 2016, A&A, 590, A1

Ross, R. R. & Fabian, A. C. 2005, MNRAS, 358, 211

Runnoe, J. C., Brotherton, M. S., & Shang, Z. 2012, MNRAS, 422, 478

Rybicki, G. B. & Lightman, A. P. 1979, Radiative processes in astrophysics

Ryle, M., Smith, F. G., & Elsmore, B. 1950, MNRAS, 110, 508

Salpeter, E. E. 1964, ApJ, 140, 796

Salvato, M., Buchner, J., Budavári, T., et al. 2018, MNRAS, 473, 4937

Salvato, M., Ilbert, O., & Hoyle, B. 2019a, Nature Astronomy, 3, 212

Salvato, M., Ilbert, O., & Hoyle, B. 2019b, Nature Astronomy, 3, 212

Salvato, M., Wolf, J., Dwelly, T., et al. 2022, A&A, 661, A3

Saxton, R. D., Read, A. M., Esquej, P., et al. 2008, A&A, 480, 611

Schlafly, E. F., Meisner, A. M., & Green, G. M. 2019, ApJS, 240, 30

Schmidt, M. 1968, ApJ, 151, 393

Schwartz, D. A. 2002, ApJ, 569, L23

Schwarzschild, K. 1916, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 189

Scott, A. E., Stewart, G. C., & Mateos, S. 2012, MNRAS, 423, 2633

Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1

Seabold, S. & Perktold, J. 2010, in 9th Python in Science Conference

Seppi, R., Comparat, J., Bulbul, E., et al. 2022, A&A, 665, A78

Seyfert, C. K. 1943, ApJ, 97, 28

Shakura, N. I. & Sunyaev, R. A. 1973, A&A, 24, 337

Shang, C., Bryan, G. L., & Haiman, Z. 2010, MNRAS, 402, 1249

Shang, Z., Brotherton, M. S., Wills, B. J., et al. 2011, ApJS, 196, 2

Shang, Z., Wills, B. J., Robinson, E. L., et al. 2003, ApJ, 586, 52

Sharov, A. S. & Efremov, Y. N. 1963, Information Bulletin on Variable Stars, 23, 1

Shastri, P., Wilkes, B. J., Elvis, M., & McDowell, J. 1993, ApJ, 410, 29

Shemmer, O., Brandt, W. N., Netzer, H., Maiolino, R., & Kaspi, S. 2006, ApJ, 646, L29

Shen, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2020, MNRAS, 495, 3252

Shen, Y. 2013, Bull. Astron. Soc. India, 41, 61

Shen, Y. & Ho, L. C. 2014, Nature, 513, 210

Shen, Y. & Liu, X. 2012, ApJ, 753, 125

Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, ApJS, 194, 45

Shen, Y., Wu, J., Jiang, L., et al. 2019, ApJ, 873, 35

Shi, Y., Rieke, G. H., Hines, D. C., et al. 2006, ApJ, 653, 127

Shields, G. A. 1974, ApJ, 191, 309

Shimwell, T. W., Röttgering, H. J. A., Best, P. N., et al. 2017, A&A, 598, A104

Shimwell, T. W., Tasse, C., Hardcastle, M. J., et al. 2019, A&A, 622, A1

Sicilian, D., Civano, F., Cappelluti, N., Buchner, J., & Peca, A. 2022, ApJ, 936, 39

Siemiginowska, A., LaMassa, S., Aldcroft, T. L., Bechtold, J., & Elvis, M. 2008, ApJ, 684, 811

Sigut, T. A. A. & Pradhan, A. K. 1998, ApJ, 499, L139

Sigut, T. A. A. & Pradhan, A. K. 2003, ApJS, 145, 15

Sijacki, D., Vogelsberger, M., Genel, S., et al. 2015, MNRAS, 452, 575

Silverman, B. W. 1981, Journal of the Royal Statistical Society. Series B (Methodological), 43, 97

Silverman, B. W. 1986, Density Estimation for Statistics and Data Analysis (London: Chapman & Hall)

Silverman, J. D., Green, P. J., Barkhouse, W. A., et al. 2005, ApJ, 624, 630

Simmonds, C., Buchner, J., Salvato, M., Hsu, L. T., & Bauer, F. E. 2018, A&A, 618, A66

Singh, K. P., Garmire, G. P., & Nousek, J. 1985, ApJ, 297, 633

Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, AJ, 146, 32

Smirnov, O. M. & Tasse, C. 2015, MNRAS, 449, 2668

Smith, F. G. 1952, MNRAS, 112, 497

Soffrin, S. 1969, A&A, 1, 305

Soltan, A. 1982, MNRAS, 200, 115

Songaila, A. 2004, AJ, 127, 2598

Souffrin, S. 1969, A&A, 1, 414

Springel, V., Di Matteo, T., & Hernquist, L. 2005, MNRAS, 361, 776

Stark, A. A., Gammie, C. F., Wilson, R. W., et al. 1992, ApJS, 79, 77

Steffen, A. T., Strateva, I., Brandt, W. N., et al. 2006, AJ, 131, 2826

Stern, D., Hall, P. B., Barrientos, L. F., et al. 2003, ApJ, 596, L39

Stoica, O. C. 2018, in 2nd Karl Schwarzschild Meeting on Gravitational Physics, ed. P. Nicolini, M. Kaminski, J. Mureika, & M. Bleicher (Cham: Springer International Publishing), 75–81

Storchi-Bergmann, T., Schimoia, J., Peterson, B., et al. 2016, ApJ, 835

Strateva, I. V., Brandt, W. N., Schneider, D. P., Vanden Berk, D. G., & Vignali, C. 2005, AJ, 130, 387

Sturm, E., Dexter, J., Pfuhl, O., et al. 2018, Nature, 563

Suganuma, M., Yoshii, Y., Kobayashi, Y., et al. 2006, ApJ, 639, 46

Sulentic, J., Marziani, P., & Zamfir, S. 2011, Baltic Astronomy, 20, 427

Sulentic, J. W., Bachev, R., Marziani, P., Negrete, C. A., & Dultzin, D. 2007a, ApJ, 666, 757

Sulentic, J. W., del Olmo, A., Marziani, P., et al. 2017, A&A, 608, A122

Sulentic, J. W., Dultzin-Hacyan, D., & Marziani, P. 2007b, in Revista Mexicana de Astronomia y Astrofisica, vol. 27, Vol. 28, Revista Mexicana de Astronomia y Astrofisica Conference Series, ed. S. Kurtz, 83–88

Sulentic, J. W. & Marziani, P. 2015, Frontiers in Astronomy and Space Sciences, 2, 6

Sulentic, J. W., Marziani, P., & Dultzin-Hacyan, D. 2000a, Annual Review of Astronomy and Astrophysics, 38, 521

Sulentic, J. W., Marziani, P., Zamanov, R., et al. 2002, ApJ, 566, L71

Sulentic, J. W., Zwitter, T., Marziani, P., & Dultzin-Hacyan, D. 2000b, ApJ, 536, L5

Sulentic, J. W., Zwitter, T., Marziani, P., & Dultzin-Hacyan, D. 2000c, ApJ, 536, L5

Sun, J. & Shen, Y. 2015, ApJ, 804, L15

Sun, W.-H. & Malkan, M. A. 1989, ApJ, 346, 68

Sunyaev, R., Arefiev, V., Babyshkin, V., et al. 2021, A&A, 656, A132

Sutherland, W. & Saunders, W. 1992, MNRAS, 259, 413

Suzuki, N., Tytler, D., Kirkman, D., O'Meara, J. M., & Lubin, D. 2005, ApJ, 618, 592

Suzuki, R., Tokoku, C., Ichikawa, T., et al. 2008, PASJ, 60, 1347

Swain, M., Vasisht, G., Akeson, R., et al. 2003, ApJ, 596, L163

Takahashi, H. R. & Ohsuga, K. 2015, PASJ, 67, 60

Tamura, N., Ohta, K., & Ueda, Y. 2006, MNRAS, 365, 134

Tanaka, T. L. & Li, M. 2014, MNRAS, 439, 1092

Tananbaum, H., Avni, Y., Branduardi, G., et al. 1979, ApJ, 234, L9

Tang, B., Shang, Z., Gu, Q., Brotherton, M. S., & Runnoe, J. C. 2012, ApJS, 201, 38

Tasse, C. 2014a, arXiv e-prints, arXiv:1410.8706

Tasse, C. 2014b, A&A, 566, A127

Tasse, C., Hugo, B., Mirmont, M., et al. 2018, A&A, 611, 1

Tavecchio, F., Maraschi, L., Sambruna, R. M., & Urry, C. M. 2000, ApJ, 544, L23

Tegmark, M., Silk, J., Rees, M. J., et al. 1997, ApJ, 474, 1

Thorne, K. S. 1974, ApJ, 191, 507

Tingay, S. J. & de Kool, M. 2003, AJ, 126, 723

Toyouchi, D., Inayoshi, K., Li, W., Haiman, Z., & Kuiper, R. 2023, MNRAS, 518, 1601

Trakhtenbrot, B. & Netzer, H. 2012, MNRAS, 427, 3081

Trenti, M. & Stiavelli, M. 2008, ApJ, 676, 767

Tristram, K. R. W., Burtscher, L., Jaffe, W., et al. 2014, A&A, 563, A82

Tristram, K. R. W., Meisenheimer, K., Jaffe, W., et al. 2007, A&A, 474, 837

Truemper, J. 1982, Advances in Space Research, 2, 241

Tsuzuki, Y., Kawara, K., Yoshii, Y., et al. 2006, ApJ, 650, 57

Turner, T. J. & Pounds, K. A. 1989, MNRAS, 240, 833

Turner, T. J., Romano, P., George, I. M., et al. 2001, ApJ, 561, 131

Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, ApJ, 786, 104

Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, ApJ, 598, 886

Urry, C. M. & Padovani, P. 1995, PASP, 107, 803

Uttley, P., Cackett, E. M., Fabian, A. C., Kara, E., & Wilkins, D. R. 2014, A&A Rev., 22, 72

van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2

van Weeren, R. J., Williams, W. L., Hardcastle, M. J., et al. 2016, ApJS, 223, 2

Venemans, B. P., Findlay, J. R., Sutherland, W. J., et al. 2013, ApJ, 779, 24, publisher: IOP Publishing

Venemans, B. P., Verdoes Kleijn, G. A., Mwebaze, J., et al. 2015, MNRAS, 453, 2259

Vestergaard, M. 2002, ApJ, 571, 733

Vestergaard, M. & Osmer, P. S. 2009, ApJ, 699, 800

Vestergaard, M. & Peterson, B. M. 2006, ApJ, 641, 689

Vignali, C., Bauer, F. E., Alexander, D. M., et al. 2002, ApJ, 580, L105

Vignali, C., Brandt, W. N., Schneider, D. P., Garmire, G. P., & Kaspi, S. 2003, AJ, 125, 418

Vito, F., Brandt, W. N., Bauer, F. E., et al. 2019a, A&A, 630, A118

Vito, F., Brandt, W. N., Bauer, F. E., et al. 2019b, A&A, 628, L6

Vito, F., Brandt, W. N., Yang, G., et al. 2018, MNRAS, 473, 2378

Vito, F., Gilli, R., Vignali, C., et al. 2016, MNRAS, 463, 348

Vito, F., Gilli, R., Vignali, C., et al. 2014, MNRAS, 445, 3557

Voges, W., Aschenbach, B., Boller, T., et al. 1999, A&A, 349, 389

Voges, W. et al. 1999, A&A, 349, 389

Volonteri, M. 2010, A&A, 18, 279

Volonteri, M., Dubois, Y., Pichon, C., & Devriendt, J. 2016, MNRAS, 460, 2979

Volonteri, M., Habouzit, M., & Colpi, M. 2021, Nature Reviews Physics, 3, 732

Waddell, S. G. H. & Gallo, L. C. 2020, MNRAS, 498, 5207

Waller, L. & Turnbull, B. 1992, American Statistician - AMER STATIST, 46, 5

Walton, D. J., Nardini, E., Fabian, A. C., Gallo, L. C., & Reis, R. C. 2013, MNRAS, 428, 2901

Wandel, A., Peterson, B. M., & Malkan, M. A. 1999, ApJ, 526, 579

Wanders, I. & Peterson, B. M. 1996, ApJ, 466, 174

Wang, F., Fan, X., Yang, J., et al. 2017, ApJ, 839, 27

Wang, F., Yang, J., Fan, X., et al. 2021, ApJ, 907, L1

Wang, F., Yang, J., Fan, X., et al. 2019, ApJ, 884, 30

Wang, J., Xu, D. W., & Wei, J. Y. 2018, ApJ, 852, 26

Wang, J.-M., Du, P., Hu, C., et al. 2014, ApJ, 793, 108

Watson, M. G., Schröder, A. C., Fyfe, D., et al. 2009, A&A, 493, 339

Webb, N. A., Coriat, M., Traulsen, I., et al. 2020, A&A, 641, A136

Webster, B. L. & Murdin, P. 1972, Nature, 235, 37

Weigelt, G., Hofmann, K. H., Kishimoto, M., et al. 2012, A&A, 541, L9

Weinberger, R., Springel, V., Pakmor, R., et al. 2018, MNRAS, 479, 4056

Weisskopf, M. C., Brinkman, B., Canizares, C., et al. 2002, PASP, 114, 1

Weisskopf, M. C., Wu, K., Trimble, V., et al. 2007, ApJ, 657, 1026

Wenzl, L., Schindler, J.-T., Fan, X., et al. 2021, AJ, 162, 72

White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D. 1997, ApJ, 475, 479

Whiting, M. & Humphreys, B. 2012, Publications of the Astronomical Society of Australia, 29, 371–381

Whitney, A. R., Shapiro, I. I., Rogers, A. E. E., et al. 1971, Science, 173, 225

Wilkes, B. J. & Elvis, M. 1987, ApJ, 323, 243

Williams, W. L., van Weeren, R. J., Röttgering, H. J. A., et al. 2016, MNRAS, 460, 2385

Willott, C. J., Albert, L., Arzoumanian, D., et al. 2010, AJ, 140, 546

Willott, C. J., Delorme, P., Omont, A., et al. 2007, AJ, 134, 2435

Willott, C. J., Delorme, P., Reylé, C., et al. 2009, AJ, 137, 3541

Wills, B. J. & Browne, I. W. A. 1986, ApJ, 302, 56

Wills, B. J., Netzer, H., & Wills, D. 1985, ApJ, 288, 94

Wilms, J., Allen, A., & McCray, R. 2000, ApJ, 542, 914

Winkler, H. & Chauke, T. 2014, ArXiv e-prints [`arXiv:1412.4498`]

Winkler, P. F., J. & White, A. E. 1975, ApJ, 199, L139

Wise, J. H. 2019, arXiv e-prints, arXiv:1907.06653

Wolf, J., Nandra, K., Salvato, M., et al. 2023, A&A, 669, A127

Wolf, J., Nandra, K., Salvato, M., et al. 2021, A&A, 647, A5

Wolf, J., Salvato, M., Coffey, D., et al. 2020, MNRAS, 492, 3580

Woltjer, L. 1959, ApJ, 130, 38

Woo, J.-H. & Urry, C. M. 2002, ApJ, 579, 530

Wright, A. H., Hildebrandt, H., Kuijken, K., et al. 2019, A&A, 632, A34

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868

Wu, X.-B., Wang, F., Fan, X., et al. 2015, Nature, 518, 512

Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, ApJS, 195, 10

Yang, G., Boquien, M., Buat, V., et al. 2020, MNRAS, 491, 740

Yang, J., Fan, X., Wang, F., et al. 2022, ApJ, 924, L25

Yang, J., Wang, F., Fan, X., et al. 2021, ApJ, 923, 262

York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579

Yoshida, N., Abel, T., Hernquist, L., & Sugiyama, N. 2003, ApJ, 592, 645

Yuan, M. J. & Wills, B. J. 2003, ApJ, 593, L11

Yue, B., Ferrara, A., Pacucci, F., & Omukai, K. 2017, ApJ, 838, 111

Zakamska, N. L., Hamann, F., Pâris, I., et al. 2016, MNRAS, 459, 3144

Zamanov, R., Marziani, P., Sulentic, J. W., et al. 2002, ApJ, 576, L9

Zamfir, S., Sulentic, J. W., Marziani, P., & Dultzin, D. 2010, MNRAS, 403, 1759

Zel'dovich, Y. B. 1964, Soviet Physics Doklady, 9, 195

Zel'dovich, Y. B. & Novikov, I. D. 1967, Soviet Ast., 10, 602

Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, Research in Astronomy and Astrophysics, 12, 723

Zhou, X.-L. & Zhang, S.-N. 2010, ApJ, 713, L11

Zhu, S. F., Brandt, W. N., Luo, B., et al. 2020, MNRAS, 496, 245

Zoghbi, A., Fabian, A. C., Uttley, P., et al. 2010, MNRAS, 401, 2419

Acknowledgements

"Not all those who wander are lost."
J.R.R. Tolkien

I hope the reader of these last pages will indulge me in a final citation from an epic tale to appropriately close this chapter of my life. It has taken some years to get here, and at times, I did meander around in a labyrinth of new concepts and impossible choices. However, the end of the project, ever at a distant and quasi-virtual future, is now within reach. I want to acknowledge the helping hands, the encouraging words and the wise mentorship that made this work possible.

I would like to express my deepest gratitude to my mentor, Dr Mara Salvato, who has always found the right words to keep me focused and proactive. Working with Dr Salvato has been a privilege and has laid a strong foundation for my academic career. I greatly appreciated our work as a team, which was rooted in mutual trust and respect. Beyond patiently initiating me to the technical aspects of the job and sharing her deep knowledge of survey astronomy in a language I could understand, she has taught me the value of intuition and common sense when facing complex problems. I am forever grateful for her unyielding support.

I am deeply indebted to Prof. Dr Kirpal Nandra for being a patient supervisor and for always keeping a watchful eye on my scientific wanderings. Under his guidance, I have developed a fascination for the very early Universe, which I hope to further explore in my career. I am immensely grateful for the trust he placed in my work.

I am very grateful to the members of my PhD examination committee, Prof. Dr Jochen Weller, Prof. Dr Gerhard Buchalla, and Dr Ariel Sánchez, for their time and for helping me complete this journey. I am particularly grateful to Prof. Weller, who made me believe in this career path early on and under whose mentorship I have grown as an aspiring scientist during my bachelor and master studies at the LMU. I would also like to extend my gratitude to my master's thesis supervisor Dr Ben Hoyle for his guidance on my early work in cosmology, helping me improve my coding skills and introducing me to modern deep learning models. Many thanks also to Dr Paolo Padovani for sharing his deep knowledge of AGN over the last years as a member of my thesis committee.

My time at the High-Energy Group of the MPE was an enriching and genuinely stimulating period. I am grateful to the scientific and technical staff, postdocs and students for upholding a great team mentality and open-door policy. I am particularly grateful to the entire eROSITA team, from operations to data analysis, at MPE and in the broader collaboration, for enabling the research in this work.

Very special thanks to:

Dr Andrea Merloni, who, as a leader of the eROSITA project, has always enabled early career researchers of the HEG to optimally benefit from the mission's great success. His consistently inspiring and relevant insights on AGN physics and astronomy at large have truly shaped the work presented in this thesis.

Dr Johannes Buchner, for always being willing (and able) to help. I thank him for very helpful insights into astrostatistics and for encouraging creative approaches when tackling scientific or data-related problems.

Dr Tom Dwelly, Dr Arne Rau, Dr Teng Liu and Dr Johan Comparat for the help they were willing to provide on my various projects.

Dr Peter Predehl for leading the development of the groundbreaking eROSITA instrument and its all-sky survey.

Dr Stefano Ciroi and Dr Francesco Di Mille for playing a pivotal role in our discovery of distant quasars by confirming them with the Magellan telescopes at the Las Campanas Observatory. And sorry for all the brown dwarfs!

Dr Tanya Urrutia for her invaluable help with the interpretation of our optical and NIR spectroscopic data.

Dr Marcella Brusa for helping with the interpretation of our work on high-z AGN space densities.

Dr Eduardo Bañados for stimulating discussions on the quasar discovery programme presented at the end of this thesis.

Dr Helmut Steinle for introducing me to hands-on astronomical observations at the ESO/MPG 2.2m telescope in La Silla.

Thanks should also go to Dr Riccardo Arcodia and Dr Adam Malyali. Their feedback as fellow students and esteemed colleagues has been crucial in the success of this project. I have also enjoyed our *Bierspaziergänge* to Little Berlin with the dog and our gourmet discoveries (a special mention to the upside down tzatziki, *blabla?*). Thank you for not abandoning me on top of that rock in Athens.

I am grateful to Dr Joseph O'Leary, who often provided the cosmological "big picture" and precious career advice. I am counting on you as pâtissier when we open our French/Chicken Karaage/Tacos restaurant in a parallel universe or distant future.

I thank Dr Damien Coffey for tutoring my first steps in AGN research and for the passionate discussions on Fender Custom Shop guitars and the history of the Roman-Gallic Wars.

I am grateful to Sophia Waddell, Dr Riccardo Seppi, Dr Jacob Ider Chitham, Dr Felix Widmann, Martin Mayer, Yrupe Fresco, Dr Nicola Locatelli, Toska, Dr Sebastiano von

Fellenberg, Zsofi Igo, William Roster and many more for the helpful discussions on science and the daily life of early career researchers.

I want to express my gratitude to Birgit Boller for making the administrative aspects of the thesis smooth sailing. Thanks also to Annette Hilbert for her kind support.

I thank Harald Baumgartner and Jonas Haase for significantly speeding up the project by patiently helping me numerous times with technical questions related to hardware, servers and SciServer.

I gratefully acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC-2094 - 390783311.

Submitting the present thesis punctuates many years of studies in Paris and Munich. I cannot express how grateful I am for the support of the people who have accompanied me beyond the academic framework. Each of them deserves a paragraph on its own here.

I am grateful to Quirin for being a great friend, an authentic *Grantler* (or *Münchner Original* as they call them) and the best Wirt possible. Thanks also to my colleagues and friends at the Old Irish: Markus, Oscar, Christoph and Kili for many, many shifts in our good old pub.

To my friends: Marc, Till, Lea, Pint, Chris, Raph, Michi, Iris, Manu, Francois, Lorenzo, Anna, Max, Karim, Hedi, Basile, Matthias, Julie, Christian, Pascal, Tibo, Pauline, Roland, Max, Klara, Fiona; I cannot stress enough how lucky I was to share good times with you. Your support and generosity have been my fuel over the last few years. Thanks for barbecues *im Garten*, cooking extravaganzas, the finest cultural moments at black metal festivals and Heilung rituals, Bastelstunden, 8-hour crusades on Mecatol Rex, non-sense Cyberpunk campaigns, taking *raccourcis*, pushing me down far too steep downhill slopes (*"Le Trail!!"*) and forming the least productive metal band in Munich.

Thanks to Sissi and Martin for being so welcoming and making me feel part of the family.

To my partner Nora, thank you. Thank you for your grounding and calming presence. Even when I spent days knee-deep in papers, eyes locked on some hieroglyphic code on the screen and mumbling unintelligible things to myself, you were always patient and supportive. Thank you for being there these last seven years. I could not have undertaken this journey without you and little Beppo.

Finally, I owe everything to my family.

Vielen Dank an Anne, Roland, Oliver, Toni und Marco für die Unterstützung, die ihr mir jederzeit zuteil werden lasst.

Merci à ma tante, Marie, pour l'aide qu'elle m'a apporté durant mes études.

Merci à ma soeur, Johanna, et mon frère, Nicolas. Merci pour votre soutien infaillible et vos mots réconfortants toutes ces années. Les schwawarma à Neukölln, les soirées confort canap' avec le chien grincheux et une myriade de petits instants de bonheur à vos côtés m'ont donné la force d'accomplir cette tâche.

À mon père, qui a rejoint les étoiles, et à ma mère, merci d'avoir tant sacrifié pour me permettre d'étudier et d'arriver ici aujourd'hui. Je dois tout à votre générosité. Merci pour

votre force face aux épreuves de ces dernières années et d'avoir toujours trouvé la force de nous encourager.