Population studies of accreting supermasive black holes in the eROSITA all-sky survey

Sophia G. H. Waddell

München 2024

Population studies of accreting supermasive black holes in the eROSITA all-sky survey

Sophia G. H. Waddell

Dissertation
der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Sophia G. H. Waddell
aus Halifax, NS, Canada

München, den 02/10/2024

Erstgutachter: Prof. Dr. Kirpal Nandra
Zweitgutachter: Prof. Dr. Andreas Burkert
Tag der mündlichen Prüfung: 13/11/2024

To my fiancé, James Lloyd, for the endless love, support, happy memories in our little Garching studio, and adventures in Germany and beyond,
to my parents, Tanya and Greg (mom and dad), and my siblings, Lucas and Charlotte, for your love and belief in me, and for visiting me so far from home,
to all my friends and loved ones, especially the Not So Secret Santa Co., for making my four years here filled with fun adventures, good food, and lots of laughs.
I could not have done it without you!

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Zusammenfassung

In dieser Arbeit werden die Natur und die Eigenschaften supermassereicher Schwarzer Löcher untersucht, welche im Zentrum der meisten Galaxien vermutet werden. Ein gewisser Prozentsatz ($\sim$10%) dieser Schwarzen Löcher nimmt aktiv Material aus der Umgebung auf. Solche Objekte werden als aktive Galaxiekerne (AGN von englisch active galactic nucleus) bezeichnet. Diese Arbeit konzentriert sich auf die Röntgeneigenschaften von AGN und untersucht die Emission, die in den innersten Regionen dieser Systeme erzeugt wird. Die Röntgenstrahlung wird in der heißen Korona erzeugt, einer Population relativistischer Elektronen mit umstrittener Geometrie, die sich in der Nähe des zentralen Schwarzen Loches befindet. In dieser Arbeit werden die Röntgeneigenschaften anhand der Daten von eROSITA untersucht. eROSITA, an Bord von SRG, ist der Detektor für weiche Röntgenstrahlen, der eine Durchmusterung der Röntgenemission am gesamten Himmel durchgeführt hat. In dieser Arbeit werden insbesondere Quellen untersucht, die im hochenergetischen (harten) Röntgenbereich von $2,3-5,0\,\mathrm{keV}$ ausgewählt wurden, wo die geringere Empfindlichkeit von eROSITA zur Erzeugung einer Stichprobe von hellen Röntgenquellen führt, von denen die meisten AGN sind. Die Untersuchung dieser AGN ermöglicht die direkte Untersuchung der heißen Korona, die das Röntgenspektrum von AGN oberhalb von $2\,\mathrm{keV}$ fast durchgängig dominiert, und liefert eine Stichprobe, die relativ wenig durch Absorption beeinträchtigt wird, insbesondere im Vergleich zur Auswahl der weichen (Haupt-)Röntgenstrahlung von eROSITA. Entscheidend ist, dass die genaue Modellierung der Korona die weitere Untersuchung zusätzlicher Emissions- und Absorptionskomponenten jenseits der koronalen Emission ermöglicht, was eine detailliertere Untersuchung der inneren Regionen der AGN erlaubt. Ich werde diese Eigenschaft nutzen, um die Eigenschaften der Korona, der Emission und der Absorption von Proben ausgewählter AGN im harten Röntgenlicht zu untersuchen, um die individuellen und die Ensemble-Eigenschaften zu untersuchen.

Vor Beginn der geplanten Ganzhimmelsdurchmusterung führte eROSITA eine Analyse eines kleineren Feldes von 140 deg^2 in der Äquatorebene durch, das als eROSITA Final Equatorial Depth Survey (eFEDS) bekannt ist. Dies geschah, um die Leistung von eROSITA in der geplanten Durchmusterung des gesamten Himmels in acht Durchgängen zu ermitteln. Aus dieser Durchmusterung wurde eine Stichprobe von 254 ausgewählten Quellen für harte Röntgenstrahlung gebildet. In Kapitel 2 beschreibe ich die Bemühungen, diese Quellen zu identifizieren, zuzuordnen und zu klassifizieren, indem ich ihre Eigenschaften bei verschiedenen Wellenlängen vergleiche und spektroskopische und photometrische Rotverschiebungen verwende. Als Ergebnis wurde eine Stichprobe von 200 AGN identifiziert.

Diese AGN-Stichprobe weist eine hohe spektroskopische Vollständigkeit auf und bildet eine Stichprobe vorwiegend heller und lokaler Quellen über einen großen Himmelsbereich, die sich ideal für den Vergleich mit tieferen Durchmusterungen mit höherer Rotverschiebung und vergleichbarem kosmologischen Volumen eignet. Dabei zeigt sich, dass die harten AGN der eFEDS-Stichprobe vergleichbare Massen und Helligkeiten von Schwarzen Löchern aufweisen, aber geringere Akkretionsraten als eine komplementäre, höher rotverschobene Stichprobe von COSMOS, was im Gegensatz zu anderen Arbeiten auf einen fehlenden Downsizing-Effekt hinweist. Während die meisten AGN auch große Flüsse weicher Röntgenstrahlung und weiche Spektren aufwiesen, enthüllte die Analyse auch eine Gruppe von AGN, die nur Anzeichen für Röntgenemission oberhalb von 2.3 keV zeigen. Diese sind als hard-only AGN bekannt und sind Kandidaten für mäßig oder stark verdeckte AGN, die in den Hauptproben mit weicher Röntgenstrahlung nicht auftauchen würden. Die in diesem Kapitel entwickelte und vorgestellte Methodik zur Klassifizierung und Untersuchung ausgewählter Proben mit harter Röntgenstrahlung ist von entscheidender Bedeutung für die Anwendung auf die viel größere Ganzhimmelsdurchmusterung eRASS:1.

In Kapitel 3 wird dieselbe Stichprobe von 200 AGN, die in Kapitel 2 erstellt wurde, spektroskopisch modelliert, wobei eine Vielzahl von phänomenologischen und physikalischen Modellen verwendet wird, um auf thermische Emission, nicht-thermische Emission, ionisierte warme Absorption, neutrale fleckige Absorption, relativistische unscharfe Reflexion von der inneren Akkretionsscheibe und auf das Vorhandensein einer sekundären, kühleren (warmen) Korona zu testen. Für jede Quelle wird zunächst ein durch die galaktische Absorption und die Absorption der Wirtsgalaxie modifiziertes koronales Leistungsgesetz zugrunde gelegt und dann jede zusätzliche Emissions- oder Absorptionskomponente einzeln berechnet. Ein Bayesianischer Ansatz wird verwendet, um die besten Anpassungsparameter für jedes Modell zu bestimmen, und der Bayesianische Modellvergleich wird genutzt, um einen Modellvergleich durchzuführen. Durch eine umfangreiche Reihe von Simulationen wird ermittelt, wo eine Modellkomponente bei verschiedenen Signifikanzniveaus als statistisch signifikante Anpassungsverbesserung angesehen werden sollte. Das Ergebnis dieser Analyse ist, dass das allgemein berichtete Phänomen des weichen Röntgenüberschusses, also des Überschusses an Photonen unterhalb von $\sim 2\,\mathrm{keV}$, nicht durch einen einzigen Mechanismus erzeugt wird, sondern durch eine Vielzahl von Mechanismen, die von der Quelle abhängen. Wir stellen fest, dass weiche Überschüsse fast zu gleichen Teilen durch eine sekundäre Korona, einen warmen Absorber und einen neutralen, teilweise bedeckenden Absorber erzeugt werden. Der thermische Ursprung wird stark abgelehnt, und ein relativistisches Modell der unscharfen Reflexion wird nur für eine kleine Minderheit von Quellen bevorzugt. Die Analyse zeigt den Reichtum und die Komplexität des weichen Überschusses und verdeutlicht die Notwendigkeit einer sorgfältigen Spektralanalyse selbst bei Spektren mit geringem Signal, um die zugrunde liegende spektrale Form korrekt zu messen.

Die in den Kapiteln 2 und 3 für das eFEDS-Feld gewonnenen Erkenntnisse werden in Kapitel 4 zunächst auf die erste eROSITA Ganzhimmelsdurchmusterung (eRASS:1 von englisch eROSITA all-sky survey) angewendet. Es werden insgesamt 5466 ausgewählte Quellen im $2,3-5\,\mathrm{keV}$-Band dargelegt. Nach der Entfernung von falschen Entdeckungen und ausgedehnten Quellen, bei denen es sich wahrscheinlich nicht um AGN handelt, verblei-

ben 4985 Quellen. Je nachdem, wo am Himmel sie sich befinden, werden diese Quellen zur Identifizierung mit Mehrwellenlängenkatalogen abgeglichen und anschließend anhand ihrer Eigenschaften klassifiziert, wobei der Ansatz aus Kapitel 2 verfolgt wird. Nach dem Entfernen der Kandidaten für gebeamte AGN/Blazare identifiziere ich eine Stichprobe von 1328 AGN mit spektroskopischen Rotverschiebungen, die von sehr lokalen AGN von $z \sim 0.003$ bis $z > 3$ reichen. Ähnlich wie in eFEDS bilden die harten Röntgen-AGN eine helle und lokale Stichprobe, die sich jedoch auf höhere Rotverschiebungen und geringere Helligkeiten als andere All-Sky-Röntgendurchmusterungen in verschiedenen Energiebereichen (z.B. HEAO-1, Swift-BAT) erstreckt und somit eine äußerst komplementäre Stichprobe für zukünftige Untersuchungen darstellt. Im Vergleich zu Swift-BAT erweist sich eROSITA als sehr empfindlich für nicht verdeckte AGN vom Typ 1. Wie bei eFEDS leuchten die meisten Quellen hell im weichen Röntgenlicht. Eine Minderheit (29) der AGN mit bekannten spektroskopischen Rotverschiebungen (spec-z) strahlen nur oberhalb von $2,3\,\mathrm{keV}$ und bilden die „hard-only"-Stichprobe. Diese Quellen sind eher in AGN vom Typ 2 zu finden und haben Säulendichten in der Größenordnung von $10^{23-23.5}$ cm^{-2} , was zeigt, dass eROSITA in der Lage ist, verdeckte, Compton-dünne AGN zu entdecken. Diese Quellen befinden sich nachweislich in von der Seite betrachteten Galaxien, und zeigen eine geringere Streuung weicher Röntgenstrahlung als erwartet, was auf einzigartige Eigenschaften im Vergleich zu anderen ausgewählten Quellen für harte Röntgenstrahlung hindeutet. Die eRASS1-Stichprobe ist nachweislich ein reichhaltiger Ausgangspunkt für die spektrale Modellierung oder für die Auswahl einzigartiger Unterstichproben für eine eingehende Analyse.

Kapitel 5 baut auf der in Kapitel 4 identifizierten Stichprobe von 29 ausschließlich harten AGN mit spec-z auf, mit dem Ziel, die einzigartigen Eigenschaften von Typ-2-AGN mit geringer Streuung im weichen Röntgenlicht besser zu verstehen. Dazu wird zunächst eine weitere Quelle mit spec-z identifiziert und eine ergänzende Stichprobe sehr harter AGN hinzugefügt. Diese werden dann mit bestehenden NuSTAR-Datensätzen abgeglichen, um Röntgenspektralinformationen im Bereich von $3 - 79\,\mathrm{keV}$ zu erhalten. Für die 16 Quellen, für die NuSTAR-Daten verfügbar sind, wird eine detaillierte Spektralmodellierung durchgeführt, bei der mehrere Absorberstrukturen und -geometrien getestet werden. Ich zeige, dass die Mehrheit dieser AGN mit geringer Streuung ein klumpiges Absorbermodell stark bevorzugt, was auf eine komplexe Struktur hindeutet, die wahrscheinlich mit der Absorption auf pc-Ebene im staubigen Torus zusammenhängt. Vorläufige Unterschiede werden zwischen Quellen unterschiedlicher Wirtsgalaxiemorphologien gefunden, wobei frontale Spiralgalaxien und elliptische Galaxien einen sphärischen klumpigen Torus benötigen, um die beobachtete Spektralform zu erzeugen, während von der Seite betrachtete Spiralgalaxien Hinweise auf großräumige neutrale Absorption zeigen, wahrscheinlich aufgrund der galaktischen Scheibe. Die Modellierung zeigt auch eine Vorliebe für steilere Photonenindizes als erwartet, die denen von AGN des Typs 1 ähnlicher sind. Dies ist wahrscheinlich ein Zeichen dafür, dass die Verwendung des richtigen Spektralmodells für jede Quelle für eine korrekte Charakterisierung des koronalen Photonenindexes unerlässlich ist. Dies unterstreicht die Bedeutung einer sorgfältigen Modellierung großer Stichproben, bei denen die Eigenschaften der Korona eingeschränkt werden können.

Abstract

This thesis examines the nature and properties of supermassive black holes, which are believed to lie at the centres of most galaxies. Of these, some percentage ($\sim 10\%$) are actively accreting material from their surroundings, and are known as active galactic nuclei (AGN). This thesis concentrates on the X-ray properties of AGN, studying the emission produced in the innermost regions of these systems. X-rays are produced in the hot corona, a population of relativistic electrons of debated geometry located close to the central black hole. This thesis studies the X-ray properties using data from eROSITA, the soft X-ray detector aboard SRG, which has performed a survey of X-ray emission across the entire sky. This thesis specifically examines sources selected in the higher energy (hard) X-ray regime of $2.3 - 5\,\mathrm{keV}$, where the reduced sensitivity of eROSITA results in the production of a bright sample of X-ray sources, the majority of which are AGN. The study of these AGN permits the direct probing of the hot corona, which almost universally dominates the X-ray spectrum of AGN above $2\,\mathrm{keV}$, and provides a sample which is relatively less biased to the effects of absorption, especially relative to the eROSITA soft (main) X-ray selection. Crucially, the accurate modelling of the corona enables further probing of additional emission and absorption components beyond the coronal emission, allowing for a more detailed probing of the inner regions of the AGN. I will exploit this property in order to study properties of the corona, emission, and absorption from samples of hard X-ray selected AGN to study the individual and ensemble properties.

Prior to beginning the planned all-sky survey, eROSITA performed an analysis of a smaller, 140 deg^2 field in the equatorial plane, known as the eROSITA Final Equatorial Depth Survey (eFEDS). This was done to ascertain the performance of eROSITA in the planned, eight-pass, all sky survey. From this survey, a sample of 254 hard X-ray selected sources was formed. In Chapter 2, I present the efforts to identify, associate and classify these sources by comparing multiwavelength properties and using spectroscopic and photometric redshifts. As a result, a sample of 200 AGN is identified. This sample of AGN has high spectroscopic completeness, and forms a sample of primarily bright and local sources over a large sky area, ideal for comparison with deeper, higher redshift surveys of comparable cosmological volume. Doing so reveals that the eFEDS hard sample AGN show comparable black hole masses and luminosities but lower accretion rates than a complementary, higher redshift sample from COSMOS, showing a lack of downsizing effect, contrary to other works. While most AGN also showed large soft X-ray fluxes and soft spectra, the analysis also revealed a sample of AGN which only show evidence for X-ray

emission above 2.3 keV known as hard-only AGN, which are candidates for moderately or heavily obscured AGN which would not appear in the main, soft X-ray selected eROSITA samples. The methodology developed and presented in this chapter to classify and study the hard X-ray selected samples is crucial for application to the much larger all-sky survey, eRASS:1.

In Chapter 3, the same sample of 200 AGN created in Chapter 2 is modelled spectroscopically using a variety of phenomenological and physical models to test for thermal emission, non-thermal emission, ionised warm absorption, neutral patchy absorption, relativistic blurred reflection from the inner accretion disc, and for the presence of a secondary, cooler (warm) corona. A baseline coronal power law modified by Galactic and host-galaxy absorption is first applied to each source, then each additional emission or absorption component is applied individually. A Bayesian approach is employed to determine the best fit parameters for each model and Bayesian model comparison is exploited to perform model comparison. Through a extensive series of simulations, I determine where a model component should be considered a statistically significant fit improvement at various levels of significance. The result from this analysis is that the commonly reported phenomenon of the X-ray soft excess, or surplus of photons below ~ 2 keV, is not produced by a single mechanism; but rather through a variety of mechanisms depending on the source. We find nearly equal prevalence of soft excesses produced by a secondary corona, a warm absorber, and a neutral partial covering absorber. The thermal origin is found to be strongly disfavoured, and a relativistic blurred reflection model is preferred for a small minority of sources. The analysis demonstrates the richness and complexity of the soft excess, and demonstrates the necessity of careful spectral analysis even for low signal spectra in order to properly measure the underlying spectral shape.

The lessons learned from Chapters 2 and 3 for the eFEDS field are first applied to the first eROSITA All-sky survey (eRASS:1) in Chapter 4. A total of 5466 sources selected in the $2.3-5$ keV band are presented. After removing spurious detections and extended sources which are unlikely to be AGN, 4985 sources remain. Depending on where in the sky they are located, these sources are then matched with multiwavelength catalogues for identification purposes and are subsequently classified based on their properties, following the approach in Chapter 2. I identify a sample of 1328 AGN with spectroscopic redshifts, ranging from very local AGN of $z \sim 0.003$ to $z > 3$, after removing candidate beamed AGN/blazars. Similar to what was found in eFEDS, the hard X-ray AGN form a bright and local sample, but extending to higher redshifts and lower luminosities than other all-sky X-ray surveys in different energy regimes (e.g. HEAO-1, Swift-BAT), thus forming a highly complementary sample for future study. Comparing to Swift-BAT, eROSITA is found to be highly sensitive to unobscured, type-1 AGN. Also as in eFEDS, most sources shine bright in soft X-rays, but a minority (29) AGN with spectroscopic redshifts were found to only radiate above 2.3 keV, forming the hard-only sample. These sources are more likely to be hosted in type-2 AGN, and have column densities on the order of $10^{23-23.5}\,\mathrm{cm}^{-2}$, showing that eROSITA is capable of detecting obscured, Compton thin AGN. These sources are also shown to be hosted in more edge-on galaxies, and are found to show lower than expected soft X-ray scattering, suggesting unique properties as compared to other hard X-ray selected sources.

The eRASS1 hard sample is demonstrated to be a rich starting point for spectral modelling, or for selecting unique subsamples for in-depth analysis.

Chapter 5 builds upon the sample of 29 hard-only AGN with spec-z identified in Chapter 4, with the goal of better understanding the unique properties of low soft X-ray scattering, type-2 AGN. This is begun by identifying one additional source with a spec-z, and adding a complementary sample of very hard AGN. These are then matched against existing *NuSTAR* data sets in order to obtain X-ray spectral information in the $3 - 79\,\mathrm{keV}$ range. For the 16 sources which have *NuSTAR* data available, detailed spectral modelling is performed, testing multiple absorber structures and geometries. I demonstrate that a majority of these low scattering AGN strongly prefer a clumpy absorber model, suggesting a complex structure, likely associated with pc scale absorption in the dusty torus. Tentative differences are found between sources of different host galaxy morphologies, where face-on spiral and elliptical galaxies require a spherical clumpy torus to produce the observed spectral shape, whereas edge-on spirals show evidence for large-scale neutral absorption, likely in the galactic disc. The modelling also shows preference for steeper than expected photon indices more similar to those found in type-1 AGN, which is likely a sign that using the correct spectral model for each source is essential for proper characterisation of the coronal photon index, further highlighting the importance of careful modelling of large samples where properties of the corona can be constrained.

Kapitel 1

Active Galactic Nuclei and X-ray Astronomy

In this chapter, I will introduce the main science driver of this thesis, active galactic nuclei, powered by supermassive black holes which are actively accreting material. This thesis studies AGN using X-ray observations. I will present an overview of their key emission properties and classification methods, and then provide a more detailed overview of the X-ray emission mechanisms and spectral components. I will then focus on the origins and development of X-ray astronomy and X-ray studies of AGN, and present the eROSITA instrument onboard the Spektrum-RG mission, a novel X-ray telescope which has observed and found over a million accreting supermassive black holes. I will describe eROSITA hard X-ray selected samples of AGN and their advantage. Finally, I will provide an outline of this thesis.

1.1 Black holes

1.1.1 Astrophysical black holes

In 2019, the first direct image of a black hole was produced by the Event Horizon Telescope (EHT), to the great excitement of scientists and the public alike (Event Horizon Telescope Collaboration et al. 2019). The target was Messier 87, a nearby massive galaxy with a supermassive black hole with a mass of several billion suns at its centre (Event Horizon Telescope Collaboration et al. 2019). In the dark central region, the image encompasses the event horizon, the region in which the gravitational pull of the black hole is so strong that even light cannot escape.

Black holes were first theorized in the 18th century, born from simple Newtonian mechanics arguments in a letter by John Michell, who theorized that there may be stars with sufficient gravity such that the escape velocity;

$$v_{\mathrm{esc}} = \sqrt{\frac{2GM}{r}} \tag{1.1}$$

Abbildung 1.1: First image of a black hole (M87*) released by the Event Horizon Telescope on April 10, 2019. The rings around the image show light which is bent around the black hole by extreme general relativistic effects. Credit: Event Horizon Telescope Collaboration.

where G is the gravitational constant, M is the mass of the star, and r is the radius, would exceed the speed of light $c = 2.98 \times 10^8$ cm s^{-1}. At this point, light would not be emitted, resulting a in object dubbed a dark star. The idea was abandoned once the wave nature of light was discovered, as it was not clear that gravity would affect light in this way. The idea was reprised by (Oppenheimer & Snyder 1939), who again posited that a stellar core would continue to collapse if another stable state was not created, where the gravitational radius and mass of the star would be related by;

$$r_{\mathrm{g}} = \frac{2GM}{c^2} \tag{1.2}$$

which a simple rearrangement of equation one in the boundary where the escape velocity approaches the speed of light. This is also defined as the Schwarzschild radius, $\mathrm{R_s}$. The concrete discovery of a stellar mass astrophysical black hole was later confirmed in 1971, a black hole around which a companion star is orbiting, named Cygnus X-1 (Bowyer et al. 1965; Bolton 1972; Webster & Murdin 1972).

Perhaps the most direct proof of a supermassive black hole prior to the EHT observations came from studying the centre of the Milky Way. In 2020, Reinhard Genzel and Andrea Ghez were awarded the 2020 prize in Physics for their respective studies of the black hole at the centre of out Galaxy, Sgr. A*. Both respective collaborations traced infrared emission from stars closest to the galactic centre, and by tracing their movements over long timescales, inferred the presence of an object with a mass of $\sim 4 \times 10^6$ solar masses (Genzel

et al. 1997; Ghez et al. 1998, 2000) and localized to an of 0.05±0.04 arcseconds (Ghez et al. 2000), comparable to the size of our solar system. This object can only be a supermassive black hole. Sgr A* was the observed by EHT (Event Horizon Telescope Collaboration et al. 2022), directly imaging the supermassive black hole.

Black holes span an enormous mass range from a few solar masses up to several billions, and can thus be divided into different categories based on their mass. Stellar mass black holes have masses of the order of ∼few M⊙ (Özel et al. 2010), and are the end points of the evolution of massive stars (Woosley & Weaver 1995; Heger et al. 2003). These are typically found in binary systems where they accrete mass from a donor star (e.g. see Bolton 1972; Webster & Murdin 1972, for Cyg X-1). On the other end of the mass range, supermassive black holes have masses of $\sim 10^6 - 10^{10}$ M⊙ and are found at the centres of virtually all galaxies (e.g. Schmidt 1978; Soltan 1982). These sources typically grew to such large masses via a combination of accretion from their surroundings, and via mergers with other SMBH; however, the precise growth channels of SMBH are not yet fully understood, especially in the early Universe (e.g. see discussion in Goulding et al. 2023).

1.1.2 Characterisation of black holes

Having established the existence of black holes, it becomes of great existence to learn how to describe and characterise them, and to better understand their fundamental properties. According to the no-hair theorem (Israel 1967; Robinson 1975), all black holes can be uniquely described based on three properties;

- Their mass, M
- The electric charge, Q
- The angular momentum, J

The charge Q is likely to be ~ 0 in an astrophysical context, so astrophysical studies of black holes focus on determining black hole mass and angular momentum.

The black hole mass can be measured using a variety of techniques. For stellar mass black holes, this is done by measuring the orbital period and velocity of the companion star in order to infer the mass function, the lower limit on the mass of the black hole (Remillard & McClintock 2006; Özel et al. 2010). For supermassive black holes, such a direct measurement is not typically possible, as resolving the motion of individual stars or gas orbiting the black hole is impossible in most cases; notable exceptions include the use of the EHT (e.g. Event Horizon Telescope Collaboration et al. 2019, 2022) or GRAVITY (e.g. Gravity Collaboration et al. 2018). A thorough review of the variety of direct and indirect mass measurement techniques is given in Peterson (2014). For sources where the mass cannot be directly measured, reverberation mapping (Peterson 1993), where time delays due to light travel time between different physical emission regions, can be used to precisely infer black hole masses. However, this approach requires high cadence, long baseline data sets (Peterson 2014). Instead, black hole mass measurements from reverberation mapping

campaigns can be used to calibrate scaling relations to infer black hole mass from other properties; perhaps most notably for this thesis, through measuring the widths of broad permitted optical lines (e.g. Peterson & Wandel 1999, 2000; Bentz et al. 2010; Shen et al. 2011). More discussion is also given in Chapters 2 and 3.

The angular momentum is usually characterised using the dimensionless spin parameter (e.g. see Reynolds 2021, for a recent review);

$$|a| = \frac{cJ}{GM^2} \tag{1.3}$$

where M is the black hole mass, and J is the angular momentum. The black hole spin |a| can take on values of 0 (non spinning) to 0.998 (maximum spin). A negative, or retrograde spin, implies that the black hole spin is in the opposite direction to that of the accretion disc. The black hole spin has an impact on the innermost stable circular orbit, or ISCO; the smallest stable circular orbit that a test particle can remain in. For non spinning black holes, the ISCO is $r_{ISCO} \sim 3\ R_s \sim 6GM/c^2$, but approaches $\sim 0.5\ R_s \sim GM/c^2$ for a maximum spinning black hole and extends to $\sim 4.5\ R_s \sim 9GM/c^2$ for retrograde spin (e.g. see Figure 1 of Reynolds 2021).

There are two crucial tools for performing black hole spin measurements; through modelling of the X-ray reflection spectrum, and through thermal continuum fitting. The thermal continuum fitting method (Zhang et al. 1997; McClintock et al. 2014) relies on the fact that more rapidly spinning black holes will have hotter inner disc temperatures due to the closer proximity to the black hole. This method is impractical for most accreting supermassive black holes, since other emission mechanisms will dominate over the accretion disc, and the UV peak is difficult to observe due to absorption. Instead, X-ray reflection modelling (see Fabian et al. 1989; George et al. 1998; Ross et al. 1999, Chapter 3) is preferred. The X-ray reflection spectrum produces a multitude of emission line features, including a prominent fluorescent emission line at 6.4 keV from neutral iron, Fe Kα(e.g. Nandra et al. 1997b). As the inner disc (ISCO) moves closer to the black hole, the fluorescent emission line photons are subject to stronger general relativistic effects and are gravitationally redshifted (e.g. Tanaka et al. 1995; Fabian et al. 2009). Through measuring the emission line profile, the gravitational redshift can be inferred, and hence the ISCO and therefore black hole spin (see Reynolds 2021, and references therein). A schematic showing the effects of spin on the reflection profile is shown in Fig. 1.2.

1.1.3 Active Galactic Nuclei

The study of black holes continues to be of key importance for astronomers across multiple sub-fields. Supermassive black holes are understood to be found at the centre of virtually every massive galaxy. However, only around 10% are believed to be active (Schmidt 1978; Bongiorno et al. 2012; Georgakakis et al. 2017); this means that they are actively accreting material from their surroundings (Salpeter 1964). Black holes alone in isolation do not emit electromagnetic radiation, so must be found via observing gravitational effects on

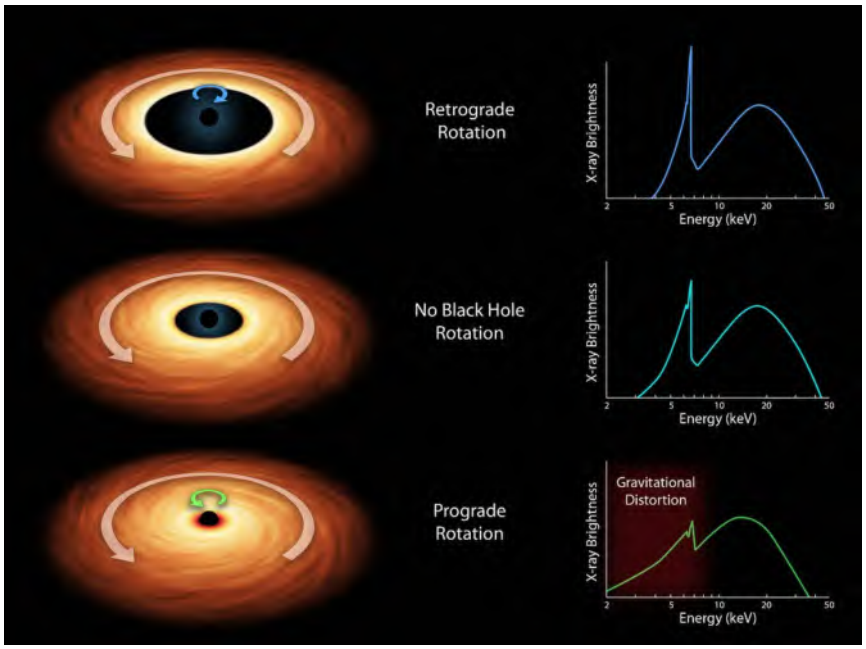

Abbildung 1.2: Schematic showing the effects of black hole spin on the reflection spectrum, in particular the Fe Kα emission line at 6.4 keV. Top: retrograde spin results in a larger ISCO, and thus a narrower Fe Kα line. Middle: a non-spinning (Schwarzchild) black hole, showing moderate gravitational redshifting of the red wing of the Fe Kα line. Bottom: maximum spin, showing the most extreme relativistic effects and an extremely broad red wing on the Fe Kα line. Image credit: NASA/JPL-Caltech

surrounding material, or through the study of accreting systems (Zel'dovich & Novikov 1971).

The first AGN to be recognized as a distant, extragalactic object emitting from an extremely compact region was 3C 48, followed closely by 3C 273 (see e.g. Schmidt 1963; Hazard et al. 1963). Both sources were first detected in radio surveys, and a search was performed to find a counterpart in optical emission. The optical spectra for both sources were unlike anything expected at the time, both objects were exceedingly bright, but showed broad optical lines that could not be associated with any known elemental transition at no cosmological redshift ($z = 0$). These sources were classified as quasi-stellar objects, or QSOs. Schmidt (1963) correctly identified that the lines in the 3C 273 spectrum exactly matched the expected spectrum from hydrogen, but at a higher redshift of $z = 0.158$; 3C 48 was found to be at an even higher redshift of $z = 0.367$. Given the large cosmological distances to the galaxies but extremely bright magnitudes, it was concluded that both galaxies hosted accreting supermassive black holes, required to produce such extreme luminosities in compact areas (e.g. see Sandage 1965). AGN radiate with luminosities:

$$L = \epsilon \dot{M} c^2 \tag{1.4}$$

where ϵ is the accretion efficiency (Thorne 1974) of $\sim 0.1 - 0.4$ depending on the spin, and $\dot{M}$ is the mass accretion rate.

The point at which the minimum outwards radiative force of these accreting sources is balanced by the inwards gravitational force, assuming only electron scattering from fully ionised Hydrogen, is known as the Eddington luminosity;

$$L_{\rm Edd} = \frac{4\pi G M m_p c}{\sigma_T} \tag{1.5}$$

where M is the mass of the black hole, m_p is the mass of a proton, and σ_T is the Thompson cross-section for electron scattering. AGN show typical luminosities of $0.01 - 1 L_{\rm Edd}$, suggesting they are accreting at a large fraction of the Eddington limit. Luminosities of 10^{46-47} ergs s^{-1} cm^{-2} can be reached for moderate SMBH with masses $M \sim 10^8 M\odot$ (e.g. see Woo & Urry 2002).

Continued study of AGN has revealed much about their nature. AGN are responsible for some of the most extreme, energetic, and luminous phenomena in the Universe. The growth of black holes by accretion is one of the most efficient mass-to-energy conversion processes, with efficiencies of 10-40%, a remarkable 10-100× more efficient than nuclear fission and fusion, respectively. AGN are found to radiate across all wavelengths of light, from radio emission (e.g. from relativistic jets; Schmidt 1963) to γ-rays (see e.g. Madejski & Sikora 2016, for a review). The inner structure of AGN is violent and complex, and numerous components and physical processes are responsible for the observed spectral energy distributions (SED). These features will be described in Section 1.2.

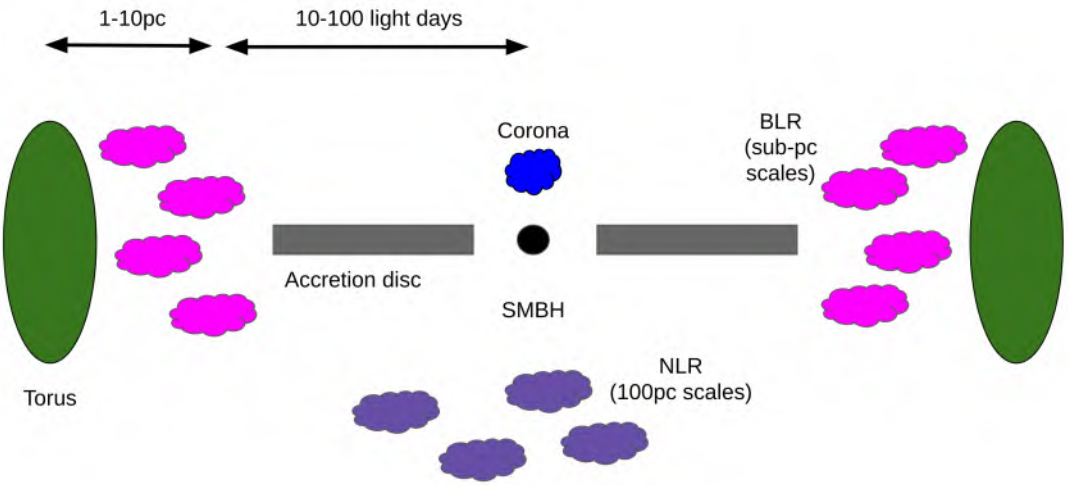

Abbildung 1.3: Cartoon image (not drawn to scale) showing a side view of the structure of emission regions of an AGN. At the centre lies the SMBH (black). The accretion disc is shown in grey, the X-ray corona is shown in dark blue, the broad line region (BLR) is shown in pink, the narrow line region (NLR) is show in purple, and the torus is shown in green. An idea of the scale of the emitting regions is shown in the top left corner.

1.2 Emission from AGN

1.2.1 The accretion disc

A key characteristic of active galactic nuclei is the presence of an accretion flow onto the black hole which radiates light isotropically. When matter comes sufficiently close, the extreme gravitational field surrounding the black hole pulls the material inwards. The accretion flow is then most likely believed to form a disc (e.g. Pringle & Rees 1972). A side-view cartoon of an accretion disc and other emitting components surrounding an SMBH is shown Fig. 1.3. In stellar mass black holes, where the accretion disc is produced from infalling material from a companion star. In the accretion disc, magnetic fields and other effects produce turbulent motions (e.g. see Papaloizou & Lin 1995; Blaes 2014, and references therein), which enables the transfer of angular momentum outwards; this loss of angular momentum then causes material to fall inwards. Gravitational potential energy from infalling material is then radiated as heat, producing the observed emission in optical and UV in AGN, and UV/X-rays in stellar mass black holes.

The accretion flow is commonly described by the thin disc model, described in detail in Shakura & Sunyaev (1973), where emission from matter in an optically thick, geometrically thin accretion disc is described, where exact properties and spectrum of the disc depends on the matter inflow rate, $\dot{M}$ (Shakura & Sunyaev 1973). From there, the disc temperature can be estimated as:

$$T \approx 50\,\mathrm{eV}\left(\frac{M}{10^8 M_\odot}\right)^{-1/4}\left(\frac{\dot{M}}{\dot{M_{\mathrm{Edd}}}}\right)^{1/4}\left(\frac{r}{r_\mathrm{g}}\right)^{-3/4} \tag{1.6}$$

where T is the disc temperature, M is the mass of the supermassive black hole, $\dot{M_{\mathrm{Edd}}}$ is the mass accretion rate at which the black hole radiates at Eddington luminosity, $\mathrm{L_{Edd}/c^2}$, and r is the radius at the point where the temperature is being estimated. The accretion disc radiates as a multi-temperature, multi-color blackbody; where each annulus of roughly constant radius and temperature radiates a blackbody spectrum:

$$B(\lambda, T) = \frac{2hc^2}{\lambda^5}\frac{1}{e^{\frac{hc}{\lambda kT}} - 1} \tag{1.7}$$

where h is Planck's constant, c is the speed of light, λ is the wavelength at which the peak of the blackbody is located, and T is the temperature. For AGN, the peak of the emission is in the ultraviolet, and larger radii produce optical continuum emission. This ultraviolet peak is sometimes referred to as the big blue bump (e.g. Malkan 1983; Czerny & Elvis 1987; Zhou et al. 1997).

1.2.2 The hot corona

The main topic of this thesis is the study of the X-ray properties of accreting supermassive black holes. The primary source of X-ray emission is the hot corona (Haardt & Maraschi 1991, 1993; Merloni & Fabian 2003), a population of relativistic electrons lying in the vicinity of the black hole. As discussed in Sect. 1.2.1, UV seed photons are emitted from the inner layers accretion disc. When these are incident upon the corona, they are Compton upscattered; the electrons and photons undergo inverse Compton collisions, and energy is transferred from electrons to the photons (Sunyaev & Truemper 1979). This processed is also referred to as Comptonisation. A cartoon of this process (not drawn to scale) is shown in Fig. 1.4. It can be seen that while some UV emission from the isotropic disc is radiated away from the corona, other UV photons feed the corona and are re-radiated as X-rays.

Since the disc can be approximated as a multi-temperature blackbody, the multiple blackbody emission components at different temperatures are up-scattered, and the resulting X-ray photons take approximately the form of a power law; $\mathrm{F = nE^{-\Gamma}}$, where Γ is the slope or index of the power law (e.g. Rybicki & Lightman 1979; Nandra et al. 1997a). This power law typically dominates the high energy or hard X-ray spectrum above $\sim 1-2\,\mathrm{keV}$, and will be discussed extensively in the following chapters.

The power law has a low-energy cut-off from the seed photon population, corresponding to the temperature of the accretion disc. This temperature depends on the radius and black hole mass, but will typically be on the order of a few eV for a typical AGN of mass $10^8 \mathrm{M}\odot$ and will not be readily measurable in the X-ray spectrum. There is also a high energy cut-off around $50-200\,\mathrm{keV}$ (see e.g. Fabian et al. 2015, and references therein), corresponding to the electron temperature in the corona. This parameter is of great interest to our understanding of the overall structure, opacity, and physical processes occurring

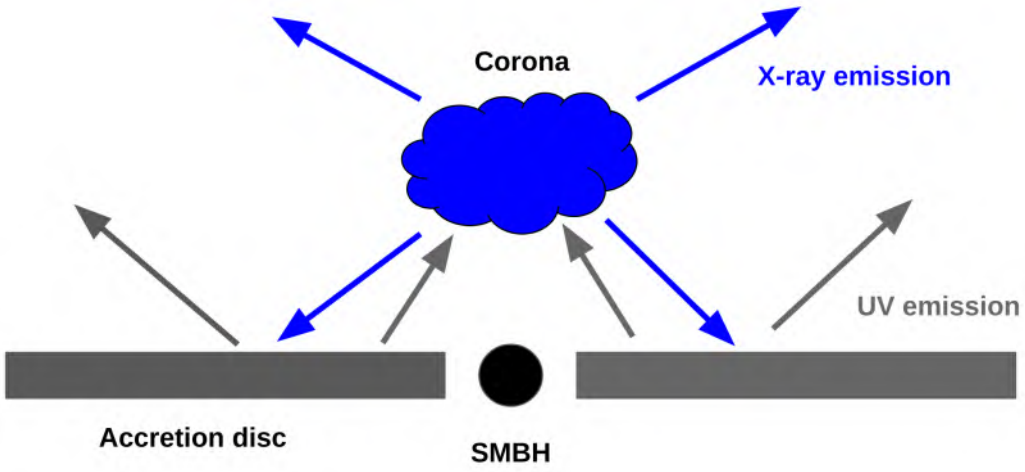

Abbildung 1.4: Cartoon image (not drawn to scale) showing the innermost few r_g of the AGN. The SMBH (black) at the centre is surrounded by the accretion disc, shown in grey. UV seed photons are emitted isotropically from the disc. Some are then emitted in the direction of the corona; these photons are then Compton up-scattered, producing the observed X-ray emission.

withing the corona, and can be measured with high energy X-ray instruments, including Integral, Swift-BAT or *NuSTAR* (see e.g. Petrucci et al. 2001; Fabian et al. 2015; Tortosa et al. 2018; Middei et al. 2019). Further investigation is also needed to study possible correlations of coronal temperature with other properties, including AGN class and black hole mass (Akylas & Georgantopoulos 2021).

Microlensing experiments have been used as a tool to probe the physical size of the corona (e.g. see Morgan et al. 2008; Mosquera et al. 2013; Reis & Miller 2013, and references therein). Microlensing is the process in which light from a background object (here a quasar) is gravitationally lensed due to the presence of a foreground object, causing the light to bend around it. The effects of microlensing increase as the scale size decrease, such that strong lensing effects would be produced in a small physical region (Reis & Miller 2013). Through the X-ray study of microlensing in lensed quasars, the corona is found to be compact, with sizes ranging from $\sim 1 - 40 \mathrm{r_g}$ (Mosquera et al. 2013; Reis & Miller 2013).

X-ray variability studies also find the corona to be compact, likely $< 10\, r_g$ (Risaliti et al. 2011; Kara et al. 2013; Reis & Miller 2013; Sanfrutos et al. 2013; Uttley et al. 2014). The X-ray varaibility studies are mostly performed using a technique known as X-ray reverberation mapping (Uttley et al. 2014), and a cartoon diagram is presented in Fig. 1.5. The physical explanation for this process is that as the corona emits, some photons are expected to be emitted in the direction of the inner accretion disc. These X-ray photons will reflect off of the inner accretion disc, producing a reflection spectrum consisting of a blend of emission features in the soft X-rays (Ross et al. 1999; Ross & Fabian 2005). As

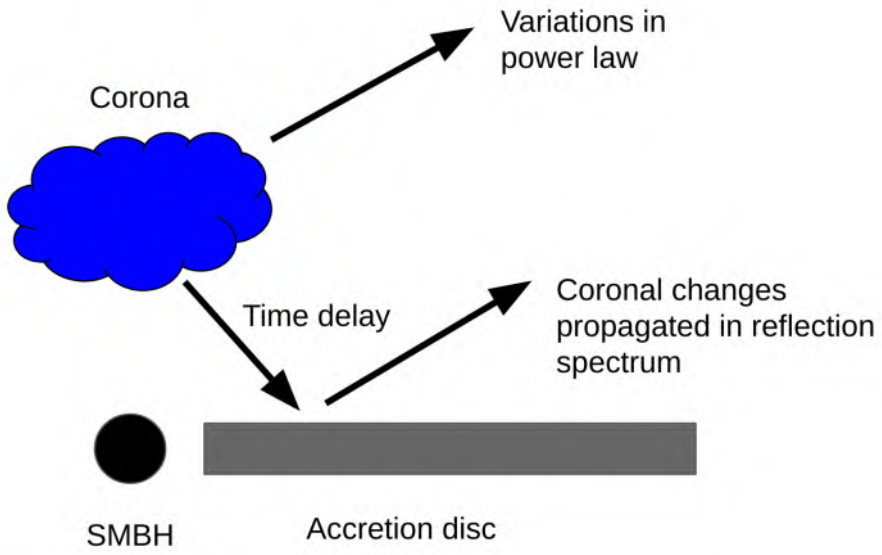

Abbildung 1.5: Cartoon image (not drawn to scale) demonstrating the physical motivation for reverberation mapping. The SMBH is shown in black, the accretion disc is shown in grey, and the corona is shown in blue. Variability in the corona is propagated to the reflected emission from the accretion disc, with a time delay corresponding to the light travel time including general relativistic effects.

fluctuations occur in the corona, this would produce variability which is observable in the power law component. These photons will then travel to the accretion disc, and the same change will be observed in the reflected component, with a delay corresponding to the light travel time including general relativistic effects. Kara et al. (2016), amongst others, have performed this analysis on numerous AGN, typically recovering a coronal height of $\sim 10\,r_{\mathrm{g}}$ above the central black hole.

This is also revealed using eclipse studies, whereby obscuring material passes between the observer and the corona (e.g. Risaliti et al. 2007, 2009; Gallo et al. 2021), which can be used to estimate both the size of the corona and the proximity of the surrounding material to the corona. It is likely that the corona is also extended - Recent results from polarization studies of AGN (e.g. Marinucci et al. 2022; Pal et al. 2023) seem to suggest a horizontally extended geometry, where the corona extends over some region of the inner disc. Such properties are hinted at in the cartoon showing emission from the corona (Fig. 1.4. Understanding the nature and structure of the corona as well as how it forms and varies over time, is a crucial element in understanding the nature of the inner structure and accretion flow of the AGN. Accurate modelling of the coronal X-ray emission is also an essential prerequisite for the study of additional features of the X-ray spectrum.

1.2.3 Narrow-line region (NLR) and broad-line region (BLR)

AGN typically show evidence for narrow emission lines produced in the so-called narrow-line region (NLR; Seyfert 1943; Antonucci & Miller 1985), gas clouds which can extend to large scales around $100 - 300$ pc (resolved for nearby sources García-Burillo et al. 2016; GRAVITY Collaboration et al. 2020a,b). The optical spectra of AGN show evidence for power law like continuum superimposed with a combination of emission lines (Khachikian & Weedman 1974). The NLR is shown in Fig. 1.3 in violet, at the bottom of the cartoon (not to scale). This gas has been heated by radiation from the AGN, and thus produces ionised fluorescent emission features. These lines typically have widths of $200 - 900$ km s^{-1}, suggesting a slow rotation speed around the central black hole (e.g. see Vaona et al. 2012).

In some sources, broad permitted optical lines (Khachikian & Weedman 1974) with full-width-half-maximum (FHWM) $1000 - 50,000$ km s^{-1} have been observed (Peterson 2006), including prominent broad Hα, Hβ, Mg II and CIV. These broad lines originate in the broad-line region (BLR), which is located closer to the black hole, likely inside the torus, on scales of $0.01 - 1$ pc (see Peterson 2006, for a review). The BLR is also shown in pink in the cartoon of Fig. 1.3, and is likely to be found at the outer regions of the disc, within the confines of the torus.

The presence or absence of broad lines have been proposed to be explained by the AGN unification model (Antonucci 1993), where sources with broad optical lines are viewed face-on Seyfert 1 or type-1 AGN and provide a direct view of the central emitting region, while sources without broad lines are viewed edge-on, known as Seyfert 2 or type-2 AGN where the torus obscures the BLR typical spectra of type-1 and type-2 AGN. Fig. 1.6 shows the striking differences between the typical optical spectra of type-1 and type-2 AGN. The Seyfert 1 AGN (top panel) shows a strong blue continuum at low wavelengths, very prominent broad lines associated with broad Hα, Hβ, Hγ MgII and CIV, as well as narrow lines associated with forbidden O[III] transitions around 5007 Angstrom. The type 2 AGN (bottom panel) shows more reddening, and shows many narrow permitted and forbidden emission lines, but no broad features.

The BLR shows evidence for complex structure, including inflowing or outflowing material (e.g. see Czerny et al. 2019). By assuming that the BLR is virialized (e.g. see Peterson & Wandel 1999) and moving in a Keplerian orbit (Baldwin 1997; Baldwin et al. 2003; Padovani et al. 2017), the FWHM of the BLR lines can be used to estimate the black hole mass; this process is described in more detail in Sect. 1.1 and 2.2.3. This mass determination makes the BLR a key area of study for better understanding AGN.

1.2.4 The dusty torus

Moving further away from the SMBH, studies of infrared emission from AGN suggests large dusty region at parsec scales. This dusty region is known as the torus Antonucci & Miller (1985), a possibly toroidal or spherical structure with varying covering factor. The torus was introduced as the source of the obscuration relevant to the AGN unification model, where the lack of observable broad lines and absorption in type-2 AGN is explained by a

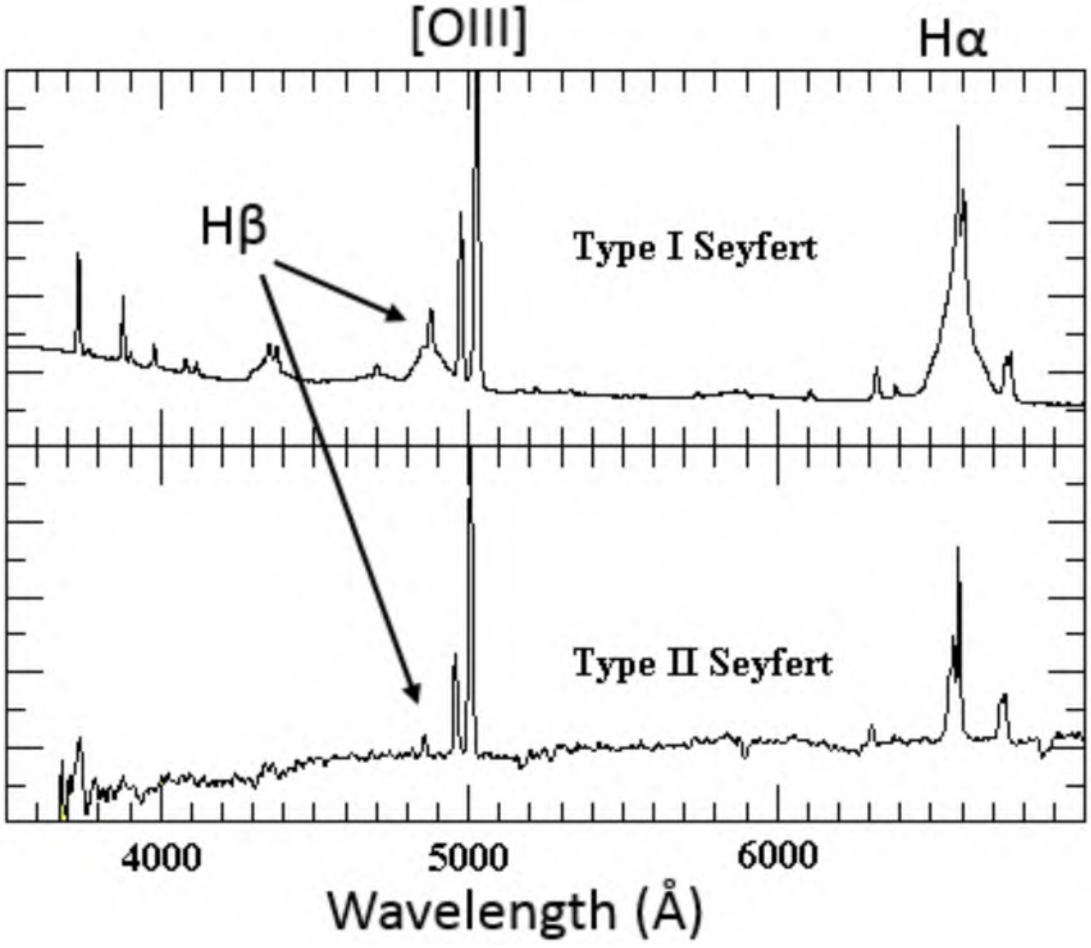

Abbildung 1.6: Top: typical spectrum of a type 1 AGN showing strong blue emission and a combination of broad and narrow emission lines. Bottom: typical type 2 AGN showing more reddening and only narrow optical lines. Prominent features including Hβ, Hα and O[III] are indicated. Image credit: Morgan, Siobahn. Distant and Weird Galaxies, course notes.

viewing angle which passes through the torus (Antonucci & Miller 1985; Urry & Padovani 1995). The torus also explains the shape of the infrared emission observed from most AGN, as the torus will re-emit optical and UV emission from the accretion disc which has been absorbed and re-emitted by dust grains, producing an infrared continuum (e.g. see Netzer 2015, for a review).

The torus is shown in green in Fig. 1.3 (not to scale), and may have a considerably larger vertical extend and/or more complex structure than shown in the cartoon. The observed infrared radiation cannot be explained through absorption in a smooth, homogenous medium, suggesting that the torus may be clumpy (see e.g. Nenkova et al. 2002; Mason et al. 2009). Some X-ray studies of AGN recover this result (Buchner et al. 2019). A review by Netzer (2015) proposes that not only is the torus likely to be clumpy, but also has hot and cold dust components. The impact of the torus on the observed X-ray spectrum is discussed in Chapter 1.3.3, and in more detail in Chapter 5.

1.3 X-ray spectroscopy of AGN

1.3.1 Continuum X-ray emission

In this thesis, I will be preparing and analysing large samples of hard X-ray selected AGN to perform X-ray spectral analysis. An example X-ray spectral shape showing some of the most commonly observed features is shown in Fig. 1.7, where the total model is shown as a solid black line, and main contributing components are shown as different colours and linestyles. The primary emission component is the corona, described in detail in Chapter 1.2.2 (Haardt & Maraschi 1991, 1993; Merloni et al. 2000; Fabian et al. 2004). The hot corona typically dominates the shape of the X-ray spectrum above $1-2\,\mathrm{keV}$, and in some cases dominates the entire observable X-ray spectrum (see e.g. Waddell & Gallo 2020; Waddell et al. 2024). A key measurable parameter of the corona using eROSITA is the photon index, with typical values of around $\Gamma \sim 1.8-2.0$ (Nandra & Pounds 1994a), possibly depending on other parameters such as accretion rate and/or spectral type (Lanzuisi et al. 2013; Liu et al. 2022b; Waddell et al. 2019, 2024; Nandra et al. 2024). The X-ray corona is shown as a dashed blue line in Fig. 1.7, and will be discussed in detail in Chapters 2-6.

1.3.2 Milky Way and host galaxy absorption

Gas within the Milky Way has column densities of order $\sim$few $\times 10^{20}\,\mathrm{cm}^{-2}$, absorbing soft X-ray photons below $\sim 0.5\,\mathrm{keV}$. These values can extend $> 10^{22}\,\mathrm{cm}^{-2}$ nearer to the Galactic disc (Willingale et al. 2013), absorbing X-rays below 1 keV. Additional absorption can also be present in the host galaxy, e.g. in the gas and dust reservoirs and/or spiral arms. This has been studied in detail by Laha et al. (2020), who show that while most sources have column densities of $10^{20}-10^{21}\,\mathrm{cm}^{-2}$, other galaxies may extend to higher values of $10^{22}-10^{23}\,\mathrm{cm}^{-2}$, further absorbing the X-rays in the $0.2-5\,\mathrm{keV}$ range. The absorption from a column density of $\sim 5\times10^{20}\,\mathrm{cm}^{-2}$ can be seen in the soft energy X-rays

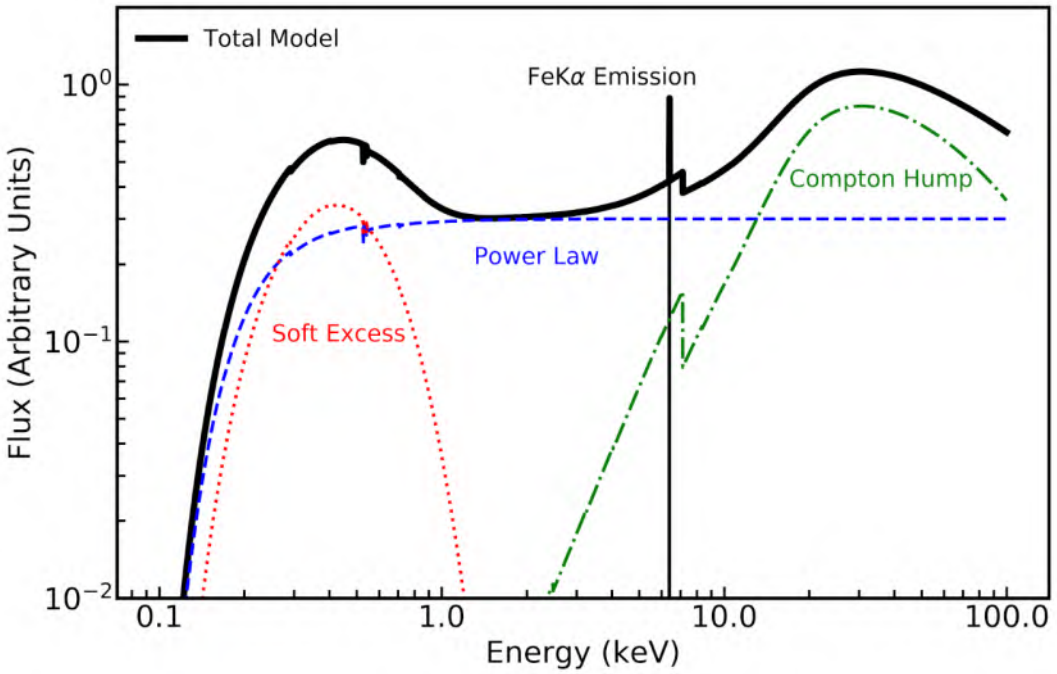

Abbildung 1.7: Typical X-ray spectral model for a mildly obscured AGN. The total model including all shown spectral components is indicated with a black solid line. The hot coronal power law is shown as a dashed blue line. The soft excess of unknown physical origin is shown as a red dotted line, contributing to the spectrum below $\sim 1\,\mathrm{keV}$. The Compton hump and iron absorption edge resulting from scattering of X-ray photons in a neutral medium (e.g. the torus) is shown as a green dash-dot line. The fluorescent Fe Kα emission line is shown as a grey solid line. Mild absorption in the Milky Way and host galaxy create a strong downturn in soft X-rays, below $\sim 0.4\,\mathrm{keV}$.

of Fig. 1.7, where a downturn in soft X-rays, below $\sim 0.4\,\mathrm{keV}$, can be observed. Absorption and measuring column densities is a central theme of this thesis and will be explored in Chapters 2-6.

1.3.3 Obscuration and reflection in the torus

Another primary source of absorption in AGN spectra is the torus (Awaki et al. 1991; Turner et al. 1997; Risaliti et al. 1999; Brightman & Nandra 2011; Baloković et al. 2019; Buchner et al. 2019), also discussed in Sect 1.2. The torus can exhibit a range of absorbing column densities; at lower column densities, we detect Compton-thin absorption on the order of $10^{22} - 10^{24}\,\mathrm{cm}^{-2}$, up to Compton thick absorption (Comastri 2004). At column densities of $\sim 1.5 \times 10^{24}\,\mathrm{cm}^{-2}$, the AGN is considered to be Compton thick, and all primary emission below 10 keV is absorbed, leaving only the reflected spectrum (Comastri 2004).

As X-ray photons from the corona are incident upon the torus, absorption and scattering interactions produce strong absorption features in the soft X-ray energies. Compton downscattering, reflection and absorption effecs are manifested in the Compton hump, peaking at $\sim 20 - 30\,\mathrm{keV}$, for neutral Fe K$\alpha$ emission line at 6.4 keV as well as an Fe Kα absorption edge at $\sim 7\,\mathrm{keV}$ (e.g. see Magdziarz & Zdziarski 1995; Murphy & Yaqoob 2009; Buchner et al. 2019, and references therein). These features can be seen in Fig. 1.7, with the Compton hump and absorption edge shown as a green dash-dot line, and the Fe Kα emission line shown as a grey solid line. A more detailed description of these features is given in Chapter 5, where torus models are applied to X-ray spectra of a sample of AGN.

1.3.4 Neutral, ionised and partial covering absorption

An often invoked physical mechanism to explain additional features in X-ray spectra is absorption, which can take on various flavours; warm absorbers (often interpreted outflowing disc winds George et al. 1998; Kaastra et al. 2000), neutral partial covering absorbers, or ionised partial covering absorbers (Tanaka et al. 2004). All suggest additional cloudy or collimated absorption zones likely to be found near to the black hole. Signatures of these components features would absorb photons around the 1 keV range, easily visible with soft X-ray detectors.

In a warm absorber or disc wind, X-ray emission passes through an ionised medium, producing absorption features concentrated in the soft X-ray spectrum (George et al. 1998; Kaastra et al. 2000; Kaspi et al. 2000a). Warm absorbers have been physically linked to outflowing, ionised disc winds (Kallman & Dorodnitsyn 2019). These warm absorbers are typically found to have moderate column densities on the order of $10^{20-23}\,\mathrm{cm}^{-2}$, and ionisations ranging from 10s to 1000s (Blustin et al. 2004, 2005; McKernan et al. 2007), but can extend to extremely hot, extremely fast, relativistic outflows known as Ultra Fast Outflows (UFO's Tombesi et al. 2010). The warm absorber model has been used to successfully model numerous X-ray spectra of AGN (e.g. Laha et al. 2014; Mizumoto et al. 2019; Waddell et al. 2024).

In a partial covering absorption scenario, X-ray emission from the corona is blocked by obscuring material along the line of sight (Tanaka et al. 2004). This material is patchy in nature, absorbing only some emission and allowing some to pass through. This produces deep absorption features in the spectrum. The partial covering absorber may be neutral or ionised, and the ionisation can be traced by modelling the observed absorption edges. The observed variability in X-ray spectra can be reproduced through a combination of varying column densities, ionisations and covering factors (Tanaka et al. 2004; Miyakawa et al. 2012; Gallo et al. 2015; Waddell et al. 2024), and can be used to explain eclipse events Gallo et al. (2021). The model is highly flexible and multiple absorbing zones with a variety of properties can reproduce observed X-ray spectra.

These absorption models are discussed in more detail in Chapter 3, and are tested on data obtained by eROSITA.

1.3.5 The soft excess

Finally, the soft excess is shown as a red dotted line in Fig. 1.7. The soft excess is an apparent surplus of X-ray photons above the coronal power law in the soft X-rays, below $\sim 1-2\,\mathrm{keV}$ (Pravdo et al. 1981; Singh et al. 1985; Arnaud et al. 1985), which is typically only visible in low obscuration sources. The physical mechanism responsible for the origin of the soft excess is a matter of dispute, and multiple models have been proposed to explain the appearance of this component. An early proposal was that the soft excess may be produced in the innermost accretion flow, where the highest temperatures are found; however, fitting a thermal blackbody to the observed spectral shape produces temperatures of $\sim 0.1\,\mathrm{keV}$ for all sources (e.g. Gierliński & Done 2004), which is too high for the disc temperatures observed from a typical Shakura-Sunyaev disc (Shakura & Sunyaev 1973).

Instead, multiple non-thermal models have been proposed to explain the origin of this component. First, the soft excess may originate including a secondary Comptonisation region known as the warm corona, which is more optically thick and cooler than the primary, hot corona (e.g. Done et al. 2012; Petrucci et al. 2018, 2020). This produces a smooth, featureless soft excess which begins to dominate over the primary power law around $0.5-2\,\mathrm{keV}$. Typical temperatures of $0.5-2\,\mathrm{keV}$ are recovered, as well as optical depths on the order of $10-20$ (Petrucci et al. 2018, 2020). This model has been successfully used in order to model the smooth soft excess observed in many type-1 AGN (e.g. Ehler et al. 2018; Tripathi et al. 2019; Waddell et al. 2024).

An alternative proposal is that the soft excess is produced by blurred relativistic reflection from photons reflecting off of the inner accretion disc (Ross et al. 1999; Ross & Fabian 2005). In this process, photons from the corona are incident upon the ionised inner layers of the accretion disc. The resulting reflection spectrum consists of a multitude of fluorescent emission lines and absorption edges concentrated in soft X-rays, but also produces a fluorescent Fe Kα emission line and Compton hump. These features are then relativistically blurred due to the proximity to the SMBH (Ross et al. 1999; Ross & Fabian 2005). Blurred reflection is an important physical mechanism in AGN, and this model also allows for study of the coronal height and geometry, reverberation mapping, and black hole

spin measurements (see Sect. 1.2). This model has also been used to model the soft excess shape in numerous type-1 AGN (Fabian et al. 2004; Gallo et al. 2019; Waddell et al. 2019).

The soft excess is the central theme of Chapter 3, and more detailed description and testing of different soft excess models is therefore given in Chapter 3.

1.4 X-ray observatories

1.4.1 Evolution of extrasolar X-ray astronomy

Extrasolar X-ray astronomy began with the launch of the first X-ray detector onboard a sounding rocket in 1962, detecting a bright X-ray source which was not coincident with the sun or moon. This was later identified to lie in the Scorpio constellation, and the source was named Scorpius X-1 (Gursky et al. 1966). The rocket was followed up further missions, discovering a plethora of additional X-ray sources and following up on the apparently diffuse X-ray emission across the entire sky discovered by Giacconi et al. (1962), which is now known as the X-ray background comprised mainly of distant and/or obscured AGN (e.g. see Comastri 2004, for a review).

A major advance in AGN science occurred with the launch of the HEAO-1 mission (Rothschild et al. 1979), which surveyed the entire X-ray sky over a broad energy range, with a limiting flux of $\sim 3 \times 10^{11}$ ergs s^{-1} cm^{-2}. From there, a sample of hard X-ray AGN was created, and presented in Piccinotti et al. (1982). Numerous other X-ray missions have launched throughout the last 50+ years, including EXOSAT (Taylor et al. 1981), GINGA (Tsunemi et al. 1989), *ASCA* (Tanaka et al. 1994), *BeppoSAX* (Boella et al. 1997), *Chandra* (Weisskopf et al. 2000), *XMM-Newton* (Jansen et al. 2001), Swift-BAT Barthelmy et al. (2005), and *NuSTAR* (Harrison et al. 2013) among others. Some missions have been used to perform specific follow-ups on the Piccinotti et al. (1982) reference sample (e.g. Turner & Pounds 1989; Nandra & Pounds 1994b; Schartel et al. 1997; Waddell et al. 2024, among others).

Prior to the launch of eROSITA the ROSAT (Truemper 1982) telescope surveyed the X-ray sky in soft X-rays, covering $\sim 0.1 - 2.4$ keV. This mission clearly demonstrated the power of surveys conducted by a sensitive X-ray telescope with good spatial resolution, enabling the study of clusters of galaxies, coronally active stars, X-ray binaries, and AGN (e.g. Ciliegi & Maccacaro 1996; Motch 2001; Böhringer et al. 2017; Xu et al. 2022, among others). The survey also shed more light on statistical properties of AGN, as well as improved our understanding of the soft excess, further revealing a population of sources with extreme ultra-soft X-ray emission (Boller et al. 1996). However, with only a limited time spent observing each source and limited coverage of higher energy X-rays, detailed spectral modelling was not possible for most AGN.

In 1999 came the launch of NASA's *Chandra* mission followed closely by ESA's *XMM-Newton* telescope. Both telescopes are active to this date, and continue to provide new data on AGN and other X-ray sources. Both missions are able to perform studies of individual objects and surveys (Fiore et al. 2003; Hasinger et al. 2007a; Nandra et al. 2015; Civano

et al. 2016; Pierre et al. 2016; Luo et al. 2017), but with excellent spatial resolution and energy resolution (in particular with *Chandra*) and very high effective are (especially *XMM-Newton*). Through these in depth studies AGN, the variability across both short and long timescales can more readily be probed (e.g. see Middei et al. 2017; Zheng et al. 2017). High quality spectra have also encouraged the development of more complex and physically motivated spectral models, as emission and absorption features beyond the simple coronal power law are often found (e.g. Waddell et al. 2019; Waddell & Gallo 2020, and references therein).

1.4.2 eROSITA on SRG

To fill the gap left by a lack of sensitive, all sky-surveys in soft to hard energy X-rays, the eROSITA (extended Roentgen Survey with an Imaging Telescope Array; Predehl et al. 2021b) instrument was designed. eROSITA is one of two instruments onboard the Spectrum Roentgen Gamma (SRG; Sunyaev et al. 2021) mission. ART-XC is the hard X-ray detector, sensitive in the $\sim 5-30\,\mathrm{keV}$ range (Pavlinsky et al. 2021). eROSITA is a complementary soft X-ray detector, sensitive in the $0.2-8\,\mathrm{keV}$ energy range. The instrument is comprised of 7 identical mirror assemblies, each with their own camera (Telescope Module; TM). These mirror assemblies can be seen in Fig. 1.8, showing the layout of each TM onboard eROSITA(see Predehl et al. 2021a). Wolter type mirrors using a parabolic/hyperbolic design are used to focus X-rays. This is because X-rays have extremely high energies, and require low incidence angles in order to be reflected and focussed onto the detector. Each TM contains 54 mirror shells with a gold coating, which produces a deep absorption edge at $\sim 2.3\,\mathrm{keV}$. This energy thus defines the separation between the soft (low energy X-ray) and hard (high energy X-ray) eROSITA bands.

A key feature of eROSITA is the high energy resolution provided by each TM. Each of the 7 telescope modules focus X-rays to a distinct pnCCD (Predehl et al. 2021a), based on technology developed for *XMM-Newton* but with updates to improve energy resolution and permit spectral coverage down to 0.1 keV. Thanks to the pnCCDs, eROSITA has excellent energy resolution[1], reaching $56-65\,\mathrm{eV}$ resolution at $0.525\,\mathrm{keV}$ and $75-82\,\mathrm{eV}$ resolution at $1.49\,\mathrm{keV}$ (Predehl et al. 2021a). Such low every resolutions are ideal for performing spectral analysis, because the energies and widths of spectral emission and absorption feature can more easily be determined (e.g. see Chapters 3, 4, 5).

The eROSITA instrument combines large field of view on the sky with very high soft X-ray sensitivity. An excellent measure of this is called the grasp, which multiplies the field of view (FOV) by the field of view-averaged effective area, and is a measure of survey capability. Comparing eROSITA to other past and current instruments operating in comparable energy ranges, eROSITA has a much larger grasp across almost the full energy range of the instrument, and particularly in the $\sim 0.3-5\,\mathrm{keV}$ band (see Fig. 1.9. This highlights eROSITA's ability to yield new insights on X-ray emission from astrophysical sources and to perform large area surveys and population census.

[1] https://www.mpe.mpg.de/455799/instrument

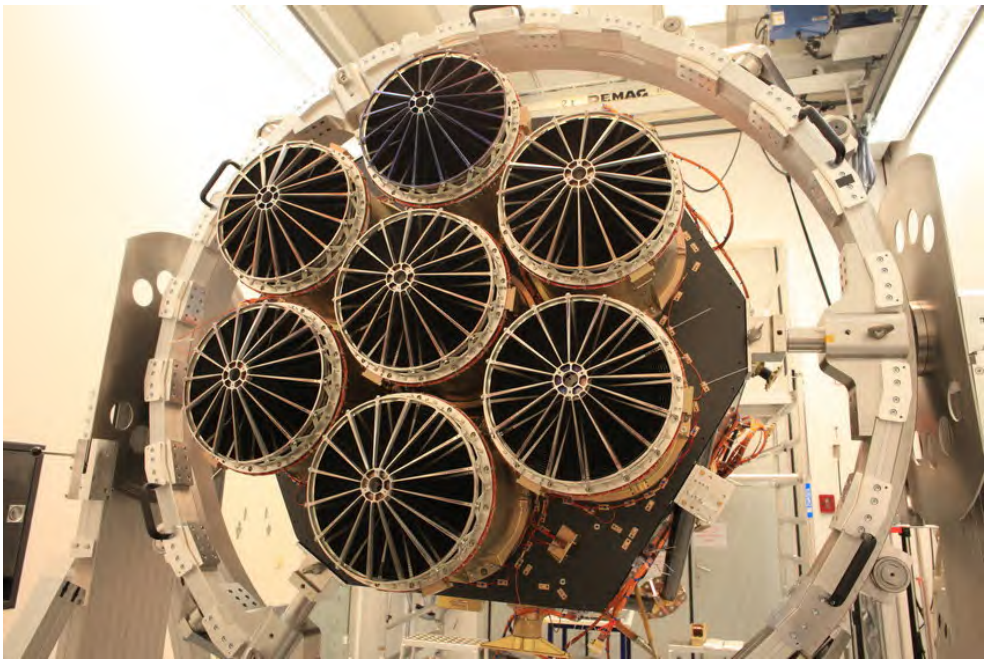

Abbildung 1.8: Photograph of the eROSITA assembly process pre-flight, showing the 7 telescope modules, each having 54 assembled gold-coated mirrors. Image credit: MPE

eROSITA was intended to perform an 8-pass, all-sky survey over the course of 4 years, completing one survey every 6 months. However, on February 26th, 2022, eROSITA was placed into safe mode[2]. Until this point, approximately 4 and a half all-sky surveys over the course of 2 years and three months had been performed. Nevertheless, eROSITA has detected X-ray emission from more than one million of AGN across a range of redshifts, enabling population studies of AGN as a function of luminosity and distance (Merloni et al. 2024). The survey strategy involved completing arcs of the sky covering a great circle once every four hours, and precessing slowly over the course of 6 months to cover the full sky. This strategy meant lower exposure at the equatorial plane of the survey, but very high coverage at the apexes, known as the North Ecliptic pole (NEP), and South Ecliptic Pole (SEP). Data rights are divided between German collaboration (Western Galactic sky) and Russian collaboration (Eastern Galactic sky), with each collaboration processing and analysing data on their respective regions.

The primary science objective of eROSITA is to detect X-ray emission from $\sim 100,000$ clusters of galaxies, allowing for cluster cosmology to be performed (Bulbul et al. 2024; Ghirardini et al. 2024). However, a secondary product will be the detection of $> 1,000,000$ AGN, suitable for population analysis and other large statistical studies (Merloni et al. 2012a). AGN will be by far the most prevalent source in the all-sky surveys, likely $\sim 80\%$ (Merloni et al. 2012a) of all detected sources. For sources with sufficient photon counts (>20; see Liu et al. 2022b), X-ray spectral analysis can give hints as to the presence of soft

[2]https://www.mpe.mpg.de/7856215/news20220303

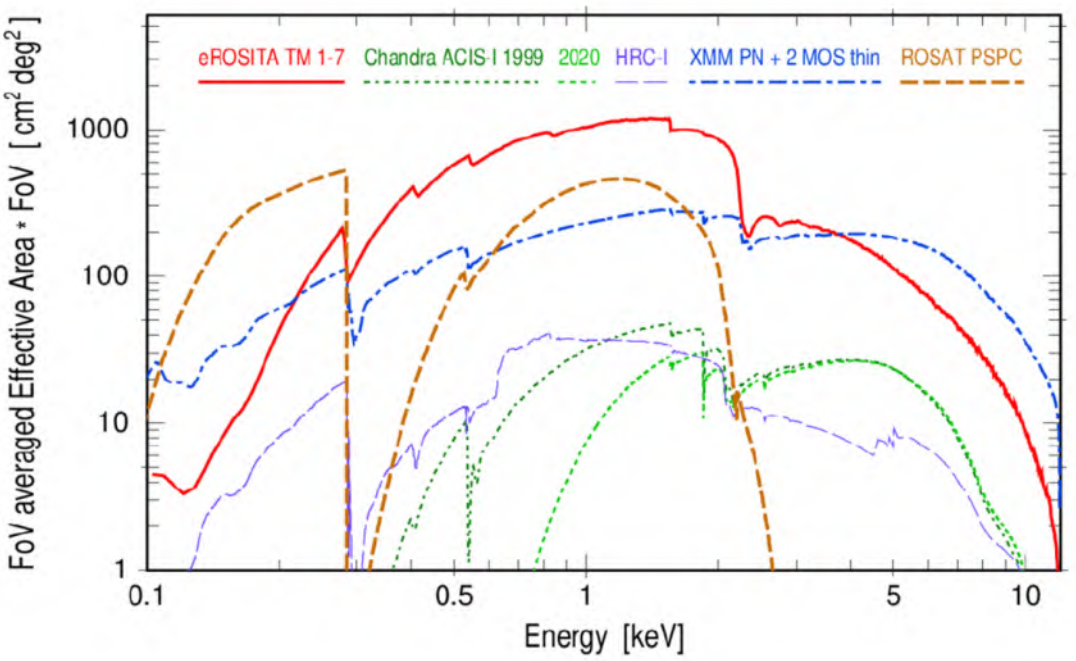

Abbildung 1.9: Grasp; product of the FOV averaged effective area and the field of view, shown comparing different past and contemporary X-ray observatories; eROSITA (all TM combined) is shown as a red solid line, *Chandra* ACIS-I is shown at launch in a dark green dotted line and in 2020 as a light green dotted line, *Chandra* HRC-I is shown in a violet thin dashed line, *XMM-Newton* (all instruments) is shown as a blue dot-dash line, and ROSAT PSPC is shown as a brown-orange thick dashed line. Image is taken from Figure 10 of Predehl et al. (2021a).

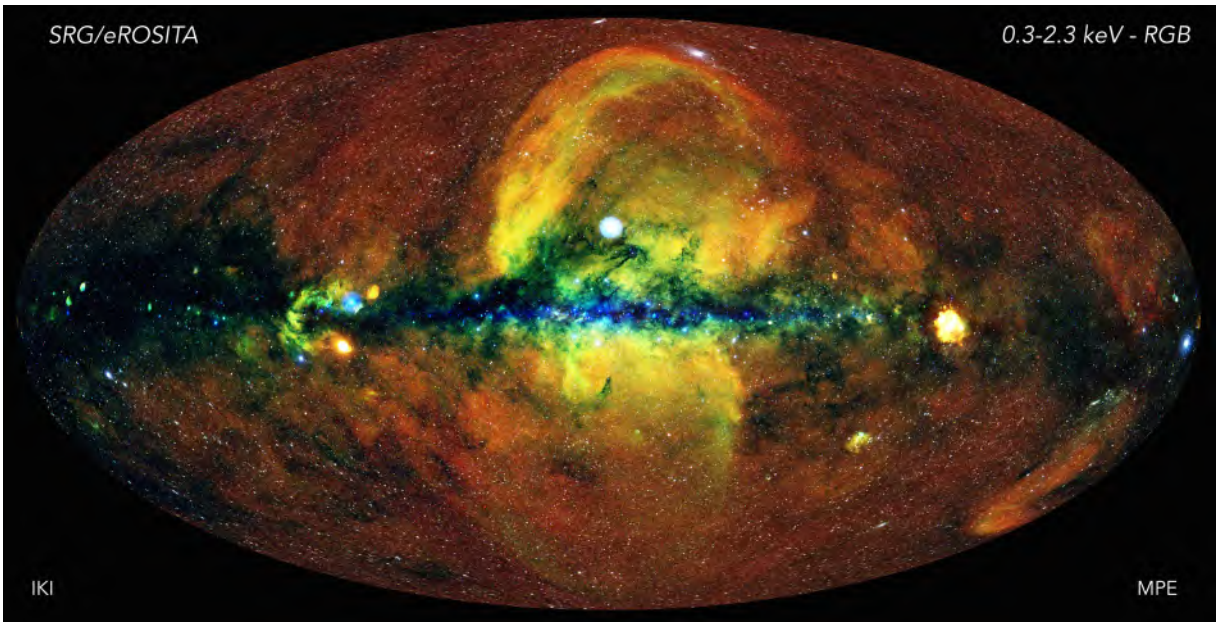

Abbildung 1.10: The map of the first eROSITA all-sky survey, eRAS1, in X-ray energies between $0.3 - 2.3\,\mathrm{keV}$. The image has been smoothed using a 10 arcminute Gaussian. Different colours represent different X-ray energies, where $0.3 - 0.6\,\mathrm{keV}$ is shown in red, green for $0.6 - 1\,\mathrm{keV}$ is shown in green, and $1.0 - 2.3\,\mathrm{keV}$ is shown in blue. Image credit: Jeremy Sanders, Hermann Brunner and the eSASS team (MPE); Eugene Churazov, Marat Gilfanov (on behalf of IKI).

X-ray absorption and the shape of the coronal power law, or other features (see Sect. 1.3 and Chapter 3).

The image of the first eROSITA all-sky survey is shown in Fig. 1.10, taken from the eRASS1 presskit[3]. Different colours represent different X-ray energies, where $0.3 - 0.6\,\mathrm{keV}$ is shown in red, green for $0.6 - 1\,\mathrm{keV}$ is shown in green, and $1.0 - 2.3\,\mathrm{keV}$ is shown in blue. There are many large scale structures visible, mostly associated with clusters of galaxies or the eROSITA bubbles (e.g. Predehl et al. 2020; Liu et al. 2024). Absorption is clearly visible along the central plane of the image, corresponding to the Galactic plane, where increased absorbing column densities block the softest energy photons. The vast majority of the point sources in the image are AGN, with a smaller fraction being coronally active stars or X-ray binaries.

1.4.3 NuSTAR

In 2016, *NuSTAR* (Harrison et al. 2010; Harrison et al. 2013, 2016) was launched, the first imaging telescope in the $3 - 79\,\mathrm{keV}$ range. It features two Focal Plane Modules (FPM), labelled A and B. The two focal plane modules have slightly differing energy calibrations,

[3]https://www.mpe.mpg.de/7461950/erass1-presskit

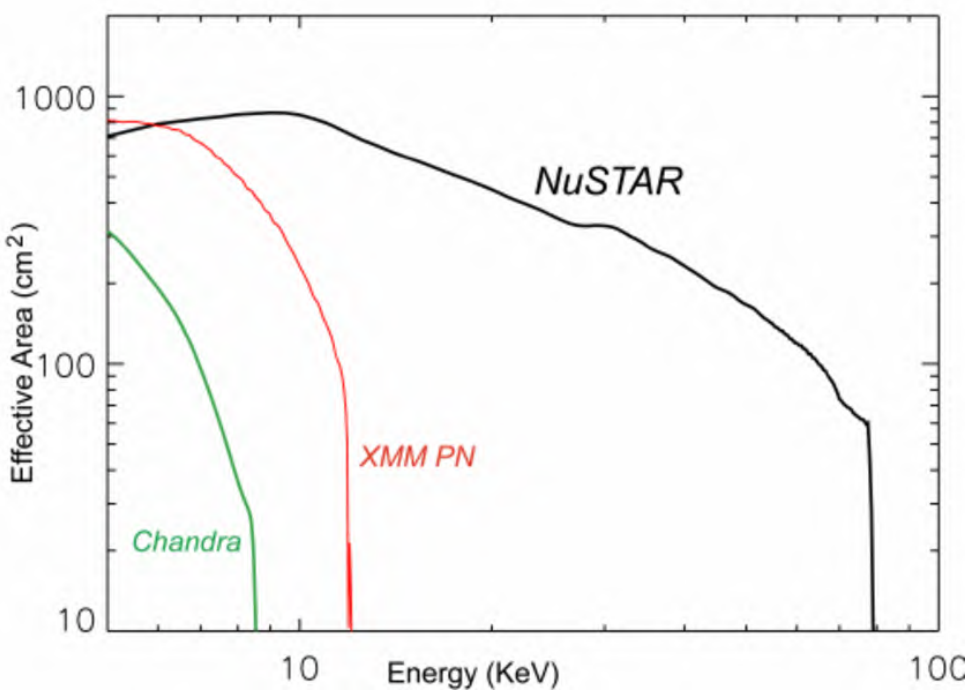

Abbildung 1.11: Effective area comparison between *NuSTAR* (black), *XMM-Newton* (red) and *Chandra* (green). Image taken from (Harrison et al. 2010).

but can be modelled simultaneously using a cross-calibration constant. It uses a modified Wolter 1 mirror design, where a conical approximation is used instead of the paraboloid-hyperboloid used by eROSITAThe mirrors are coated in Pt/SiC and W/Si layers, which are in part responsible for enabling the energy range to be extended to 79 keV, where a Pt edge produces a sharp drop off in the effective area. A comparison of the effective are of *NuSTAR* compared to *XMM-Newton* and *Chandra* is shown in Fig. 1.11, clearly demonstrating that *NuSTAR* is much more sensitive to hard X-ray photons even above 6 keV as compared to *XMM*, with an energy range extending up until the steep drop off at 79 keV and showing the effective area peaking at ~ 10 keV.

This mission significantly expanded on the science that could be performed with higher energy X-rays, and is often used in conjunction with Swift-XRT, *Chandra* and *XMM-Newton* (e.g. see Kang & Wang 2023, for a review of joint spectral modelling between *NuSTAR* and *XMM-Newton*). In particular, *NuSTAR* is able to simultaneously measure emission from the power law, the Fe Kα emission and absorption profile, and the Compton hump peaking around $\sim 20 - 30$ keV, and is therefore extremely well suited to studies of the AGN torus (Panagiotou & Walter 2019; Panagiotou et al. 2021; Zhao et al. 2021). This will be exploited in Chapter 5, where I will use *NuSTAR* data to study the geometry and orientation of the torus in a number of eROSITA selected sources. A more detailed description of *NuSTAR* data processing and analysis is also presented in Chapter 5.

1.5 Hard X-ray selected AGN with eROSITA

Various selection criteria for AGN samples can reveal different interesting properties of these systems. One such selection is in hard X-rays, the precise energies of which depend on the goals of the mission and structure of the telescope (e.g. see Nandra et al. 2024; Waddell et al. 2024). For eROSITA in particular, this is considered to be the $2.3-5\,\mathrm{keV}$ band. This is due to the gold coating of the mirrors, which produces a steep drop-off in the effective area at 2.3 keV, but also due to the rapid drop-off of effective area and higher background above 5 keV. As a result, eROSITA all sky-surveys should produce a flux-limited sample of hard X-ray emitting AGN.

There are several key advantages to the use of hard X-ray selected surveys to study AGN. First, they are less sensitive to obscuration effects, which can entirely suppress emission below $\sim 10\,\mathrm{keV}$, but allow transmission of higher energy photons. This is problematic, since obscured AGN are known to be very common in the local Universe (Lanzuisi et al. 2013) and across cosmic time. Consequently, a large fraction of the accretion power in the local Universe and beyond is known to lie in obscured AGN (e.g. Hasinger et al. 2007b; Ueda et al. 2014; Aird et al. 2015). The bias against obscured sources is therefore of key importance to overcome to understand global AGN demographics.

For X-ray spectral modelling, hard X-ray selected samples provide major benefits. Crucially, as explored above, the hot X-ray corona dominates the X-ray spectrum of most sources at energies above $\sim 1-2\,\mathrm{keV}$. This means that data above 2 keV are necessary to be able to properly constrain the luminosity and the photon index of the hot corona. By doing so, one can then attempt to model additional spectral components, including soft X-ray absorption, Fe Kα emission and absorption features, the Compton hump, and the soft excess (see e.g. Waddell et al. 2019; Waddell & Gallo 2020, 2022, and Chapters 2-6 of this thesis). By using the hard X-ray selected sample, we can therefore study all of these components, an impossible feat with e.g. ROSAT or eROSITA data below 2 keV. This point will be exploited throughout the following chapters.

1.6 Outline of the thesis

The launch of the eROSITA telescope onboard SRG promises to transform our ability to study AGN in X-rays. Aside from enabling large statistical studies, eROSITA will perform the first all-sky imaging survey above 2 keV, providing the first all-sky census of AGN with hard X-ray emission above the eRASS:8 flux limit. Hard X-ray samples of AGN are less biased to obscuration than samples selected only from soft X-ray emission, and enable constraints to be placed on the spectral index of the hot X-ray corona. This then permits the subsequent modelling of additional X-ray spectral components, including absorption, the X-ray soft excess, or outflows. To-date, eROSITA has completed four and a half of the planned all-sky surveys, with the first all-sky survey providing nearly one million detected sources (see Fig. 1.10).

In Chapter 2, I present the hard ($2.3-5\,\mathrm{keV}$) X-ray selected sample of sources detected

in the eROSITA Final Equatorial Depth Survey, eFEDS, a performance verification survey designed to mimic the final equatorial depth of the 8-pass, all sky survey. The counterparts of X-ray sources are found and classified, and an AGN sample is constructed. I examine the basic properties of the resulting bright local AGN sample, and compare to the properties of other samples.

In Chapter 3, I perform an in-depth X-ray spectral analysis of the 200 hard X-ray selected AGN identified and classified in Chapter 2. I test different X-ray models to study the nature of the soft excess, considering a thermal or non-thermal origin, as well as absorption models. I then perform model comparison to identify the most likely physical mechanism responsible for the observed emission, and compare the properties of different models.

In Chapter 4, I apply what was learned in Chapter 2 to the hard ($2.3 - 5\,\mathrm{keV}$) X-ray selected sample from he first eROSITA all-sky survey, eRASS1. I begin by cleaning the X-ray sample, then I perform counterpart identification and classification using a variety of techniques and multiwavelength catalogs. I then identify a sample of 1328 AGN with spectroscopic redshifts, thus defining a half-sky sample of bright, nearby AGN. A total of 29 of these sources were found to be only detected above 2.3 keV, and were thus dubbed hard-only.

In Chapter 5, I combine the sample of 29 hard-only AGN with other AGN showing a majority of the emission above 2.3 keV, and search for those which have corresponding *NuSTAR* data sets. Using a sample of 16 sources, I then investigate the properties of the observed absorption. I examine the structure and geometry of the torus, and for some sources identify that absorption may originate in the host galactic disc.

Throughout this thesis, I adopt a flat in the ΛCDM cosmology with $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, and $H_0 = 70$ km s^{-1} Mpc^{-1}.

Kapitel 2

eFEDS: The hard X-ray selected sample

After the launch of the eROSITA telescope onboard SRG on July 13th, 2019, and prior to beginning the planned all-sky survey, eROSITA underwent a calibration and performance phase. During this time, eROSITA surveyed a 140 deg^2 area of the equatorial sky. The resulting eROSITA Final Equatorial Depth Survey (eFEDS) field provided a first look at the expected performance of eROSITA after the completion of the 8-pass survey. One key output of the all-sky survey is the eROSITA hard X-ray selected sample, providing the first all-sky hard (2 keV) focussing imaging survey. Based on the energy range selected and learning from other hard X-ray surveys, we may expect that a large percentage of these sources are AGN; through characterisation of hard X-ray selected sources in eFEDS, knowledge can then be applied to the all-sky survey. A sample of 246 point-like, hard X-ray selected sources is constructed from our $2.3 - 5$ keV selection (Brunner et al. 2022). Counterparts are then identified for these sources, and they are classified based on spectroscopic redshifts and multiwavelength properties as extragalactic (90%; predominantly AGN) or Galactic (10%; active stars). The AGN are then studied in more detail by combining X-ray properties with optical/UV spectroscopy, and insights are identified to aid with the first all-sky survey. These AGN then form the base sample for analysis in Chapter 3. The lessons learned from this analysis will be applied in Chapter 4, where I present a similar analysis to the first all-sky survey (eRASS:1) hard X-ray selected sample.

This Chapter is derived from Nandra, Waddell, Liu et al. (2024), in the co-authors created, classified, and studied the 246 sources in the eFEDS hard X-ray selected sample. Sample creation, classification, and optical spectral modelling were performed by co-authors. I assisted with extensive catalog validation and analysis, and was responsible for all figures in this Chapter and comparison to other samples. Optical spectral fitting including line measurements and black hole mass estimates was performed by Q. Wu and Y. Shen. The text has been abridged and re-written in order to focus on key aspects of relevance for later chapters. Sections on X-ray spectral modelling have been removed, as this will be done more extensively in Chapter 3.

2.1 The eROSITA Final Equatorial Depth Survey

The details of the eFEDS X-ray survey were presented and released by Brunner et al. (2022), who described the source detection and subsequent catalog creation in detail. The processing and reduction of the X-ray data were conducted using the eROSITA Science Analysis Software system (eSASS) with the early data release version (EDR). The source detection process is performed using a sliding box detection algorithm, followed by PSF-fitting to determine source counts and significance, using a likelihood ratio approach. This method allows for source detection across multiple energy bands simultaneously, which in turn allows for better positional accuracy and for characterization of the detection significance across different X-ray energy bands. The main eFEDS sample is constructed using a soft X-ray selection in the $0.2-2.3\,\mathrm{keV}$ band, where the instrument sensitivity is highest. This resulting catalog includes 27,910 sources with a $0.2-2.3\,\mathrm{keV}$ detection likelihood (`DET_LIKE`) greater than 6, where `DET_LIKE` parameterises the chance probability of the source being a background fluctuation. Brunner et al. (2022) also preesented a supplementary catalog of sources less robustly detected in this band, having `DET_LIKE` $>$ 5.

The algorithm also assesses whether sources are extended, assigning an `EXT_LIKE` likelihood. Sources with `EXT_LIKE`$\geq$ 6 are considered extended and are candidate galaxy clusters, as detailed by Liu et al. (2022a). Sources with `EXT_LIKE` $<$ 6 are considered point-like X-ray sources and are treated as such in the Brunner et al. (2022) catalog. These point-like X-ray sources comprise the bulk of the X-ray catalog, and are of key interest for the study and identification of AGN; but also include stars, distant galaxy clusters, and X-ray binaries.

In addition to the main, soft X-ray catalog, Brunner et al. (2022) also published a hard X-ray selected catalog, in the $2.3-5\,\mathrm{keV}$ band. In order to construct this sample, separate source detection runs are performed on three energy bands; $0.2-0.6\,\mathrm{keV}$ (band 1), $0.6-2.3\,\mathrm{keV}$ (band 2) and $2.3-5\,\mathrm{keV}$ (band 3), resulting in source detection likelihoods for each band as well as a combined likelihood for all bands ($0.2-5\,\mathrm{keV}$; band 0). To create this sample, we first apply a cut-off of `DET_LIKE_0`>5, to minimize spurious sources identified as a result of background fluctuations. We then apply a hard band cut of `DET_LIKE_3`>10, to select only sources which have robust detections above 2.3 keV. The estimated spurious fraction as a result of these cuts is low, at $\sim 2.5\%$ (Liu et al. 2022c). The resulting catalog contains 254 sources, eight of which are extended sources with `EXT_LIKE`$>=6$ which are candidate galaxy clusters (Liu et al. 2022a). Since we are aiming to construct a sample of AGN, we remove these cluster candidates.

The resulting hard X-ray point-like sample contains 246 sources. This is a factor of more than 100 fewer sources than in the main, soft X-ray selected sample, although with similar expected purity. One primary reason for this is that the effective area of eROSITA drops significantly above 2.3 keV, due to M absorption edges produced through interactions with the gold coating on the eROSITA mirrors (ee Chapter 1, c.f. Figure 9 of Predehl et al. 2021a). The eROSITA background has also been found to be higher than pre-launch estimates, in part believed to be due to the solar cycle (see Brunner et al. 2022, for a

detailed explanation). Another interesting property of the hard sample is that it contains 20 sources which are only present in the hard catalog, as we do not find any matched source in the soft catalog. These sources are dubbed hard-only, and are candidate heavily obscured AGN.

Various other surveys (e.g. see Chapter 1) have covered this energy range across various sky areas and depths. One early survey of importance for future works was HEAO-A1/A2 (Rothschild et al. 1979), which surveyed the entire sky and produced a uniform, high-flux sample of the brightest sources in the hard X-ray sky. A key finding from this sample was that a majority of sources were found to be AGN, and we expect the same from our eFEDS sample. This sample of AGN was first presented by Piccinotti et al. (1982), and has been followed-up by numerous other X-ray missions and surveys; these sources will also be of interest in eRASS1 (see Chapter 4).

2.2 Optical identifications, classifications and redshifts

2.2.1 Counterpart identification

Having now assembled a sample of hard sources, it is of interest to identify the multiwavelength counterparts of these sources, with the ultimate goal of constructing a sample of AGN for further analysis. To this end, optical counterparts for the hard X-ray sample have been identified following the method of Salvato et al. (2022), which was applied to the eFEDS main sample. A total of three different matching algorithms were applied to the hard sample: the Bayesian NWAY code (Salvato et al. 2018), a maximum likelihood approach (e.g., Sutherland & Saunders 1992; Brusa et al. 2007) and HamStar, a dedicated matching algorithm designed to identify X-ray emitting stars (Schneider et al. 2022). The key data set to which the X-ray source positions are matched is the DESI Legacy Survey public Data Release 8 (LS8), which includes forced photometry from unWISE (Lang 2014), and Gaia EDR3. Gaia also provides parallaxes and proper motion information that is useful in distinguishing between Galactic and extragalactic counterparts to the X-ray sources. These matching algorithms and their applications to the data are described in detail in Salvato et al. (2022). A quality or strength of the association is then assigned based on the agreement between the various methods, where a counterpart quality of 4 is a very strong association, 1 is a very weak association, and 0 implies that no counterpart was found. For this purpose of this chapter, we consider secure counterparts to have quality of > 2.

The number of sources given each counterpart quality is shown in Table 2.1, for hard, hard-only and main samples. Comparing the main and hard samples, we find generally that there is higher quality in hard sample. This result is perhaps unsurprising, since the hard sample is generally comprised of the X-ray brightest sources. However, we also see that the counterpart quality is low for hard-only sources. This result will be discussed further in Sect. 2.3.2.

Tabelle 2.1: Optical counterpart quality, shown separately for all 246 sources in the hard band, for the hard-only sample, and for the main sample. Secure counterparts are defined as having quality > 2.

Sample	Total	CTP_quality					Secure
		0	1	2	3	4	%
Hard	246	4	11	3	0	228	92
Hard-only	20	4	7	1	0	8	40
Main	27,369	1225	1370	2552	1379	20837	81

2.2.2 Source classification

Having assigned multiwavelength counterparts to the eFEDS X-ray sources, one can then move onto classify these sources. We focus our efforts only on sources which have secure counterparts, as defined in the previous section. The source classifications are again performed in the same manner as Salvato et al. (2022), following the decision tree presented in their Figure 8. The primary tool for classification is spectroscopic redshift (spec-z), where sources with spec-z>0.002 are confirmed to be extragalactic (mostly AGN or galaxies), whereas sources with spec-z<0.002 are confirmed Galactic (e.g. X-ray binaries or stars). We therefore focus our effort on collecting spec-z, both with dedicated follow-up programs with SDSS (from which a majority of our redshifts are obtained; see Gunn et al. (2006); Smee et al. (2013); Blanton et al. (2017); Dwelly et al. (2017); Kollmeier et al. (2017); Comparat et al. (2020)) and through searching public archival spectra. High confidence redshifts come from a variety of sources (see Table 2.2). In total, we find 197/246 have an available spec-z; these are tabulated in Table 2.2, along with the references as appropriate. A small minority, only 12 sources, are stars, while the rest are secure extragalactic sources (AGN). Overall, we have much higher spectroscopic completeness (80%) than the main catalog (24%), owing to our much smaller and brighter sample; this is a common phenomenon in large surveys (see also e.g. Akiyama et al. 2003; Della Ceca et al. 2004; Eckart et al. 2006; Cocchia et al. 2007). For 3 sources with spectra, no redshift could be determined; these are found to be blazars (see Sect. 2.4.4).

For sources with no available spectrum, we must rely on other methods (see Salvato et al. 2022). First, considering stars; sources with 5σ parallax measurements from Gaia are secure Galactic, as are sources flagged with the stellar counterpart identification schema HamStars. Then, sources which are optically extended based on optical light profile fitting in LS8 are also considered to be secure extragalactic. For the remainder, when possible, a separation scheme based on optical to mid-infrared colors is applied (Salvato et al. 2022). If these are not available, an X-ray/mid-IR cut is instead used (Salvato et al. 2022). These sources are classified as likely Galactic or extragalactic, since their classifications are less certain than those with spec-z, parallax or extent measurements. Source classifications for the secure counterparts are summarized in Table 2.3, separated by secure and likely extragalactic counterparts, and secure and likely Galactic counterparts. We show the same values for the main sample. A majority (93.5%) of our sample is found to be extragalactic,

Tabelle 2.2: Optical spectroscopy of the eFEDS Hard sample, shown by reference catalog or survey. Where multiple catalogs contain the same object, the order of preference used is: (1) Blanton et al. (2017); (2) Almeida et al. (2023); (3) Baldry et al. (2018); (4) Jones et al. (2009); (5) Wenger et al. (2000); (6) Huchra et al. (2012);

Origin	Redshifts z	Stellar ($z < 0.001$)	Ref
SDSS	180	2	(1,2)
Gaia-RVS	8	8	–
GAMA	2	–	(3)
6dFGS	2	–	(4)
LAMOST	2	1	–
Simbad	2	1	(5)
2mrs	1	–	(6)
Total	197	12	–

Tabelle 2.3: Source classifications for the 232 hard-band sources with `CTP_quality` 3 or higher with comparison to classifications of the main sample.

Sample	Extragalactic			Galactic		
	Secure	Likely	All	Secure	Likely	All
Hard	181	36	217	14	1	15
Hard (%)	78.0	15.5	93.5	6.0	0.4	6.4
Main (%)	48.4	46.6	95.0	3.3	1.7	5.0

which is unsurprising, since we expect a majority of sources to be AGN. The main sample has a similarly high (95%) fraction of extragalactic sources. However, also of interest is the relative percentages of secure versus likely counterparts; examining this, it is found that the hard sample has a higher fraction of secure classifications. This is mostly due to the high spectroscopic completeness of the hard sample, and demonstrates that this sample is of key interest for X-ray spectral analysis.

In order to find redshifts for sources without spec-z, which are necessary for computing the source luminosity, photometric redshifts (photo-z) are computed for the extragalactic sources (Salvato et al. 2022). This is done using template fitting of the IR-optical-UV Spectral Energy Distribution (SED) using the LePhare code (Ilbert et al. 2006). To ensure that only secure photo-z are selected, a redshift grade based on scatter and outlier fraction is assigned, and only those sources with `REDSHIFT_GRADE` ≥ 4 (Salvato et al. 2022) are considered. These are then combined with the spec-z to finalize the redshifts for the sample.

2.2.3 Line Measurements and Black Hole Mass estimates

With high quality spectra obtained from SDSS used to measure spectroscopic redshifts, it is also possible to estimate the mass of the central black hole (e.g., Kaspi et al. 2000b).

This applies to the sources which show broad optical lines, specifically Hβ, MgII and CIV, which are produced in the broad line region (BLR). Assuming the BLR is virialized (see also Chapter 1), we can assume that the line width is produced due to its rotation around the central black hole, allowing us to estimate the black hole mass.

To model the optical spectra, the code PyQSOFit (Guo et al. 2018; Wu & Shen 2022) is used. This python package uses χ^2 fitting, employing several components; a power law and a polynomial continuum, the Balmer continuum, FeII emission in the optical and UV, and finally emission from broad and narrow lines using gaussians. The continuum model is first fit to spectral regions which do not contain strong emission features. FeII emission is then modelled using various templates depending on the wavelength regime:

- $1000 - 2200$ Å: Vestergaard & Wilkes (2001)
- $2200 - 3090$ Å: Salviander et al. (2007)
- $3090 - 3500$ Å: Tsuzuki et al. (2006)
- optical: Boroson & Green (1992)

Once these components have been modelled, they are subtracted from the data, leaving only the emission lines, which are then fit with the Gaussian components. Error bars are then estimated based on Monte Carlo simulations, which are then later used to assess the confidence that can be placed in the derived quantities. Using the derived widths of the broad lines and the measured continuum luminosities, the black hole masses are then calculated using:

$$\log\left(\frac{M_{BH}}{M_\odot}\right) = A + B\log\left(\frac{\lambda L_\lambda}{10^{44}\text{erg s}^{-1}}\right) + C\log\left(\frac{\text{FWHM}}{\text{km s}^{-1}}\right) \tag{2.1}$$

where the constants A, B and C are calibration constants depending on which line is used for the black hole mass. Based on the redshift of the source, we use either the Hβ line, calibration constants from Vestergaard & Peterson (2006), and 5100Å continuum luminosity (low redshift), or the MgII line, calibration by Shen & Liu (2012), and luminosity at 3000Å (high redshift). These luminosities are also used to derive Bolometric luminosities, using the bolometric corrections derived from fitting SDSS quasars (Richards et al. 2006). The Monte Carlo errors from the fit are then propagated to derive uncertainties for the black hole mass. These single-epoch black hole mass estimates also carry a systematic uncertainty of around 0.4 dex (Shen 2013). The black hole masses and bolometric luminosities can then be used to estimate the Eddington ratio;

$$\lambda_{\text{Edd}} = \frac{L_{\text{Bol}}}{L_{\text{Edd}}} = \frac{L_{\text{Bol}}}{1.26 \times 10^{38}(M/M\odot)} \tag{2.2}$$

which can be used to compare relative accretion rates between black holes of different masses. These properties will be discussed further in section 2.4.5 and in Chapter 3, where I describe the process of visually inspecting the spectra to perform quality control.

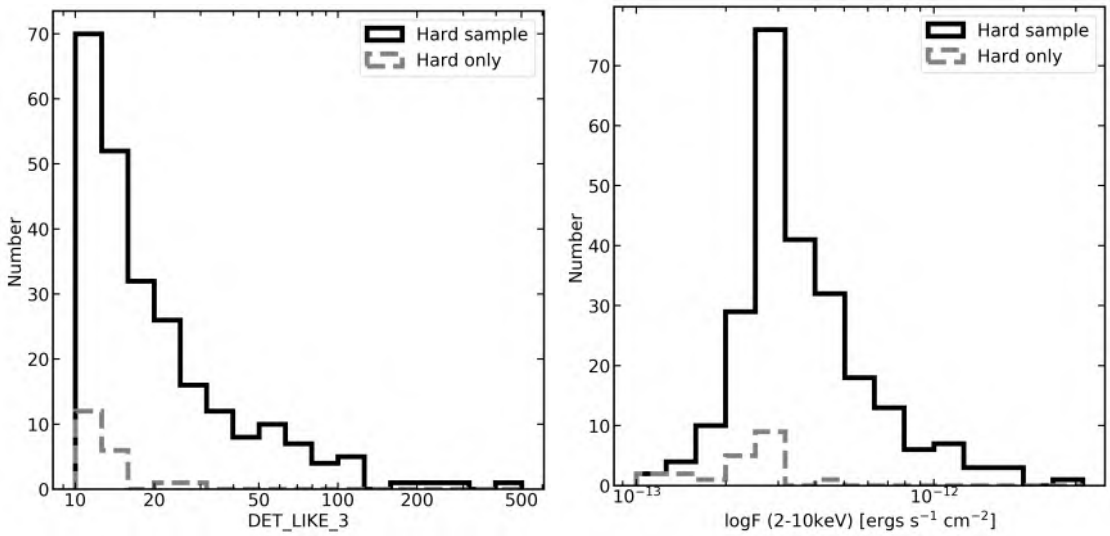

Abbildung 2.1: Left: detection likelihoods `DET_LIKE_3` in the $2.3 - 5\,\mathrm{keV}$ band. Right: Histograms of the X-ray flux ($2 - 10\,\mathrm{keV}$). All sources in the sample are shown in black, while those detected only in the hard band (hard-only) are shown with grey dashed lines.

2.3 Basic properties of the hard sample

2.3.1 X-ray properties

There are several key characteristics of the eFEDS hard sample which distinguish it from the main, soft X-ray selection. Notably, with secure detections in the hard X-ray regime, we are able to investigate the luminosity and spectral shape above 2 keV. Fig. 2.1 shows the detection likelihoods (`DET_LIKE_3`; left) and the $2 - 10\,\mathrm{keV}$ fluxes (right) derived from the $2.3 - 5\,\mathrm{keV}$ counts. Sources in the hard sample are shown with black lines, and hard-only sources are shown with grey dashed lines. Most sources also have relatively low detection likelihoods, with values around 10-20; this is particularly true for the hard-only sources, with very few having values > 20. This may point towards issues with the hard-only sample, discussed in Sect. 2.3.2. The peak of the flux distributions is around $2 - 3 \times 10^{-13}$ ergs s^{-1} cm^{-2}, and is comparable between the hard and hard-only sources.

We also seek to study the differences between fluxes in the main and hard samples. A comparison of the soft X-ray ($0.5 - 2\,\mathrm{keV}$) fluxes between the hard and main sample is therefore shown in Fig. 2.2, with the main sample shown in orange dashed lines and the hard sample shown in blue solid line. As previously stated, the hard sample is brighter by nearly 1.5 orders of magnitude, with some fluxes extending up to $> 10^{-12}$ ergs s^{-1} cm^{-2}. There are also some dimmer sources with fluxes between $10^{-14} - 10^{-15}$ ergs s^{-1} cm^{-2}, most of which are the hard-only sources which have poorly constrained soft X-ray fluxes and large error bars. In the main sample, most sources are dimmer, with $\sim 40\%$ of sources having fluxes of $5 \times 10^{-15} - 10^{-14}$ ergs s^{-1} cm^{-2}. A very few sources extend to brighter fluxes, most of which are likely found in the hard sample.

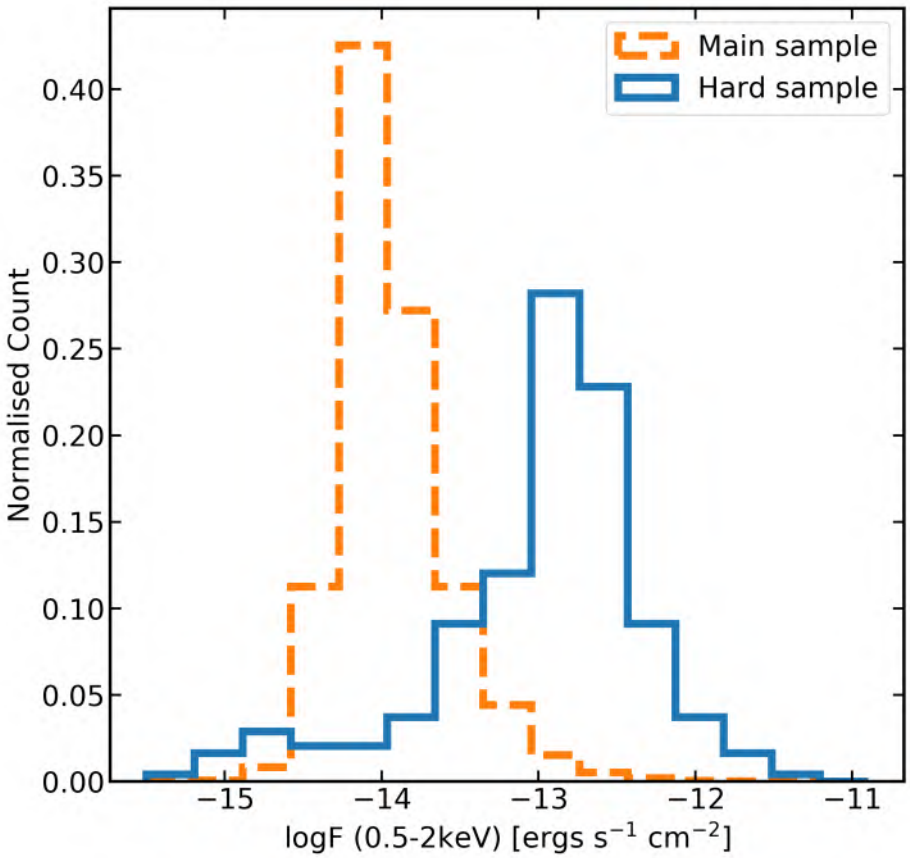

Abbildung 2.2: Normalised histogram of the soft X-ray ($0.5 - 2\,\mathrm{keV}$) fluxes shown for the main sample (orange dashed line) and hard sample (blue solid line).

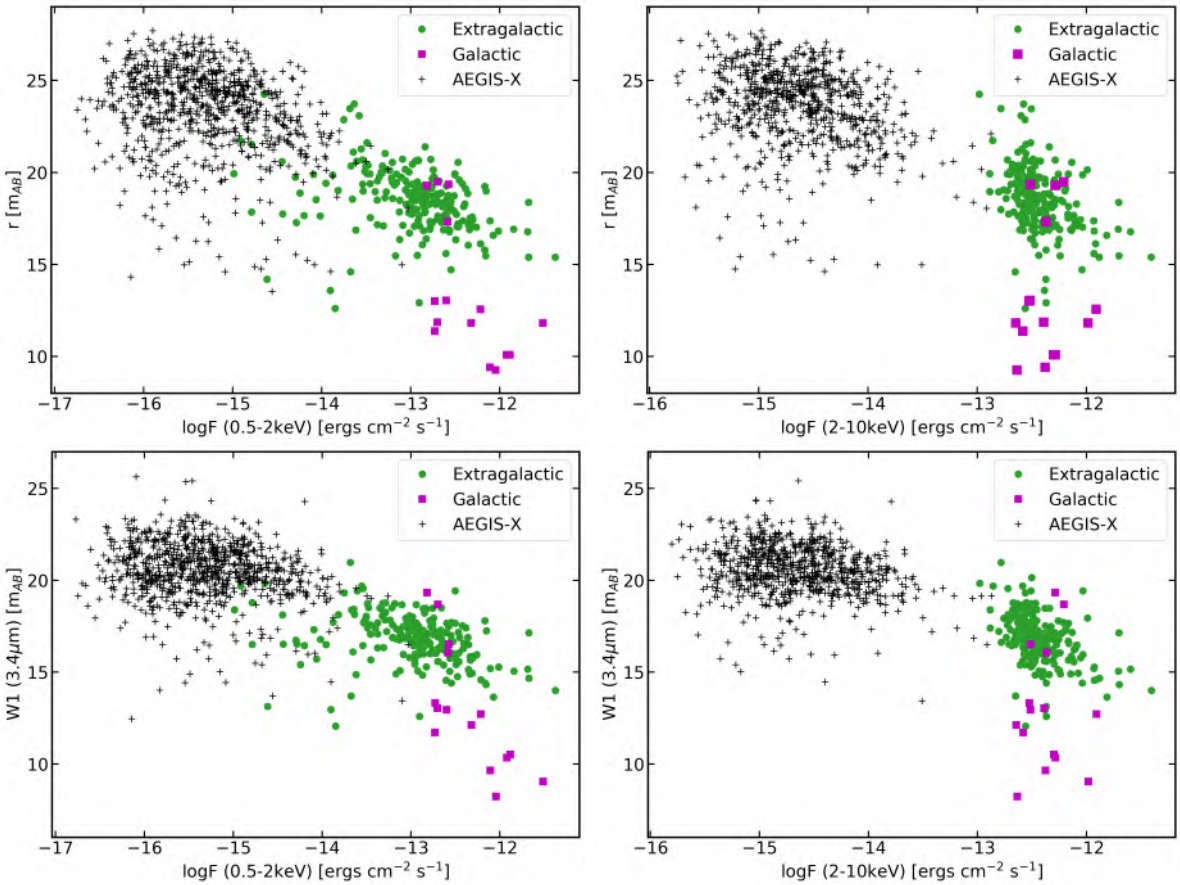

Abbildung 2.3: Legacy survey r-band (uppers panels) and Wise W1 3.4μm (lower panels) magnitude versus the 0.5–2 keV (left panels) and 2-10 keV (right panels) X-ray flux for the eFEDS hard sample. Source classified as being extragalactic (green symbols) and Galactic (purple symbols) are shown separately. As a comparison the data for the AEGIS-XD survey (Nandra et al. 2015) are shown (small black symbols), which cover slightly different bands in the optical/IR, specifically Subaru R_c and Spitzer/IRAC 3.6μm.

Next, we wish to compare the X-ray and multiwavelength properties between different classes. This is done in Fig. 2.3, where we compare the soft ($0.5-2\,\mathrm{keV}$) and hard ($2-10\,\mathrm{keV}$) X-ray fluxes with the LS8 r magnitudes and the WISE W1 magnitudes. These are shown for both extragalactic (green circles) and Galactic (magenta squares) sources. For comparison, we also show data from AEGIS-XD (Nandra et al. 2015), a deep Chandra field. Immediately, it can be seen that the eFEDS hard sample is much brighter in all bands. The optical magnitude range for the eFEDS hard sample is also broader, spanning $r_{AB} = 9-24$. The AGN/extragalactic samples are generally fainter in X-rays and r/W1. Most of the sources classified as Galactic are much brighter in both r and W1, however there are four sources which show more extragalactic-like AGN optical/IR magnitudes; these could be highly accreting compact binaries or similar.

2.3.2 Hard-only sources and catalog integrity

Out of the 246 sources in our point-like, hard-band X-ray catalog, there are 20 sources which do not appear in the main catalog (Brunner et al. 2022), which are dubbed hard-only. These sources are good candidates to be heavily obscured AGN, or other rare phenomena, as eROSITA is highly sensitive to soft X-ray emission, and thus the hard-only sources must have high levels of photoelectric absorption. It must be cautioned however that the purity of our sample is not 100%; indeed, simulations performed by Liu et al. (2022c) suggest a spurious source fraction of $\sim 2.5\%$ at `DET_LIKE_3`> 10. Based on this, around 6/246 sources in the sample should be spurious.

Hard-only sources are more likely to be spurious, since the simultaneous detection in the soft and hard X-ray bands means that the source is more likely to have a true astrophysical origin rather than a background fluctuation. The spurious fraction decreases with increasing `DET_LIKE_3`, and hard-only sources are mostly found at low `DET_LIKE_3` (see Fig. 2.1). Furthermore, these sources typically have much lower counterpart quality; only 40% have secure counterparts (`CTP_QUALITY`>2), compared to $> 98\%$ of sources with both hard and soft X-ray detections,suggeting that some of them may be spurious. Conversely, the 9 hard-only sources with secure counterparts are very likely to be real, given that the low probability of spurious association also supports that the X-ray detection is real. Of these 9, all are extragalactic; this makes sense, since we expect these sources to be mostly obscured AGN.

The properties of the hard-only sources including their spectra are discussed in further detail below; these sources are also investigated in more detail in Chapter 3 for eFEDS and in Chapters 4 and 5 eRASS1.

2.3.3 Hardness ratios

To perform a preliminary spectral shape characterisation, we define a hardness ratio;

$$\mathrm{HR} = (H - S)/(H + S) \tag{2.3}$$

where H is the flux in the $2.3-5\,\mathrm{keV}$ band and S is the flux in the in the $0.6-2.3\,\mathrm{keV}$ band. This hardness ratio can be used to give a first clue as to the overall spectral shape for each source, using values directly derived from the catalog. A more thorough X-ray spectral analysis would enable improved characterisation of the spectral shape, including proper analysis of potential emission and absorption components, as well as more accurate estimation of the coronal power law shape; these results are presented in Nandra et al. (2024) following the approach described in Liu et al. (2022b), however here we relegate this discussion to Chapter 3, where a more thorough analysis is presented.

Fig. 2.4 shows the HR distribution of our sample. All sources are shown as a black solid line, Galactic sources with reliable classifications are shown as a dashed pink line, and extragalactic sources with reliable classifications are shown as a green dash-dot line. The grey dotted line at ~ -0.18 shows the expected hardness ratio for a photon index of $\Gamma = 1.8$ and a typical Galactic column density of $3 \times 10^{20}\,\mathrm{cm}^{-2}$. It can be seen that most sources in the extragalactic sub-sample are harder than the dotted line; this seems reasonable, since many sources are likely to be AGN which show additional absorption from various components within the host galaxy. There is also a strong peak at $HR = 1$ which can be seen in the black histogram including all sources, many of these are the hard-only sources, a large fraction of which did not have reliable classifications, but also lack detectable soft X-ray emission. Finally, we see that the Galactic sources generally appear to be very soft; this is also as expected, as many of these are expected to be coronally emitting stars which are known to be very bright in soft X-rays.

2.4 The AGN sub-sample

2.4.1 Sample definition and construction

The primary objective of this chapter is to construct a sample of AGN in the hard sample, which will be explored in greater detail in Chapter 3. Therefore, for the remainder of this chapter, we restrict our discussion to a sub-sample of the sources which meet all the following criteria:

- Are found within the 90% exposure area region of eFEDS. This excludes the edges of the field where there is a steep exposure gradient, and provides a clean sample to define a selection function (241/246 sources)
- Have a counterpart quality > 2 (228/241 sources)
- Are classified as secure or likely extragalactic sources (213/228 sources)
- Have a spectroscopic redshift or secure photometric redshift (see above) (200/213 sources)

The resulting sample of 200 objects should consist primarily of AGN. The properties of these sources are discussed in more detail in the following subsections, as well as in Chapter 3.

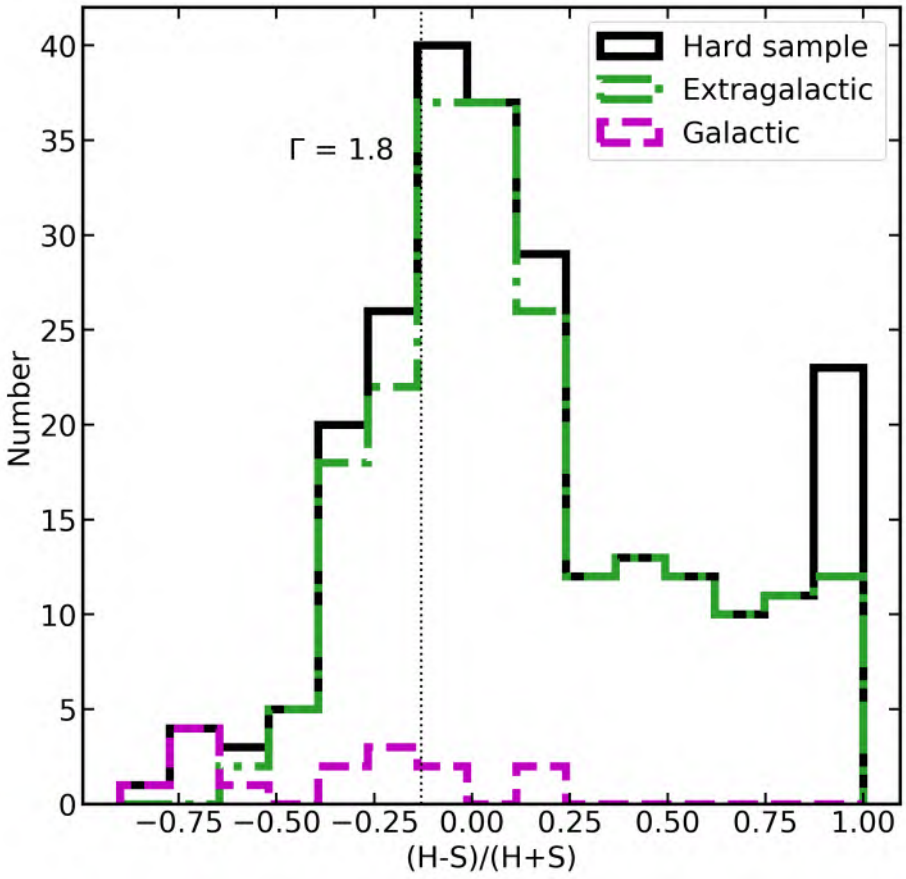

Abbildung 2.4: Histogram of the hardness ratios, shown for the full sample (black solid line), extragalactic sources (green dash-dot lines) and Galactic sources (magenta dashed lines). The vertical dotted line indicates the expected hardness ratio for an absorbed power law with photon index $\Gamma = 1.8$ and Galactic column density of $N_{\rm H} = 3 \times 10^{20}$ cm^{-2}, typical for eFEDS sources.

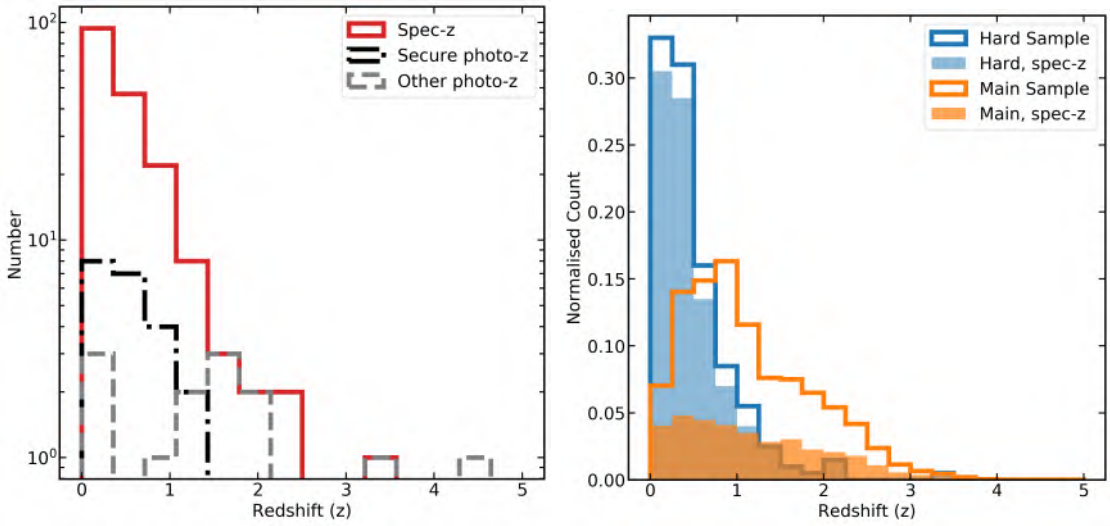

Abbildung 2.5: Left: Redshift distribution for the extragalactic AGN sample. Sources with spectroscopic redshift are shown as solid red lines, and secure photometric redshifts are shown with black dash-dot lines. Sources with less secure photo-z are also shown in grey dashed lines for comparison. Right: Normalised histograms showing spec-z and all reliable redshifts comparing the hard and main samples.

2.4.2 Redshifts and luminosities

The eFEDS hard sample is highly spectroscopically complete, and for the few sources which lack a spec-z, photometric redshifts are available. Fig. 2.5 (left panel) shows a histogram of the redshift distribution separating the sources into three categories; sources with spectroscopic redshifts are shown as a red solid line and sources with secure photo-z are shown as black dash-dot lines. For reference, sources with lower confidence photo-z are also shown as grey dashed lines, but are not included in our sample of 200 AGN. Two of the photo-z sources have very high redshifts $z > 3$, but these are not likely to be the correct redshifts. The remainder of secure photo-z have values < 1.5, which is more reasonable for this sample. There is also one source with a very high spec-z of 3.277, which is likely to be a blazar (see Section 2.4.4).

Fig. 2.5 (right) shows a comparison of the hard (blue) and main (orange) sample redshift distributions, shown for all reliable redshifts (unfilled histograms) and spec-z (filled histograms). We show only redshifts for the candidate AGN in the main sample, using the same criteria described above applied to the hard sample. It is immediately apparent that the hard sample occupies a lower redshift space with a median value of $< z >= 0.34$, meaning it is comprised mostly of nearby AGN. This, along with the Swift-BAT all-sky survey (e.g. Koss et al. 2022), provide information about the AGN population in the local Universe, which is harder to study due to the large area required to find large numbers of such nearby sources. The main sample has much lower spectroscopic completeness, and peaks at a higher median redshift of $< z >= 0.94$, and extending up to larger values,

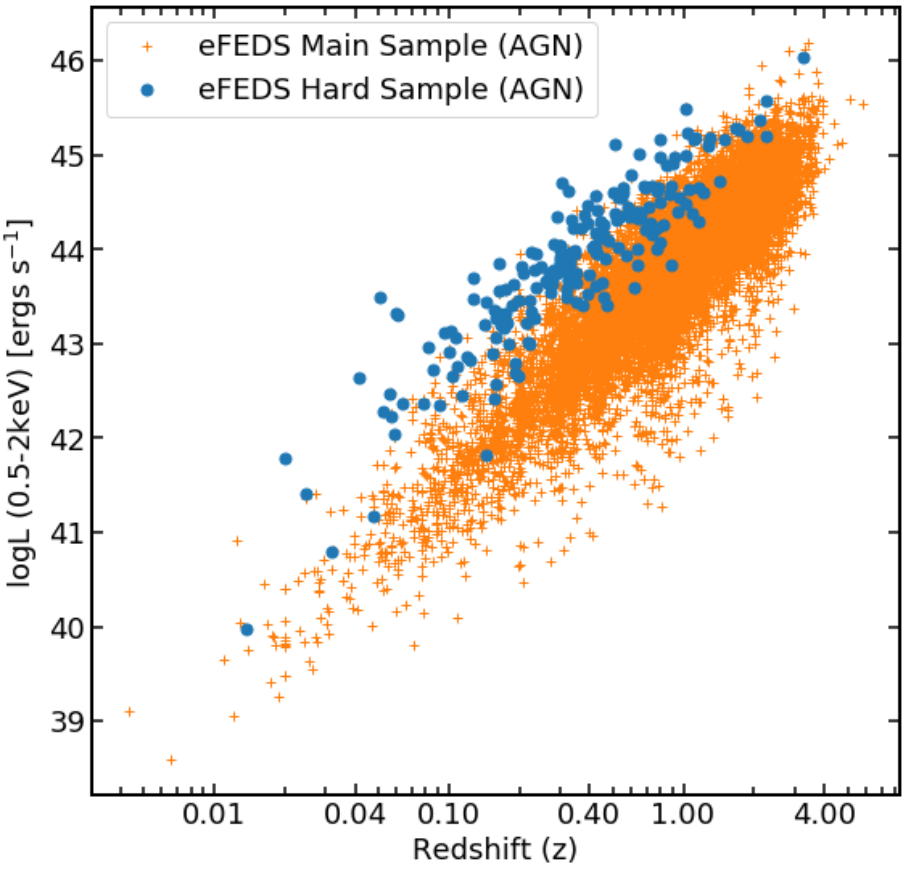

Abbildung 2.6: Soft X-ray (0.5–2 keV) luminosity-redshift relation for the eFEDS hard sample (blue circles) compared to the main sample (orange crosses). The luminosities are corrected for absorption and show the sub-samples with either spectroscopic redshift or high quality photometric redshift.

including a few % of the sample at $z > 2$.

Having compiled the fluxes and redshifts for the main and hard samples, we can then compute the luminosities for each source. This is done by measuring the fluxes from spectral fitting performed and described in Nandra et al. (2024), and using the redshift to compute the luminosity distance. Fig. 2.6 shows the luminosity-redshift relation, with soft X-ray luminosities shown since these are available for all sources in the main sample. The hard sample mostly lies at luminosities $\log L_{\rm X} = 42 - 45$, with a dimmer tail comprised mostly of hard-only sources or very local AGN, and the one high spec-z source at a luminosity $\log L_{\rm X} = 46$, brighter even than most sources in the main sample. In general, the main sample lies at higher redshift and dimmer luminosities, while the hard sample populates the brighter and low-z regime. Improved spectral modelling using complex spectral models will help to further confirm the luminosities of these AGN.

2.4.3 Optical spectroscopic properties

For the 172 sources in the eFEDS hard sample which have optical spectra from SDSS, visual inspection can provide some clues as to the nature of the galaxy. Specifically, the presence of broad lines, specifically $H\alpha$, $H\beta$, MgII or CIV, where different lines will be visible based on the source redshift. These broad lines are a clear indicator of AGN activity. These broad lines will be present in type-1 (Seyfert 1) AGN, where the orientation provides a direct view of the central region (see Chapter 1). The visual inspection reveals that 159 objects show clear evidence for at least one broad line, with the majority showing broad $H\beta$ and/or MgII. This demonstrates that, despite being sensitive to obscured, type-2 AGN, the eFEDS Hard sample primarily consists of unobscured type-1 AGN. This is also in agreement with the observed soft X-ray flux distributions, which demonstrated that a majority of the sources are very bright (and thus low obscuration) AGN.

Out of these 159 broad line sources, 25 only show evidence for weak broad Hα with no corresponding broad $H\beta$. These sources are classified as type 1.9 AGN (see also Koss et al. 2022), AGN which typically show evidence for reddening in the optical spectrum and absorption in the X-ray spectra, yet still appear to have broad Hα. The visual inspection also reveals 13 sources with no evidence for any broad emission components. Out of these, 12 show evidence for narrow emission lines and are thus either dominated by the host galaxy in the optical, or type-2 AGN, and one is an absorption line galaxy. These sources are also more likely to be obscured in soft X-rays. The optical spectral properties will be discussed further in Sect. 2.4.5.

To provide a first look at the X-ray properties of sources separated by optical type, Fig. 2.7 shows the distribution of sources showing broad lines (blue solid lines) and lacking broad emission (black dashed lines). For comparison, sources without a spec-z are also shown in grey dash-dot lines. For reference, the hardness ratio derived from an absorbed power law with photon index $\Gamma = 1.8$ and Galactic column density of $N_{\rm H} = 3 \times 10^{20}$ cm^{-2} is also shown. A majority of the broad-line AGN lie near this line, suggesting that many of these sources are relatively unobscured, typical AGN. Some show slightly lower hardness ratios, suggesting a steeper power law component, or the presence of surplus soft emission known as the soft excess (see Chapters 1 and 3), common in type-1 AGN. There is also a significant population with slightly higher HR, suggesting an additional absorption component, and a small population with much larger HR, suggesting significant absorption. These will be modelled in more detail in Chapter 3.

In general, the non-broad-line AGN and photo-z sources show higher hardness ratios, suggesting that these sources may be X-ray absorbed. The non-broad-line AGN are likely to be classified as Seyfert 1.9 or Seyfert 2 galaxies, which are known to show heavier absorption in the X-rays, since under the AGN unification model (see Chapter 1) they should be viewed through the dusty torus (see e.g. Ricci et al. 2017), while photo-z sources are likely to be dimmer, type-2 AGN not targeted in previous spectroscopic follow-up campaigns. The small population with HR of 1 are hard-only AGN which do not show any counts in the soft X-ray band. More detailed spectral modelling is required for these sources to better understand the physical origin and properties of the X-ray absorber, and

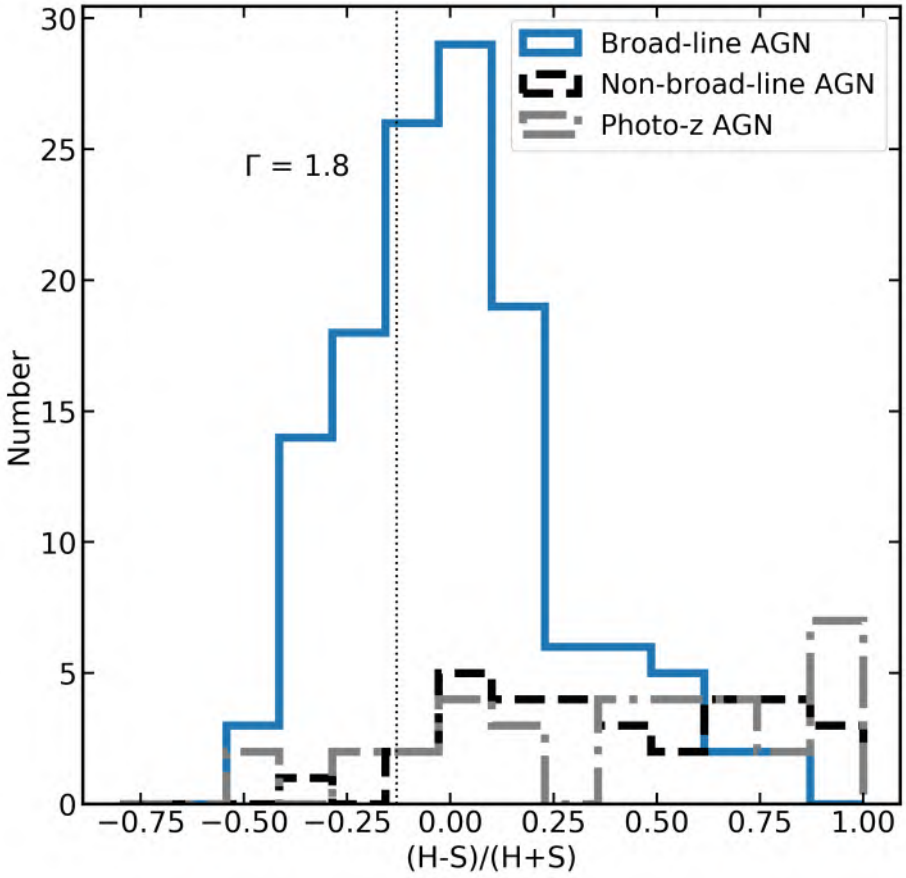

Abbildung 2.7: Distribution of hardness ratios shown for AGN with broad lines in the SDSS spectra (blue solid line), AGN with no broad lines (black dashed line), and source with photo-z (grey dash-dot line). The vertical dotted line indicates the hardness ratio derived from an absorbed power law with photon index $\Gamma = 1.8$ and Galactic column density of $N_{\rm H} = 3 \times 10^{20}$ cm^{-2}.

continued follow-up of the eFEDS hard sample to obtain the final spec-z will help to classify these sources as either type-1 or type-2 AGN.

2.4.4 Blazars

The large area of the eFEDS survey means it is sensitive to rare AGN, including very high luminosity sources such as those seen in blazars. In these sources, the luminosity appears brighter in our line of sight due to Doppler boosting from a relativistic jet, which must be pointing nearly directly towards the observer. A more detailed presentation and analysis of blazars across all eFEDS sources will be given in Collmar et al. (in prep), but here the beamed AGN candidates found in the hard sample are presented. It should be noted that for these sources, the X-ray luminosities as well as the optical luminosities may be boosted or increased due to the presence of the jet, and are thus not representative of the underlying disc-corona emission, so should be treated carefully.

Blazars typically appear abnormally bright in γ-rays and in radio emission. Therefore, they can typically be found by cross-matching with radio and γ surveys. Here, the Fermi Large Area Telescope Fourth Source Catalog (4GL; Abdollahi et al. 2020); the ROMA-BZCAT Multifrequency Catalog of Blazars (Massaro et al. 2015); the 3HSP catalog of extreme blazars (Chang et al. 2019);the Swift-XRT catalog of blazars (Giommi et al. 2019); the Blazar Radio and Optical Survey (BROS; Itoh et al. 2020) and the WISE Blazar-like Radio-Loud Sources catalog (D'Abrusco et al. 2019) are used. A positional cross-match is performed to search to matches within 15 arcseconds of the optical counterpart for the eFEDS hard sources.

In all, 15 matches are found, and are listed in Table 2.4. Two sources are found outside of the 90% exposure area, and were thus removed from the AGN sample presented in this chapter; both are BL Lac objects. Of the 13 within the 90% area, 7 are spectroscopically confirmed flat spectrum radio quasars (FSRQ), and four are of unknown type. A number of these sources were also identified in the visual inspection of the SDSS spectra, as discussed in Sect. 2.4.3.

The analysis confirms that the highest redshift (spec-z 3.2), highest luminosity source discussed above is indeed a beamed AGN; eFEDS J090915.8+035442 was matched with the blazar candidate BROS J0909.2+0354 (Itoh et al. 2020). Examining the optical spectrum, there is clear evidence for optical broad lines; however, the X-ray spectrum is absorbed, with a measured absorption column of $\log N_{\rm H} = 10^{21.4}$ cm^{-2} (see Liu et al. 2022b; Nandra et al. 2024). It is possible that the absorption is intrinsic to the quasar or host galaxy, but the source should be viewed relatively face-on, given the low jet angle required for classification as a blazar and for the beaming required to produced the high observed X-ray luminosity. Instead, the absorption may also originate in the intergalactic medium (see e.g. Arcodia et al. 2018).

Tabelle 2.4: Beamed AGN identified in the hard sample. [a] Outside Area90. [b] bll= BL Lac; fsrq = Flat Spectrum Radio Quasar; fsrqc = fsrq candidate; bcu = Unclassified blazar; bcuc = Unclassified blazar candidate.

Name	SRCID	DET_LIKE_3	Name	z	z_orig	Class[b]
eFEDS J083148.9+042939[a]	65	29.8	4FGL J0831.8+0429	0.174	SDSS	bll
eFEDS J083949.5+010427	112	57.1	4FGL J0839.8+0105	1.124	SDSS	fsrq
eFEDS J085301.2-015049	148	41.7	5BZQ J0853-0150	1.498	Simbad	fsrq
eFEDS J085920.5+004711	19	45.4	4FGL J0859.2+0047	1.47	photz	bll
eFEDS J090111.8+044900	1116	10.8	5BZQ J0901+0448	1.862	SDSS	fsrq
eFEDS J090910.2+012135	25	68.4	4FGL J0909.1+0121	1.024	SDSS	fsrq
eFEDS J090915.8+035442	78	75.1	BROS J0909.2+0354	3.262	SDSS	bcuc
eFEDS J090939.9+020005	354	15.4	4FGL J0909.6+0159	0.15	photz	bll
eFEDS J091408.2-015944	10	61.9	4FGL J0914.1-0202	1.15	photz	bcu
eFEDS J091437.8+024558	55	41.9	4FGL J0914.4+0249	0.427	SDSS	fsrq
eFEDS J092203.5-004442	16185	11.3	WIBRaLS2 J092203.20 -004443.3	0.576	SDSS	fsrqc
eFEDS J092400.9+053345[a]	20	50.4	4FGL J0924.0+0534	1.99	photz	bll
eFEDS J092414.7+030859	23	114.1	5BZQ J0924+0309	0.128	SDSS	fsrq
eFEDS J092507.8+001913	457	19.9	5BZQ J0925+0019	1.721	SDSS	fsrq
eFEDS J093303.3+045235	41	27.9	3HSPJ093303.0+045235	0.378	SDSS	bcu

2.4.5 Black hole masses and Eddington ratios

In sections 2.2.3 and 2.4.3, an overview was given of the optical spectroscopy available for sources in the eFEDS hard sample. Having visually inspected each available SDSS spectrum, it was found that a majority of sources showed evidence for broad lines. Using these sources, bolometric luminosities can be estimated from the continuum luminosities, and black hole masses can be measured from a combination of the full-width-half-maximum of the broad components and the measure luminosities. From there, the accretion rate relative to the Eddington luminosity can also be estimated.

This analysis is performed only for the sources with broad Hβ, Mg II and CIV, since the type-1.9 AGN found only to have broad Hα likely suffer from reddening, making it more difficult to estimate both the continuum luminosity as well as the black hole mass. I also want to select only sources where the Eddington ratio is well constrained from the data available, following the method outlined in Chapter 3. Blazars are also considered separately, since these sources may have beamed emission which complicates the estimation of optical spectral parameters, so these results should be considered with caution and better modelling should be used for more accurate parameter estimation of these sources. Applying these criteria, this selects a sample of 154 AGN suitable for analysis.

The distribution of redshifts, bolometric luminosities, black hole masses and Eddington ratios are shown in blue in Fig. 2.8. Blazars are plotted separately, using filled in blue histograms. The median redshift of the non-beamed AGN sample is at $z = 0.39$, similar but slightly higher than the extragalactic eFEDS sample, but still representative of local AGN. In order to show how these properties evolve over time, distributions from the *Chandra* COSMOS sample presented in Suh et al. (2020) are shown as black dashed lines.

This sample has a higher median redshift of $z = 1.58$, but since the sample is constructed using a soft X-ray selection, both samples are constructed from approximately the same rest-frame energy range. Both samples also have a similar number of sources, making them ideal for understanding the evolution of the population of black hole masses, luminosities, and Eddington ratios over time.

Examining first the distributions of bolometric luminosities shown in Fig. 2.8 (top right), it can be seen that the two distributions are similar, with the COSMOS sources having slight brighter luminosities ($\log(L_{bol}) \sim 43.5$) than eFEDS ($\log(L_{bol}) \sim 43.2$). To test the significance of this difference, a Kolmogorow-Smirnov test (KS-test) can be used, which assesses the significance of the median separations between two distributions. The KS-test yields a p-value of 0.0078, suggesting that the differences are statistically significant. Moving then to the observed black hole mass distributions in Fig. 2.8 (bottom left) however, the two distributions are remarkably similar, with median values of $\log(M_{BH}) \sim 8.63$ in eFEDS and $\log(M_{BH}) \sim 8.55$ in COSMOS. Again performing the KS-test, the result is not statistically significant, with a p-value of 0.08, implying that the black hole mass distributions of X-ray selected AGN has not evolved over this time period.

Moving to the observed Eddington ratios shown in Fig. 2.8 (bottom right), a dramatic difference is seen between the two distributions, with much higher λ_{Edd} values found in the COSMOS AGN as compared to eFEDS. The median Eddington ratio for COSMOS is $\lambda_{Edd} \sim 0.09$, and is nearly a factor of three lower in eFEDS at $\lambda_{Edd} \sim 0.03$. Performing the KS-test yields a p-value of $< 4 \times 10^{-5}$, suggesting that the observed differences are highly statistically significant. This result is clearly not biased by the presence of unidentified blazars, since these cover the full range of measured Eddington ratios. Instead, the observed differences appear to be due to an evolution in Eddington ratio in cosmic time; this result will be discussed further in Section 2.5.2.

2.4.6 Comparison with other samples

The eFEDS field was designed to better characterise the expected performance of eROSITA during the final 8-pass all-sky survey (eRASS:8), which will provide the first all-sky imaging survey in the $\sim 2-8$ keV band. Aside from the not-quite all-sky *XMM* slew survey (Saxton et al. 2008), other previous all-sky surveys in comparable energy ranges have been performed with non-imaging telescopes; for example, HEAO-1 (Rothschild et al. 1979), RXTE (Revnivtsev et al. 2004) and at higher X-ray energies, Swift-BAT (BASS; Oh et al. 2018; Koss et al. 2022).

To better show eFEDS in the context of previous hard X-ray samples, Fig. 2.9 shows the hard X-ray ($2 - 10$ keV) luminosity and redshift of the eFEDS hard sample (blue circles) compared to the BAT AGN sample (Ricci et al. 2017, grey crosses). Flux limits for the HEAO-1 AGN (Piccinotti et al. 1982) are shown with a red dashed line, and are similar to the flux limits of RXTE, and the expected eRASS1 flux limit is shown as a black solid line. The comparison shows that eFEDS probes to higher redshifts and much dimmer luminosities than these previous all-sky surveys, albeit less deep than eFEDS main, thereby opening up a new parameter space to study the hard X-ray (coronal) emission from

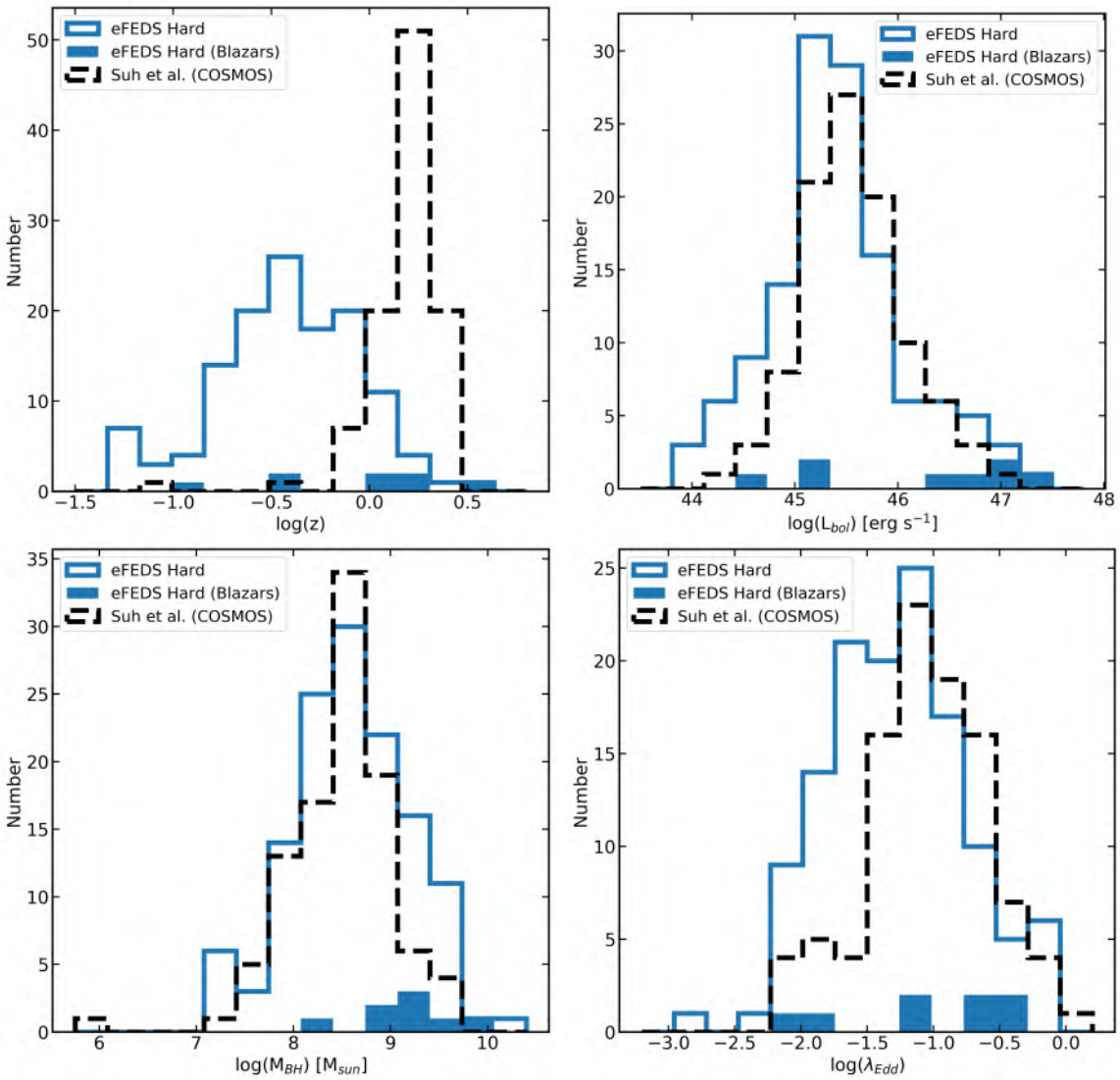

Abbildung 2.8: Top left: distributions of spectroscopic redshifts, shown for eFEDS Hard AGN where Eddington ratios can be constrained (blue solid line), for blazars in the eFEDS hard sample (blue filled histogram), and for AGN in the *Chandra* COSMOS field (Suh et al. 2020, black dashed lines). Top right: same, but shown for bolometric luminosity. Bottom left: same, but shown for black hole mass in units of solar masses. Bottom right: same, but shown for Eddington ratio ($\lambda_{\rm Edd}$).

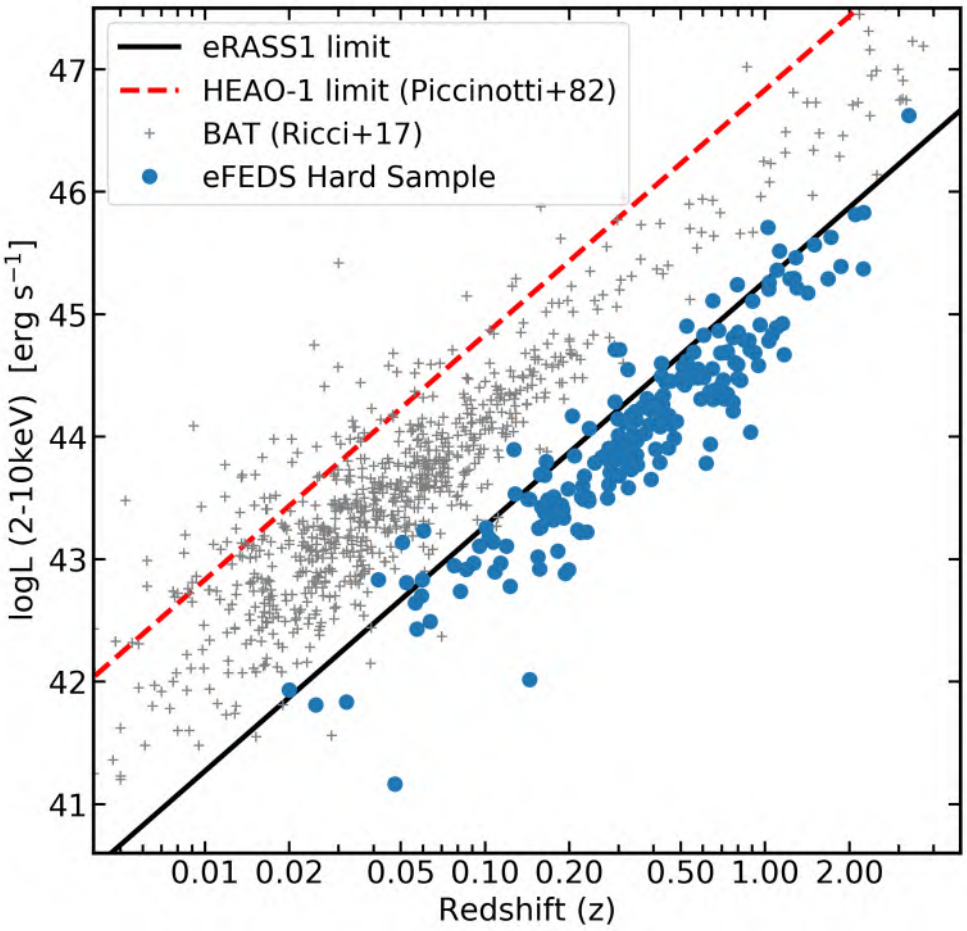

Abbildung 2.9: Luminosity redshift distribution for eFEDS hard AGN (blue circles) and Swift-BAT AGN (grey crosses). Flux limits for eRASS1 (black solid line) and HEAO-1 (red dashed line) are also shown.

AGN. The sample also provides an alternative to Swift-BAT to use as a reference sample of uniformly selected local AGN, albeit with some bias for unobscured AGN. The sample also goes much deeper than either RXTE or HEAO-1 flux limits. From the flux limit, it is clear that eRASS1 (Merloni et al. 2024, Chapter 4) will also extend to lower fluxes than HEAO-1, RXTE or Swift-BAT, and will very likely also cover a sample of local AGN with redshifts of ~ 0.3. This sample however will be much larger than eFEDS, further opening up the opportunity to study nearby, hard X-ray selected AGN.

2.5 Discussion

2.5.1 The hard X-ray sky seen by eROSITA

Analysis of the hard X-ray selected sample in eFEDS has revealed a number of interesting properties and yielded new insights into the expected results from the eRASS1 dataset. Despite the relatively lower effective area in the hard band yielding a sample of only 246 sources, a factor of ~ 100 less than the eFEDS main survey, the higher fraction of sources with secure counterparts and spectroscopic redshifts has enabled a detailed preliminary study of the AGN content of the survey. This chapter discusses the sample selection, optical counterpart identification combining multiple data sets and matching algorithms (NWAY, HamStar, and maximum likelihood), spectroscopic and photometric redshift determination, and source classification. Through these analyses and through comparison we other surveys, the data reveal;

1. The eFEDS hard sample is primarily ($> 90\%$) extragalactic sources, mostly with secure associations, counterparts and spec-z

2. Hard-only sources are mostly found at low X-ray detection likelihoods and are found to have lower quality counterpart associations, suggesting that a higher fraction of these sources are spurious

3. Constructing a sample of 200 AGN reveals primarily a local sample with a median redshift of 0.34, but covers a broad luminosity and redshift range up to $z \sim 3.2$

4. A total of 15 sources including 13 of the AGN sample are revealed to be blazars, including the highest redshift and luminosity source in the sample

5. For the objects with available SDSS spectra, visual inspection reveals that a majority of sources have broad optical lines, and that these sources typically have softer X-ray spectra than sources showing only narrow optical spectral lines

6. Comparing to a similar sample of $z \sim 1$ objects in the COSMOS field reveals a very similar distribution of black hole masses, but significantly lower Eddington ratios (by a factor ~ 3) in the lower redshift eFEDS sample, providing important hints as to the reason for the decrease in accretion luminosity in the local Universe

The analysis presented in this chapter has shown that a majority of the hard X-ray selected sample from eFEDS are AGN, most of which have broad optical lines and unabsorbed X-ray spectra with bright soft X-ray fluxes. This result is not entirely unexpected, as as previous bright, hard X-ray selected samples from other instruments have also shown this behaviour (e.g. La Franca et al. 2002; Della Ceca et al. 2008). Spectral modelling presented in Nandra et al. (2024) demonstrates that the column density and coronal photon index can be constrained with these data. Furthermore, a preliminary look at the hardness ratios shows that many sources show evidence for excess soft X-ray emission or absorption.

This implies that spectral modelling beyond a simple power law should be used to properly measure X-ray fluxes as well as to characterise the emission and to better understand the physical emission mechanisms. This will be explored in Chapter 3, where physical emission and absorption models including thermal emission, non-thermal emission from a secondary warm corona, relativistic blurred reflection, and absorption models including a warm absorber and partial covering absorber (see Chapter 1, Chapter 3) will be tested and compared to better understand the nature of the observed X-ray emission.

2.5.2 AGN evolution and downsizing

One key finding of the eFEDS hard sample is that it contains primarily local AGN, with a peak redshift of ~ 0.34. Being relatively large in size, we also obtain a sufficient number of AGN. As such, it offers an interesting local baseline for comparison with deeper, higher redshift samples, in order to study the evolution of AGN over cosmic time. In particular, I have compared the properties of the eFEDS sample to those of the *XMM-Newton* COSMOS survey, which at a median redshift of ~ 1.3, is selected in a very similar rest-frame energy range as eFEDS, and both samples occupy a similar luminosity range. Both samples contain a similar number of sources, further enabling this comparison.

The analysis reveals a striking similarity in the black hole mass distributions across both surveys. However, we observe a distinct difference in the accretion rate distributions: the typical black hole growth rate relative to the Eddington limit is about three times lower in the eFEDS sample compared to the COSMOS sample. This result is of particular interest considering the global decline in accretion rate density between the two cosmic epochs represented by the COSMOS and eFEDS samples, where the local AGN are found to be accreting at lower rates. Assuming these surveys are reflective of the overall AGN populations, this suggests that the decrease in accretion rate in the local Universe is primarily due to a reduction in the typical fueling rate of black holes, rather than a downsizing effect where active accretion shifts to objects with smaller black holes.

This observed trend aligns with other studies examining the evolution of the specific accretion rate distribution of AGN over time, as indicated by the ratio of X-ray luminosity to host stellar mass (e.g. Aird et al. 2012; Bongiorno et al. 2016; Georgakakis et al. 2017). Given the extensive optical spectroscopy available in our field, we were able to calculate the accretion rate more directly using line width to estimate the black hole mass. The result of a lack of downsizing effect in the overall black hole populations contrasts with some previous studies of optical quasars that do report such an effect (e.g. Kelly & Shen 2013; Schulze et al. 2015). This discrepancy might be attributable to differences in selection bands, but a full understanding necessitates careful modeling of the selection effects (e.g. Trump et al. 2015; Jones et al. 2016).

2.6 Lessons learned from the eFEDS hard sample

This chapter has presented the sample of 246 hard X-ray ($2.3 - 5\,\mathrm{keV}$) sources detected in eFEDS. At the time it was performed, eFEDS was the largest contiguous X-ray survey of its depth performed above 2 keV, now surpassed only by the subsequent eROSITA all-sky surveys (Merloni et al. 2024). The eFEDS data have combined this large area coverage, significant cosmological volume, and a wealth of accompanying data sets including overlapping survey fields and dedicated follow-up campaigns to produce a rich, multiwavelength data set suitable for further analysis. The optical spectral follow-up with SDSS is particularly of relevance here, as thanks to these campaigns, the eFEDS hard sample now has high optical spectroscopic completeness of $\sim 80\%$, and even increasing to 96% when only considering the AGN sample, making it an ideal sample for targeted follow-up of individual sources, for exploitation of the multiwavelength data (e.g. Brusa et al. 2022), or for detailed X-ray spectral analysis (e.g. Chapter 3).

The methods and results of this analysis are of crucial importance for the future study of the eRASS1 and later eRASS hard X-ray selected samples (see Merloni et al. 2024, Chapter 4). The comprehensive analysis presented in this chapter provides a framework for the study of eRASS1 data; beginning first with the construction of a hard X-ray selected sample, then proceeding to the optical counterpart association, then classification using a variety of tools including parallax, spectroscopic redshifts and IR/optical/UV/X-ray colour colour diagnostic tools. The spectroscopic AGN sample is of particular note, since it provides the opportunity to analyse both optical and X-ray spectral properties (see Chapter 3). The sample consists of $\sim 90\%$ AGN, and we might expect similar AGN fractions in the all-sky survey. The analysis has also shown the importance and difficulty in careful treatment of the hard-only sample, as it is shown to be more difficult to find counterparts for and to classify these sources, but they are likely to host obscured AGN which are of interest for analysis and for building a better understanding of obscured AGN in the local Universe.

The eFEDS data also reveal a path forward to later eROSITA all-sky surveys, which will be followed in Chapter 4. Extrapolating from the 200 AGN in eFEDS and applying some approximate corrections for the larger half-sky are and the slightly less deep predicted exposure for the eRASS:8 survey, one could expect 50,000 hard X-ray selected AGN in the 8-pass, all-sky survey, and around 30,000 after ~ 4 all-sky surveys. The analysis here suggests that a majority of sources will be bright, unobscured AGN with large soft X-ray fluxes and low column densities; however, heavily obscured, type-2 AGN are also identified here. The excellent soft X-ray spectral response of eROSITA along with the even richer accompanying data set, including spectroscopic follow-up with SDSS-V and 4MOST as well as counterpart matching with the DESI Legacy Survey data release 10 including unWISE and Gaia magnitudes will provide a rich data set rich for exploitation, as will be seen in the later Chapters of this thesis.

Kapitel 3

Complex absorption and soft excesses in hard X-ray–selected AGN

The eROSITA Final Equatorial Depth Survey (eFEDS) field provided a precursor to the all-sky survey, covering a ~ 140 deg^2 patch of the equatorial sky. From these data, a hard X-ray selected catalog of sources detected in the $2.3-5$ keV band was created and analysed, with methodology and results presented in Chapter 2. Crucially, we recognized that using this hard X-ray selection, the properties of the coronal power law could be constrained using spectral modelling, subsequently enabling the study of additional emission/absorption features. The primary goal of this chapter is thus to test a variety of models for the soft X-ray emission of AGN (thermal emission, non-thermal emission, ionised absorption or neutral partial covering absorption) to help identify the physical origin of the soft X-ray spectral complexity. Differences between these models are examined in the context of this sample to understand the physical properties. Bayesian X-ray Analysis is used to fit a sample of 200 AGN from the eFEDS hard X-ray selected sample with a variety of phenomenological and physically motivated models. Model selection is performed using the Bayes factor to compare the applicability of each model for individual sources as well as for the full sample, and source properties are compared and discussed. I find that 29 sources have evidence for a soft excess at a confidence level $> 97.5\%$, all of which are better modeled by an additional soft power-law, as opposed to thermal blackbody emission. Applying more physically motivated soft excess emission models, I find that 23 sources prefer a warm corona model, while only 6 sources are best fit with relativistic blurred reflection. Of the remainder of the sample, many sources show evidence for complex absorption, with 29 preferring a warm absorber, and 25 a partial covering absorber. Many (18/26) sources which show significant neutral absorption when modelled with an absorbed power law, in fact show evidence that the absorber is ionized, which has important impacts on the understanding of obscured AGN.

This Chapter was published in (Waddell et al. 2024), but I have removed Sections 5.3 and 6.4 from the published manuscript, and all references to the analysis of the optical data. This is due to the adoption of a more conservative black hole mass sample presented in (Nandra et al. 2024), which affected some conclusions presented in (Waddell et al. 2024).

The details of this are presented in sections 3.7 and 3.8.5.

3.1 The X-ray soft excess in AGN

Active Galactic Nuclei (AGN) are powered by supermassive black holes accreting material in their surroundings, giving off radiation across all wavelengths from radio to X-rays. The primary source of X-ray emission in AGN is the corona, which consists of hot and/or relativistic electrons located at some height (few r_g to 100s r_g) above the inner accretion disc (e.g. Parker et al. 2022; Haardt & Maraschi 1991, 1993; Merloni et al. 2000; Fabian et al. 2004). The resulting coronal emission arises from Compton upscattering of lower energy photons, and takes the form of a power law that dominates the X-ray spectra of AGN above energies of $\sim 2\,\mathrm{keV}$. In the soft X-ray band, the spectra of many AGN show evidence for a soft excess (e.g. Pravdo et al. 1981; Arnaud et al. 1985; Singh et al. 1985, see also Chapter 1), a surplus of photons over the primary power-law continuum below $1-2\,\mathrm{keV}$. The origin of this component is highly debated, and a variety of physical mechanisms have been proposed to explain this feature.

Initially, it was proposed that the tail of the disc blackbody from the hottest, innermost regions of the accretion disc may be responsible, as the soft excess shape is well fitted with a blackbody with temperatures of $\sim 0.1\,\mathrm{keV}$ (e.g. Gierliński & Done 2004). The temperature of the disc, however, should scale with black hole mass as $\mathrm{T} \propto \mathrm{M}_{\mathrm{BH}}^{-1/4}$ (Shakura & Sunyaev 1973). Given this relationship, the blackbody photons even from the innermost accretion discs in AGN with typical masses of 10^7 – 10^9 solar masses are highly unlikely to be visible in the X-ray regime. Furthermore, the expected trend between the fitted blackbody temperature and the black hole mass has not been found (e.g. Gierliński & Done 2004; Crummy et al. 2006).

Instead, it has been proposed that the soft excess may be due to a blurred reflection component (e.g. Ross et al. 1999; Ross & Fabian 2005). Some of the photons from the corona will be incident upon the accretion disc. These can be Compton back-scattered from the disc, or excite fluorescence or recombination processes, thus producing a multitude of absorption and emission features. If the accretion disc is highly ionised, these features are concentrated primarily at low energies below 2 keV. Due to the proximity of the inner accretion disc to the black hole, these features are relativistically broadened. This produces a smooth soft excess in excess of a power law at low energies that can have a form similar to a blackbody, as well as a broad iron line and absorption edge in the hard X-ray. Blurred reflection modelling has been used to successfully explain the spectral shape and variability of numerous type-1 AGN (e.g. Fabian et al. 2004; Zoghbi et al. 2008; Wilkins et al. 2017; Gallo et al. 2019; Jiang et al. 2019; Waddell et al. 2019; Boller et al. 2021).

In another interpretation, the soft excess can be produced by a secondary, warm corona (e.g. Done et al. 2012; Petrucci et al. 2018, 2020) which is optically thick ($\tau \sim 10$; Petrucci et al. 2020) and cooler than the primary hot X-ray corona. Blackbody seed photons from the disc undergo Comptonisation in the warm corona. Due to the lower temperature ($\sim 0.1-1\,\mathrm{keV}$) and higher optical depth of the secondary corona, compared to that producing

the hard power law, the tail of the Comptonisation spectrum can be seen in the soft X-rays, which can then produce a soft excess over the power law continuum. In general, the warm corona is interpreted to be a slab above the accretion disc (e.g. Done et al. 2012). The warm corona may only be stable under certain restrictive conditions (e.g. the gas cannot be too hot or too cold; see Ballantyne & Xiang 2020) and may also sometimes produce significant absorption features for temperatures below 10^7 K (García et al. 2019). This model has been shown to fit the spectral shape of numerous type-1 AGN (e.g. Done et al. 2012; Ehler et al. 2018; Tripathi et al. 2019; Petrucci et al. 2018, 2020).

In contrast to these emission mechanisms, it has been proposed that the soft excess is an artefact of improperly modelled absorption features (e.g. Gierliński & Done 2004; Tanaka et al. 2004; Parker et al. 2014; Fang et al. 2015; Gallo et al. 2015; Boller et al. 2021; Parker et al. 2021). In a neutral partial covering absorption scenario, X-rays from the corona pass through an absorber with moderate column density and covering fraction, producing significant absorption in the soft X-ray. This produces a flat spectrum that appears to have some excess emission. This model also produces a deep iron absorption edge at $\sim 7\,\mathrm{keV}$ (Tanaka et al. 2004; Parker et al. 2021), which can explain the hard X-ray curvature observed in some AGN spectra. Rapid variability observed in the spectra of some type-1 AGN has also been attributed to changes in the column density or covering fraction of the absorber.

Instead of an excess of soft photons, many AGN show evidence for complex ionised (warm) absorption which produces features concentrated in the soft X-ray. In particular, partially ionised neon, oxygen and iron absorption lines and edges in the $0.5-2\,\mathrm{keV}$ band are observed in many sources (e.g. George et al. 1998; Kaastra et al. 2000; Kaspi et al. 2000a; Blustin et al. 2004, 2005; Gierliński & Done 2004; McKernan et al. 2007; Laha et al. 2014; Mizumoto et al. 2019). It has been proposed that these warm absorbers can be physically connected to disc winds launched in some AGN systems, or even to large-scale outflows (e.g. Blustin et al. 2004, 2005; Kallman & Dorodnitsyn 2019). These low-velocity winds are well-studied in AGN, especially using high resolution (e.g. using gratings) spectroscopy (e.g. Blustin et al. 2004). Proper characterisation of warm absorption or partial covering absorption is essential for characterising not only the soft excess, but also the hard X-ray continuum, as absorption can create an apparent soft excess at low energies.

Distinguishing between these various models for the soft excess is not a straightforward task, and many previous attempts to do so have shown that all models can sometimes reproduce the observed X-ray spectral shape (e.g. Ehler et al. 2018; Tripathi et al. 2019; Waddell et al. 2019; Chalise et al. 2022). Each model typically has caveats, and it is often difficult to simultaneously explain the spectral shape as well as the short and long term variability. It is also likely that more than one soft excess component exists simultaneously, often complicated by the superposition of multiple absorption components (Boller et al. 2021). Since each of the different soft excess models have very different physical interpretations, consequences for the understanding of X-ray emission from AGN differ. X-ray reverberation mapping (see Uttley et al. 2014, for a review), where a search is performed for time lags between different X-ray energy bands, can also be used to probe the geometry and height of the corona above the black hole (e.g. Zoghbi et al. 2013; Wilkins & Gallo

2015; Kara et al. 2016), and the lags can be interpreted as light travel time between two coronas (Chainakun & Young 2017) or between the hot X-ray corona and the accretion disc (De Marco et al. 2013). Other timing methods including fractional variability analysis (Vaughan et al. 2003) or principal component analysis (Parker et al. 2015) can also be compared with simulations to distinguish between models.

In this chapter, the X-ray spectra of 200 AGN from the hard X-ray–selected sample (Nandra et al. 2024) of the eROSITA Final Equatorial Depth Survey (eFEDS) field (Brunner et al. 2022; Liu et al. 2022b; Salvato et al. 2022) are fit with a variety of phenomenological and physically motivated models to search for the presence of soft excesses and attempt to determine their physical origin. Therefore, the eROSITA bandpass of $0.2-8$ keV is ideal for such measurements, as it provides excellent coverage and resolution as well as high effective area for energies below 1 keV. In Sect. 3.2, the data reduction and techniques used in this chapter are described. In Sect. 3.3, the preliminary models used for spectral fitting are described, and the absorption and soft excess samples are constructed. Sect. 3.4 describes the Sect. 3.5 describes the physical models applied to the sample of sources with soft excess. In Sect. 3.6, the spectral properties identified in this chapter are presented, and in Sect. 3.7, a further discussion of results is given. Finally, conclusions are drawn in Sect. 3.8.

3.2 Data reduction and fitting

3.2.1 The eFEDS hard X-ray–selected sample

The eFEDS field was observed in the eROSITA calibration and performance verification phase and covers ~ 140 deg^2 (Brunner et al. 2022). This equatorial survey overlaps with a plethora of multi-wavelength data, facilitating source characterisation and classification, as well as redshift measurements (Salvato et al. 2022). The eFEDS field is slightly deeper than, but comparable to, the average equatorial exposure of the originally planned eROSITA All Sky Survey (eRASS:8; eight passes of the entire sky), with an average exposure per pixel of 2.2 ks. The eFEDS data are therefore representative of and can be used to predict the all-sky survey performance. More details of this survey and the resulting data products, including the source detection algorithm and data reduction techniques, are presented in Brunner et al. (2022). All eFEDS data have been made public in June 2021 with the Early Data Release (EDR) of the eROSITA German consortium[1].

The main X-ray source catalogue in the eFEDS field is assembled from sources detected in the $0.2-2.3$ keV band (Brunner et al. 2022; Liu et al. 2022b). Proper characterisation of the hard X-ray emission is crucial for understanding the strength and shape of the soft excess component, so sources which only have detections below 2.3 keV are less suitable for this analysis. Therefore, this chapter makes use of the hard X-ray–selected catalogue, based on the detection likelihood in the $2.3-5$ keV band (DET_LIKE_3 > 10; see Nandra et al. 2024). This sample is nearly spectroscopically complete, with 197/246 sources having

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

a spectroscopic redshift from a variety of sources. Since most objects which are significantly detected in the hard X-ray also have significant soft emission, the total of 246 sources in the hard X-ray–selected catalogue largely overlap with the 27,910 sources presented in the main catalogue, with 20 sources being present only in the hard catalogue. These sources are classified according to the process outlined in Salvato et al. (2022). This classification combined the results of three independent methods and considers the multi-wavelength properties of the sources' optical/IR counterparts. The best matches are determined based on a combination of the astrometric and photometric information (see Salvato et al. (2022)). I then apply the selection methods presented in Nandra et al. (2024) and Salvato et al. (2022) wherein sources with secure counterparts, which are likely extragalactic based on their redshift and colour-colour diagnostics, and which have secure photo-z or spectroscopic redshift, are considered to be AGN.

Applying these classifications and cuts, a final sample of 200 hard X-ray–selected AGN is obtained. Of the remaining 46 sources, around one-third are likely galactic, one-third do not have sufficient data quality for counterpart identification, and around one-third do not have sufficiently secure redshifts for spectral modelling. For completeness, all AGN from the sample, including those only detected in the hard band, are included.

Most AGN in this sample are high flux and low redshift (median $z \sim 0.34$) sources (see Chapter 2), and targeting of bright eFEDS sources as part of several SDSS follow-up programmes has resulted in a high level of spectroscopic coverage (Nandra et al. 2024, Merloni et al. in prep.). A total of 156 sources have usable optical spectra obtained as part of SDSS-IV (Gunn et al. 2006; Smee et al. 2013; Dawson et al. 2016; Blanton et al. 2017; Ahumada et al. 2020), or from SDSS-V (Bowen & Vaughan 1973; Gunn et al. 2006; Smee et al. 2013; Kollmeier et al. 2017; Wilson et al. 2019; Almeida et al. 2023). For sources without a spectroscopic redshift, photometric redshifts were computed according to the method outlined in Salvato et al. (2022). To ensure accurate photometric redshifts are obtained, IR, optical, and UV data are used in order to construct and SED, and this is fit to measure the redshift. Independent methods (LePhare; Ilbert et al. (2006) and DNNZ; Nishizawa et al. in press) are compared, and the most reliable redshifts are those which agree between the two methods. Only these sources are considered in spectral fitting (for more details on redshift measurements, see Salvato et al. 2022). Since there exists very high spectral coverage for the AGN sample, for sources which rely on a photo-z, the peak of the probability density function for each redshift is taken, and associated errors are not considered. Redshift and luminosity distributions are presented in section 3.5.1, and also in Nandra et al. (2024).

3.2.2 X-ray spectral analysis

X-ray spectral extraction was performed in the manner described in Liu et al. (2022b) and Nandra et al. (2024), using the eROSITA standard analaysis software system (eSASS) version c001 (Brunner et al. 2022). Spectral fits were also performed in a manner similar to those described in Liu et al. (2022b) and Nandra et al. (2024), with the exception that here the spectra were not rebinned before modelling. This maximises the spectral information

at the expense of computational speed.

Following Simmonds et al. (2018), a parametric spectral model for the eROSITA background (see Freyberg et al. 2020) is learned empirically using all eFEDS background spectra (Brunner et al. 2022; Liu et al. 2022b). First, the parametric model is determined using principle component analysis (PCA) in logarithmic count space. Next, the background spectrum of each source is iteratively fitted by adding principle components in logarithmic space, and further adding Gaussian lines in linear space, as required by the data according to the Akaike information criterion (AIC). In this way, when the addition of further Gaussian lines no longer changes the AIC, the background model is considered satisfactory. During the joint source and background fit, the normalisation of the background shape is a free parameter, while the shape parameters are kept fixed; however when the relative areas of the source and background regions are accounted for, this value is almost always 1. This technique ensures that improper subtraction of the background does not affect the spectral fits, particularly at higher energies, which is highly relevant for this sample.

After the background model has been applied, spectral fitting is performed using Bayesian X-ray analysis (BXA; Buchner et al. 2014; Buchner 2019). BXA combines the X-ray spectral fitting packages and models used in XSPEC (Arnaud 1996) with UltraNest (Buchner 2021), a nested sampling algorithm. By using BXA, the full range of parameter space can be explored to ensure that the best fit is found. Input priors on parameters are used to constrain the values to a reasonable parameter space, and posterior distributions can be examined after fitting to better understand the constraints that can be placed on parameters.

Using BXA for spectral fitting also enables model comparison. The Bayesian evidence (Z) is computed for each spectral fit. This value encompasses both the available parameter space and the fit quality, so it can be used to compare models. This is done using the Bayes factor, K;

$$K = \frac{Z_{M1}}{Z_{M2}} \tag{3.1}$$

where M1 and M2 are the models to be compared. While this value cannot directly be linked to a confidence interval or significance, Bayes factors can still be used to identify the best fitting spectral models. This method will be used in the following chapters in order to robustly compare spectral models, and select sources with soft excesses and warm absorbers on a sound statistical basis.

3.3 Preliminary spectral modelling

Before applying more physically motivated spectral models to each source, I sought a simple characterisation of the spectral shape. This baseline model can then be rejected if the data show statistical evidence in favour of a more complex model including a soft excess or absorption component. The continuum can be modelled using a power law, modified by absorption from the Milky Way (taken from Willingale et al. (2013)) as well as the host

Tabelle 3.1: Abbreviated model names as referenced in this chapter as well as their implementation in XSPEC.

Name	XSPEC implementation
PL	`tbabs` × `ztbabs` × `powerlaw`
PL+PCF	`tbabs` × `ztbabs` × `zpcfabs` × `powerlaw`
PL+WA	`tbabs` × `ztbabs` × cwa18 × `powerlaw`
PL+BB	`tbabs` × `ztbabs` × (`powerlaw` + `constant` × `blackbody`)
PL+PL	`tbabs` × `ztbabs` × (`powerlaw` + `constant` × `powerlaw`)

galaxy, and the results of this fit are discussed in section 3.3.1. The method for performing model comparison is summarised in section 3.3.2, with more details in section 3.4. Next, the complex absorption modelling is presented in section 3.3.3; here, neutral partial covering absorption and warm (ionised) absorption are compared. In section 3.3.4, two different toy models are presented and compared to characterise the shape of the soft excess; a second power law component, and a blackbody component. Section 3.3.5 discusses the model comparison in more detail and present a final sample of sources with complex absorption and soft excesses. A full list of the XSPEC models is given in Table 3.1, and an example source (ID 00011) with all models applied along with the residuals for each of the best fits is shown in Fig. 3.1.

3.3.1 Baseline model: Absorbed power law

Each source is first fit with an absorbed power law model (PL). Two absorption components are added; one component has a redshift of zero and the column density fixed to the value of the Milky Way (taken from Willingale et al. (2013) for each source), while the other has a redshift matching the host galaxy and column density, $\log(\mathrm{NH}_z)$, left free to vary in order to account for absorption in the host galaxy (e.g. originating in galaxy-scale gas, torus, other absorbing material). The column density of the host galaxy absorber is allowed to vary between $\simeq 10^{20}\,\mathrm{cm}^{-2}$ and $\simeq 10^{25}\,\mathrm{cm}^{-2}$, where the lower limit is much smaller than the absorption in the Milky Way and is thus difficult or impossible to measure, and the upper limit represents an entirely obscured spectrum. The index of the power law component is constrained to be between one and three. This will allow for the identification of very hard and soft sources while ensuring that most sources have reasonable values of $\Gamma \sim 1.8 - 2.0$ (e.g. Nandra & Pounds 1994a; Reeves & Turner 2000; Nandra et al. 2007; Waddell & Gallo 2020, 2022; Liu et al. 2022b). For the normalisation, a broad, log-uniform prior ensures that all of the broad range of AGN fluxes found in the eFEDS field can be adequately characterised. The full list of priors are summarised in Table 3.2.

Some results from this preliminary fit are shown in Fig. 3.2. The left-hand panel shows the distribution of median values of the host-galaxy column density, NH_z. There is clearly a large peak at column densities of $\mathrm{NH}_z \simeq 10^{20}\,\mathrm{cm}^{-2}$ (at the limit of the prior so consistent with no additional absorption component beyond Galactic absorption), with a smaller, secondary peak at $\simeq 10^{22.5}\,\mathrm{cm}^{-2}$. Since most sources have low column densities, this suggests that the soft excess component or warm absorption, if present, should be easily detectable

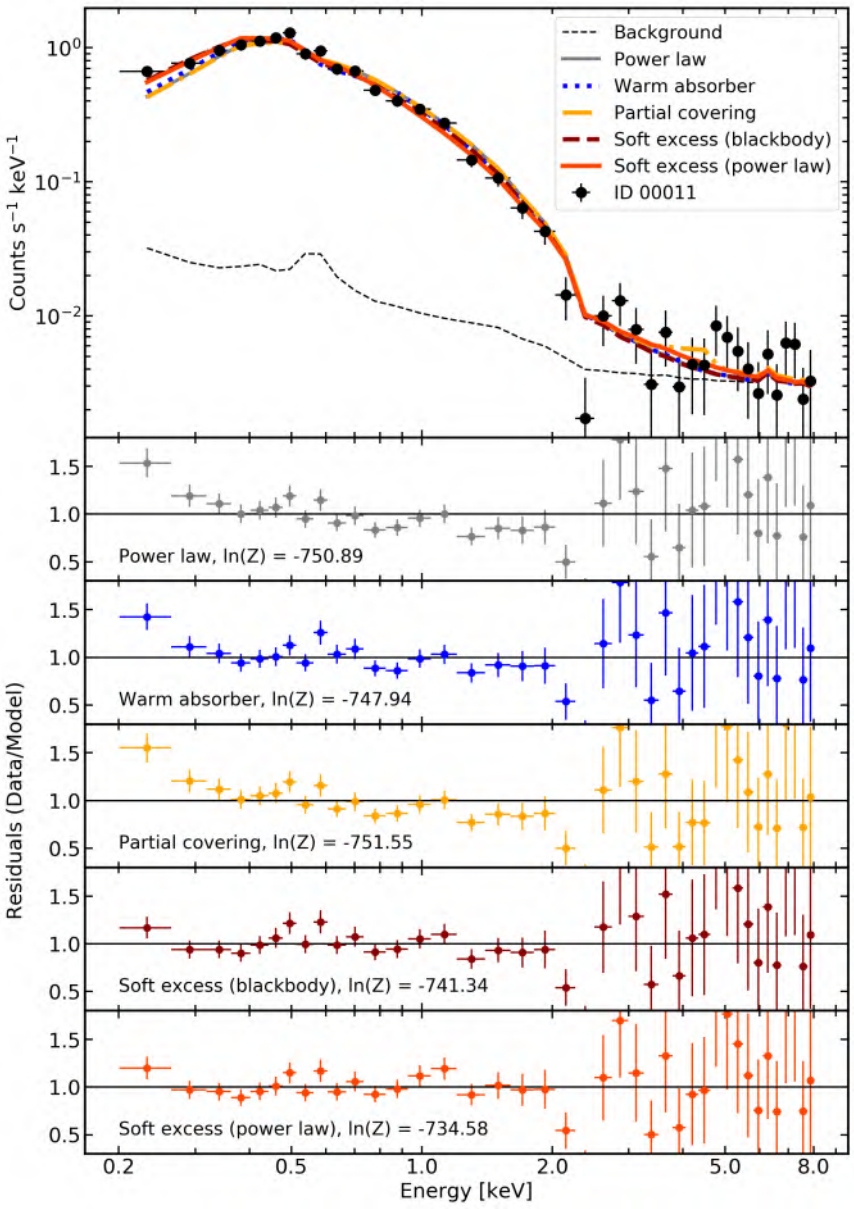

Abbildung 3.1: Comparison of different models and residuals for ID 00011 (z = 0.5121), a source best fit with a double power law soft excess model. The top panel shows the folded spectrum along with each of the best fit models, and the second, third, fourth, fifth and sixth panels show the residuals for the power law (grey), warm absorber (blue), partial covering (orange), blackbody (dark red) and double power law (red) models, respectively. The Bayesian evidence is also given for each model, to ease comparison. The spectrum and residuals are re-binned for clarity. The best fit is a power law soft excess, with a Bayes factor of $K_{pl} \sim 1.22 \times 10^7$ and a significance of >99%. Data have been re-binned for display purposes.

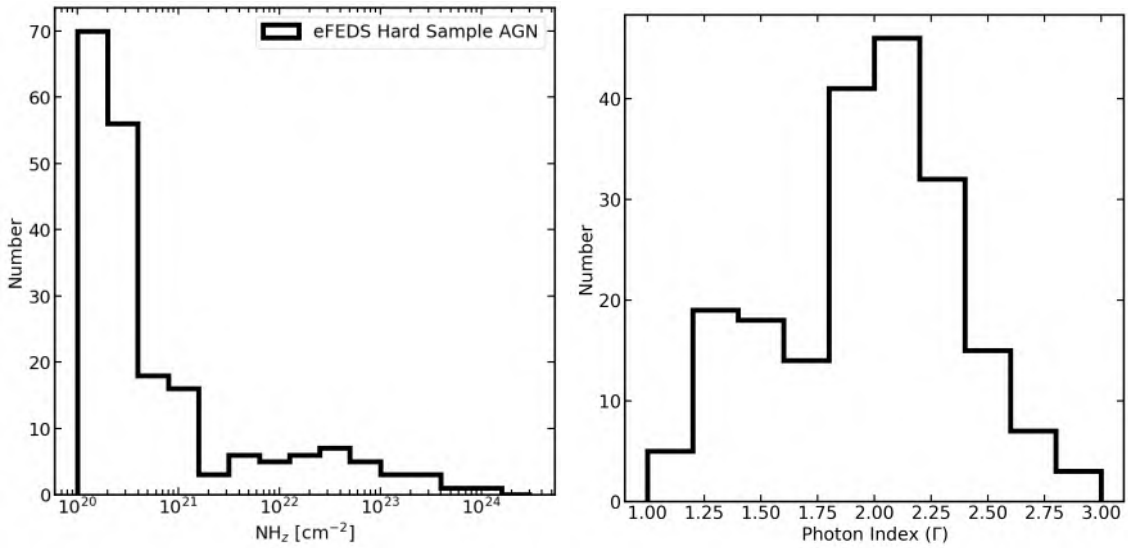

Abbildung 3.2: Distributions of absorbed power law (PL) model parameters. Left: Distribution of host galaxy column density measured using the baseline, absorbed power law model. There is a clear primary peak at $\mathrm{NH}_z \simeq 10^{20}\,\mathrm{cm}^{-2}$, with a secondary peak at $\mathrm{NH}_z \simeq 10^{22.5}\,\mathrm{cm}^{-2}$. Right: as top left, but shown for the photon index. There is a clear peak at $\Gamma \simeq 2.0$, and a second cluster of objects with an unusually flat Γ of ~ 1.4.

Tabelle 3.2: Priors used for simple spectral models. Column (1) gives the shortened name of the model. Columns (2) shows the lower and upper limit of the host galaxy absorption. Columns (3) shows the lower and upper limit of the hard (coronal) power law component. Columns (4) and (5) show the upper and lower limits of the warm or partial covering absorber column density and ionisation component, respectively. Column (6) shows the upper and lower limit placed on the neutral absorption covering fraction. Column (7) shows the upper and lower limit placed on the soft X-ray photon index. Column (8) gives the constraints placed on the cross-normalisation placed on the soft excess component, and column (9) shows the constraints placed on the blackbody temperature. Column (10) gives the constraint placed on the hard power law normalisation.

(1) Model name	(2) log(NH_z) [cm^{-2}]	(3) Γ_h	(4) log(nH_{abs}) [cm^{-2}]	(5) log(ξ_{wa})	(6) Covering fraction	(7) Γ_s	(8) log(c)	(9) kT [keV]	(10) norm
Prior	log-uniform	uniform	uniform	uniform	uniform	uniform	log-uniform	log-uniform	log-uniform
PL	20 – 25	1 – 3	-	-	-	-	-	-	-10 – 1
PL+PCF	20 – 25	1 – 3	20 – 25	-	0 – 1	-	-	-	-10 – 1
PL+WA	20 – 25	1 – 3	20 – 24	-4 – 4	-	-	-	-	-10 – 1
PL+BB	20 – 25	1 – 3	-	-	-	-	-3 – 1	0.04 – 0.4	-10 – 1
PL+PL	20 – 25	1 – 3	-	-	-	2.25 – 10	-3 – 1	-	-10 – 1

Tabelle 3.3: Definitions of Bayes factors used throughout this chapter for model comparison.

(1)	(2)	(3)	(4)	(5)
K_{wa}	K_{pcf}	K_{pl}	K_{nth}	K_{rel}
Z_{WA}/Z_{PL}	Z_{PCF}/Z_{PL}	Z_{PL+PL}/Z_{PL}	Z_{NTH}/Z_{PL}	Z_{REL}/Z_{PL}

in most sources. The right-hand panel shows the distribution of median values of the photon indices (Γ), with most having values of $\Gamma \simeq 2$. This is likely a selection effect, as the value is slightly steeper than found by some samples (e.g. Nandra & Pounds 1994a; Reeves & Turner 2000; Waddell & Gallo 2020, 2022), but in agreement with previous eROSITA modelling presented in Liu et al. (2022b) and Nandra et al. (2024). There are also a number of sources with very steep ($\Gamma > 2.3$) or very flat ($\Gamma < 1.4$) values. With typical error bars of the order of ± 0.2, these values are not in agreement with the expected median of $\Gamma \simeq 2$. These likely indicate the presence of more complex spectral features, such as soft excess emission or complex absorption, motivating further investigation.

3.3.2 Model comparison summary

In the rest of this chapter, the best model for each spectrum is identified with Bayesian model comparison. This relies on the computation of Bayes factors, which examine the Bayesian evidence for two models to determine which is preferred.

Simulations are used to assess the significance of selecting one model over another. These are described in detail in Appendix A. One-thousand simulated spectra are generated using an absorbed power law (PL) model using the average sample properties, and these spectra are then fit with each of the more complex models subsequently defined in this chapter. Purity thresholds can then be defined based on the Bayes factor values which yield a given number of instances where modelling falsely selects the more complex model as the correct one; in this chapter, thresholds of 95%, 97.5% and 99% are considered, and more detail is given on these selections in Appendix A. For these simulated spectra, false detections are defined when a model both has the lowest Bayesian evidence of all models, and the Bayes factor exceeds the threshold. In this way, for a real source to be classified as having a warm absorber, soft excess, or partial covering absorber, this model must have the lowest Bayesian evidence, and the Bayes factor must exceed the threshold. For this reason, there is no overlap between the true soft excess, partial covering, or warm absorber samples.

3.3.3 Absorption modelling

The first model used to model a complex absorption component in this chapter is a neutral partial covering scenario (PL+PCF), where emission from the corona passes through an absorber before reaching the observer (e.g. Tanaka et al. 2004). This absorbs hard X-ray photons while allowing leakage in the soft X-ray, which flattens the observed power law slope, gives the appearance of a soft excess, and can produce a deep edge at 7 keV

depending on the column density. However, often multiple absorbing zones with a variety of ionisation states, column densities and covering fractions are required to fit the observed spectral shape.

In this chapter, one neutral partial covering absorber is applied to each source. The full XSPEC implementation is given in Table 3.1. Two more free parameters are present; the absorption column density is allowed to vary between 10^{20} and $10^{25}\,\mathrm{cm}^{-2}$, and the fraction of emission which passes through the absorber (the covering fraction) is allowed to vary between zero (no absorption) and one (full covering). The redshift of the absorber is set to that of the host galaxy such that absorption in the vicinity of the corona is modelled. All other parameters and priors are the same as the baseline PL model. While the use of a single, neutral absorber to explain the observed curvature in the spectral shape may be an over-simplification of a true physical absorber, this simple implementation still allows for a preliminary model check, and the Bayesian evidence can be compared with other physical scenarios.

In the warm absorber model (PL+WA), rather than passing through a neutral absorber, the X-ray emission passes through an ionised medium, which produces absorption features in the soft X-ray spectrum due to partially ionised materials including Neon, Oxygen and Iron (e.g. George et al. 1998; Kaastra et al. 2000; Kaspi et al. 2000a; Blustin et al. 2004, 2005; Gierliński & Done 2004; McKernan et al. 2007; Laha et al. 2014; Mizumoto et al. 2019). These warm absorber features have been physically linked to low velocity (e.g. 100s-1000s km s^{-1}) outflows or disc winds which intercept the line of sight (e.g. Kallman & Dorodnitsyn 2019). To model the warm absorber, an XSPEC-compatible table model (cwa18.fits; cwa18) was generated using XSTAR. The construction of this model is described in Nandra et al. (2007). This model also has two more free parameters than the baseline PL model; the column density and the ionisation of the absorber. The ionisation of the absorber is allowed to vary broadly between 10^{-4} and 10^{4} ergs cm s^{-1} to account for a broad range of wind ionisation states, and the column density is between 10^{20} and $10^{24}\,\mathrm{cm}^{-2}$. The full list of priors for the partial covering absorber and warm absorber models is summarised in table 3.2.

The significance of each of the absorption components is assessed using simulations, described in detail in Appendix A. Using these simulations and the Bayes factor as given in table 3.3, it is found that 29/200 sources (14.5%) have evidence for a warm absorber and 25 sources (12.5%) have evidence for partial covering absorbers, both at the 97.5% confidence level. For completeness, most figures/ch3/ will show all three determined purity levels (95%, 97.5%, and 99% significance) for comparison.

Fig. 3.3 shows the distribution of the warm absorber parameters; column density and ionisation (ξ). All three purity levels are shown; sources which have purity at the 95% level ($\mathrm{K}_{wa} > 0.815$) are shown as translucent blue circles, sources at the 97.5% level ($\mathrm{K}_{wa} > 1.126$) are shown as blue rings, and sources with the 99% purity ($\mathrm{K}_{wa} > 2.040$) are shown as dark blue circles. Sources which do not show significant evidence for warm absorbers, and are indicated with black crosses. Sources lacking evidence for warm absorption have lower column densities, while sources with warm absorbers have higher column densities of $\geq 10^{21}\,\mathrm{cm}^{-2}$, with typical values around $10^{22} - 10^{23}\,\mathrm{cm}^{-2}$. In general, the column densities

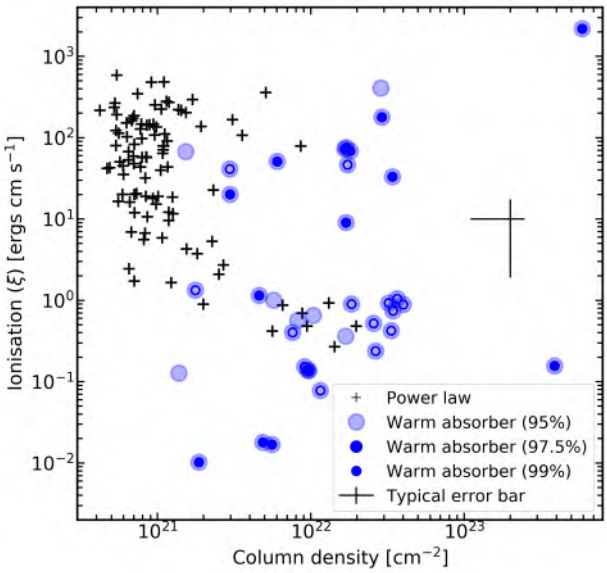

Abbildung 3.3: Warm absorption parameters for all AGN in the sample. Sources with warm absorption components of various purity levels (95%, 97.5% and 99%) are indicated with blue circles (translucent, unfilled and opaque, respectively). Typical error bars are indicated with a black cross.

are not well constrained and can extend to lower values, as there is significant degeneracy between the column density and the ionisation of the absorber as well as between the host-galaxy absorbing column density and the warm absorber column density.

Interestingly, the significant warm absorbers show a wide range of column densities and ionisations, suggesting some diversity in absorbers across different AGN. Typically, warm absorbers studied in the X-ray have been found to have ionisations of the order of $\xi \sim 10 - 1000$ ergs cm s^{-1} and column densities of the order of 10^{20}-10^{23} cm^{-2} (e.g. Blustin et al. 2004, 2005; McKernan et al. 2007; Tombesi et al. 2010; Mizumoto et al. 2019). The results from this chapter are broadly in agreement with this; however, several low ionisation absorbers with $\xi \sim 0.01 - 1$ ergs cm s^{-1} are also found, including some with very high significance (e.g. a very large improvement of the Bayesian evidence compared to a power law). The error bars on the ionisation these sources are large, and the marginal posterior probability distributions can be complex (see Sect.s 3.5.2 and 3.6.4 for more details). These results should be interpreted with caution due to the known degeneracy between the ionisation and the column density of the absorber, however they appear significantly different than sources best fit with neutral absorbers. It is also interesting to note that the sources with higher ionisations ($\sim 10^2$ ergs cm s^{-1}) have relatively low redshifts ($z < 0.5$), while the sources with lower ionisations occupy a much broader range of redshifts with many having $0.5 < z < 1$. These sources will be discussed in more detail in later chapters.

An example of a source (ID 00016) best fit with a warm absorber model is shown in Fig. 3.4. The background (black dashed line), the warm absorber model (blue) and an absorbed power law model (grey) are shown over-plotted with the folded spectrum. The

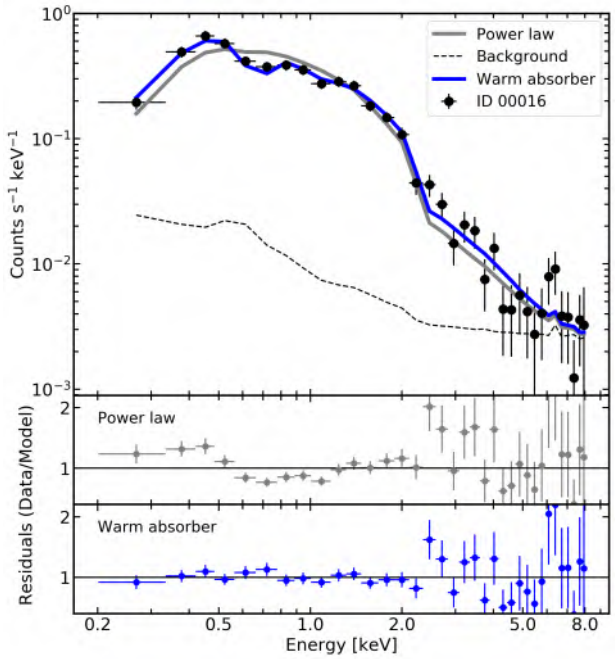

Abbildung 3.4: ID 00016 (z = 0.2907), a source best fit with a warm absorption model. The spectrum (re-binned for display) is shown in black, the background model is shown as a black dashed line, the power law model is shown as a grey line, and the warm absorber model is shown in blue. The bottom two panels show the residuals for the power law, and a warm absorber, respectively. The source has relatively high signal-to-noise, and has a warm absorber column density of $\sim 10^{22}\,\mathrm{cm}^{-2}$ and an ionisation of $\xi \sim 10^2$ ergs cm s^{-1} . Data have been re-binned for display purposes.

warm absorber model clearly provides a better fit than the simple absorbed power law (PL) model, in particular to the softest X-ray energies, as well as in the $2-5\,\mathrm{keV}$ band. In this case, the partial covering absorber model fails to reproduce the absorption features.

Fig. 3.5 is the corresponding corner plot for the warm absorber model fit to source ID 00016, with variable names provided in the figure caption. For this source, most parameters are very well independently constrained, though some degeneracy exists between the column density and the ionisation of the warm absorber. Here the column density of the warm absorber and host galaxy absorber are independently constrained, and the host galaxy column density is consistent with the minimum value.

Finally, Fig. 3.6 shows the distribution of partial covering column densities and covering fractions obtained from the PL+PCF model, with the Bayes factor as given in Table 3.3. As in Fig. 3.3, sources which have purity at the 95% level ($\mathrm{K}_{pcf} > 1.392$) are shown as translucent orange pentagons, sources at the 97.5% level ($\mathrm{K}_{pcf} > 1.555$) are shown as orange unfilled pentagons, and sources with the 99% purity ($\mathrm{K}_{pcf} > 2.646$) are shown as darker orange pentagons. It is clear that sources which are best fit with the partial covering model have higher covering fractions than those which do not; the median is ~ 0.4 for sources with

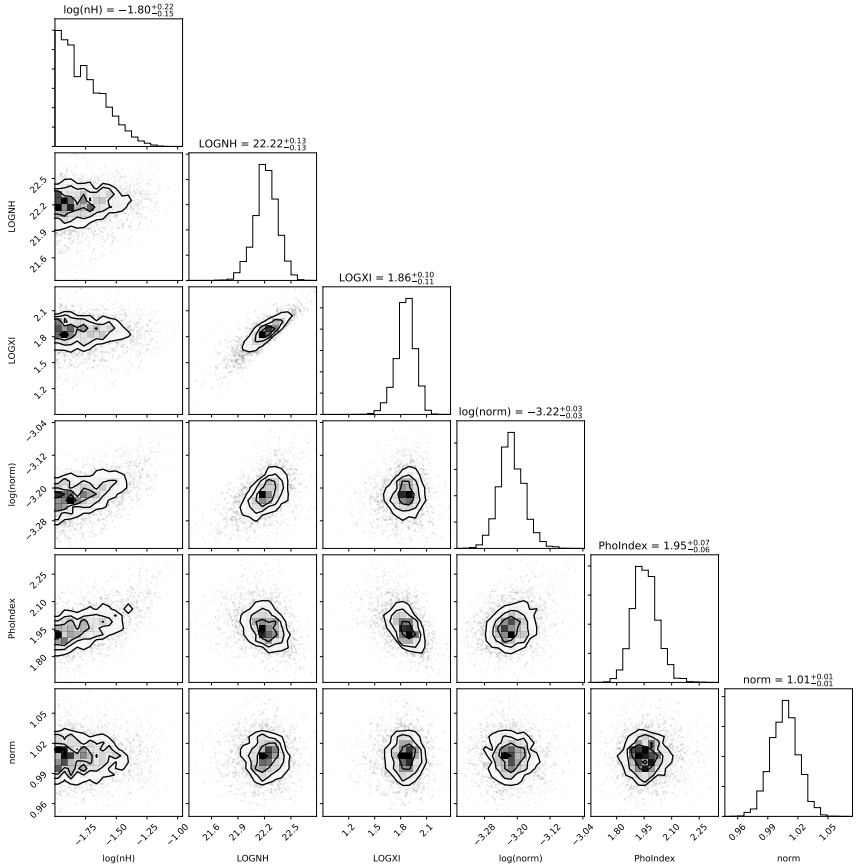

Abbildung 3.5: Corner plot for source ID 00016, best fit with a warm absorption model. The diagonal panels show the marginal posterior probability distribution for each parameter, while the other panels show the conditional probability distribution functions for each pair of parameters. Here, log(nH) is the host galaxy absorber column density (in units of $\times 10^{22}\,\mathrm{cm}^{-2}$), LOGNH is the column density of the warm absorber (cm^{-2}), LOGXI is the ionisation of the warm absorber (ergs cm s^{-1}), log(norm) is the power law normalisation, PhoIndex is the photon index of the power law, and norm is the relative renormalisation of the background model with respect to the source model, which is consistent with unity.

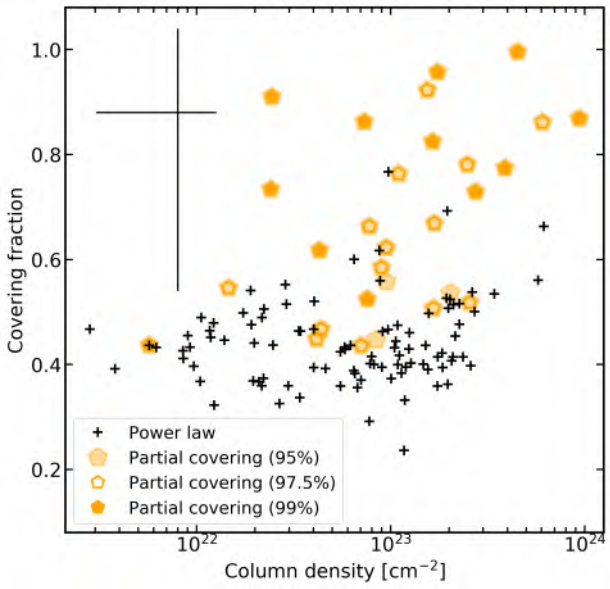

Abbildung 3.6: Partial covering parameters for all AGN in the sample. Sources with partial covering components of various purity levels (95%, 97.5% and 99%) are indicated with orange pentagons (translucent, unfilled and opaque, respectively). Typical error bars are indicated with a black cross.

no evidence for partial covering absorption, and ~ 0.7 for those with evidence for partial covering components. Typical column densities are $\sim 10^{23}\,\mathrm{cm}^{-2}$, with a few sources having higher column densities near $10^{24}\,\mathrm{cm}^{-2}$. This is similar to the warm absorbers. Many of the highest significance sources (filled orange pentagons) also have very steep photon indices (e.g. those in the top right-hand corner of Fig. 3.6). The high covering fraction creates deep absorption edges around 7 keV for sources with sufficiently high column densities, which are difficult to see in the data due to high background at high energies.

An example of a source (ID 00030) best fit with the neutral partial covering model is shown in Fig. 3.7. As for the warm absorber spectrum, the background (black dashed line), the partial covering model (orange) and an absorbed power law model (grey) are shown over-plotted with the folded spectrum. This source is also well fit with a power law soft excess model, but evidence comparison reveals that the partial covering absorption model provides the best fit, highlighting the importance of considering a variety of models to explain the soft spectrum.

The corner plot for the partial covering model of ID 00030 is shown in Fig. 3.8. Some parameters are less well constrained for this model, and the photon index is found to be extremely high compared to expected values of $\sim 1.9 - 2.0$. In this source, there are degeneracies between the partial covering fraction and photon index, as well as between the partial covering fraction and the normalisation on the coronal power law component. In some sources, there are also degeneracies between the partial covering fraction and column density.

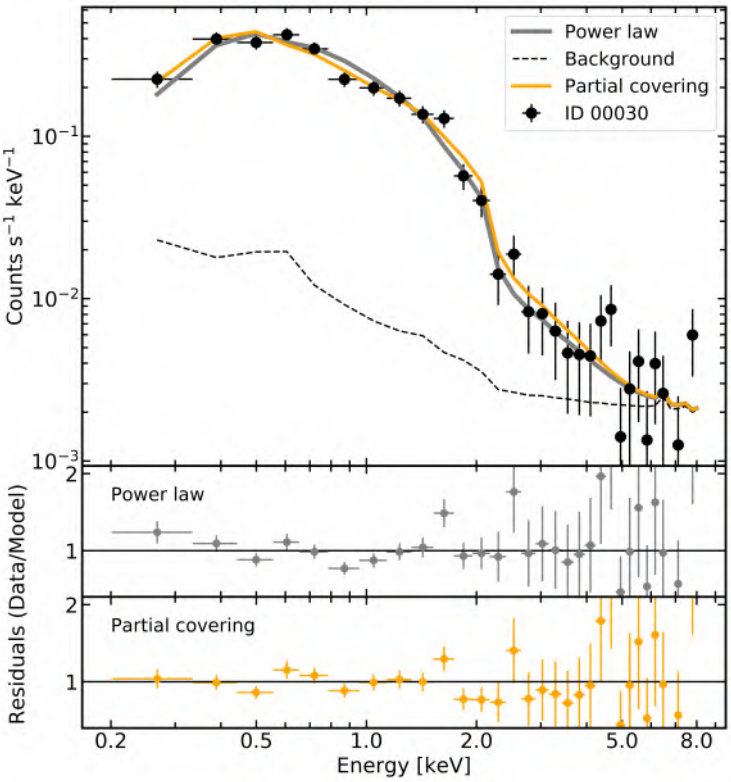

Abbildung 3.7: ID 00030 (z = 0.4263), a source best fit with a partial covering absorption model. The spectrum (re-binned for display) is shown in black, the background model is shown as a black dashed line, the power law model is shown as a grey line, and the neutral partial covering absorption model is shown in orange. The bottom two panels show the residuals for the power law, and a partial covering absorber, respectively. The source has a moderate covering fraction of ~ 0.6 and a moderate column density of $4 \times 10^{22}\,\mathrm{cm}^{-2}$. Data have been re-binned for display purposes.

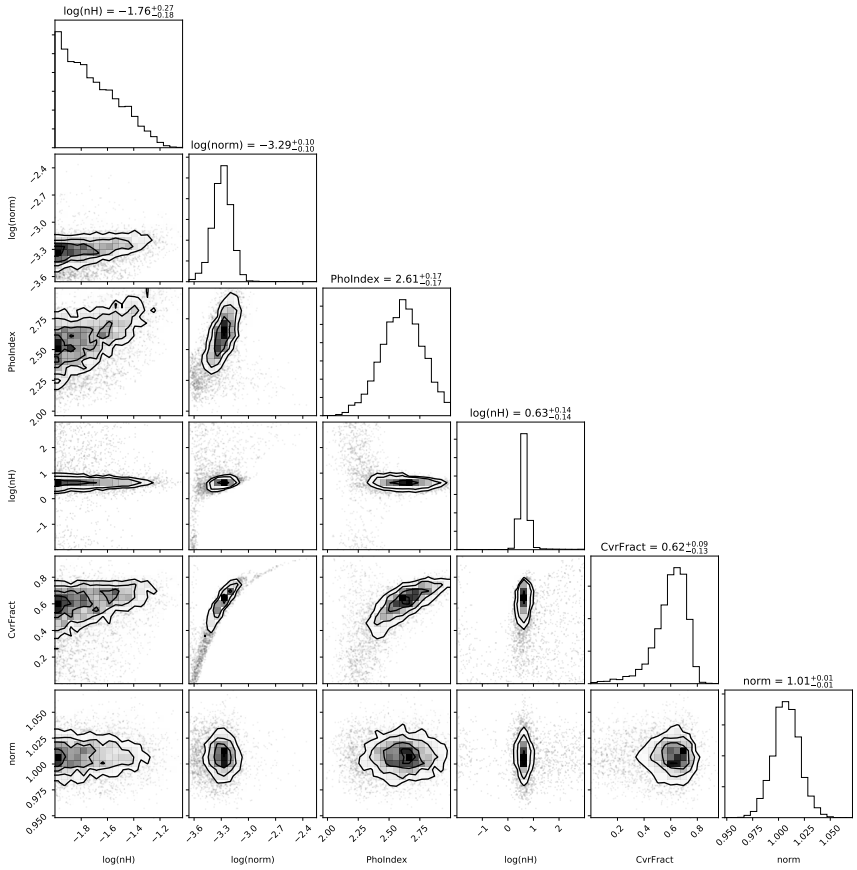

Abbildung 3.8: Corner plot for source ID 00030, best fit with a partial covering absorber. Here, the first instance of log(nH) is the host galaxy absorber column density (in units of $\times 10^{22}\,\mathrm{cm}^{-2}$), log(norm) is the power law normalisation, PhoIndex is the photon index of the power law, the second instance of log(nH) is the partial covering absorber column density ($\times 10^{22}\,\mathrm{cm}^{-2}$), CvrFrac is the covering fraction of the absorber, and norm is the relative renormalisation of the background model with respect to the source model, which is consistent with unity.

3.3.4 Soft excess modelling

In order to account for an intrinsic soft excess component, two separate spectral models are used. First, a `blackbody` component is used (PL+BB). The normalisation component of the blackbody is linked to that or the coronal power law, with a constant factor applied to set the relative spectral flux density of the power law and blackbody components at 1 keV. For the second model, the blackbody component is replaced by a soft power law component (PL+PL), where the relative normalisations of the soft and hard power law is again fit using a constant factor. Motivated by the results of Liu et al. (2022b) from fitting the eFEDS main sample, a wide range of values are adopted for the priors of the blackbody temperature kT and the soft power law index Γ_s in order to characterise a wide variety of soft excess shapes. Priors are listed in Table 3.2.

The significance of the soft excess is evaluated for each source by computing the Bayes factor (equation 3.1), as given in Table 3.3. The resulting Bayes factors are then compared between models for each source to assess which model is better able to characterise the shape of the soft excess. The results are shown in Fig. 3.9, shown for all sources, including those which show evidence for obscuration. Sources which do not show evidence for a soft excess are shown in grey. Most sources are better fit with the PL+PL model (e.g. lie below the line), and all but one of the few sources which are better fit with the PL+BB model have very low K_{pl} and K_{bb} values and therefore likely do not have a soft excess. The sole exception of this is eFEDS ID 00016, which appears to have a strong soft excess with both models, but in fact is better fit with a warm absorber. Therefore, it can be concluded that the PL+PL model is a better representation of the soft excess, therefore the PL+BB model will be discarded for the remainder of this chapter. The double power law model will be used to select sources with significant soft excesses.

As with the complex absorber modelling, the significance of each of the soft excess components is assessed using simulations, described in detail in Appendix A. The distributions of parameters for the soft excess models are shown in Fig. 3.10. All three purity levels are shown; sources which have purity at the 95% level ($K_{pl} > 1.392$) are shown as translucent red squares, sources at the 97.5% level ($K_{pl} > 2.586$) are shown as unfilled red squares, and sources with the 99% purity ($K_{pl} > 8.613$) are shown as dark red squares. Choosing the 97.5% purity level, this leaves 29/200 sources with soft excesses, or $\sim 14.5\%$ of the full sample, the same number as found for a warm absorption model and slightly more than found using a partial covering model. However, significantly more sources are found to have soft excesses at the 95% and 99% confidence levels. Sources with soft excess display a surprisingly large variation in primary photon index, which may suggest that some sources have additional absorption components not considered in this model, or that the two power law model is too simplistic to characterise the spectral shape and complexity for some sources. There also appears to be two clusters of soft photon indices for sources with soft excesses; a cluster around $\Gamma_s = 4.5$, and another around $\Gamma_s = 6.5$. Both of these values are too steep to be produced in a corona with reasonable opacity and temperature (e.g. Petrucci et al. 2018), and depending on the assumed temperature, these may exceed the steepness of the exponential cut-off. Rather, they likely highlight some diversity in the

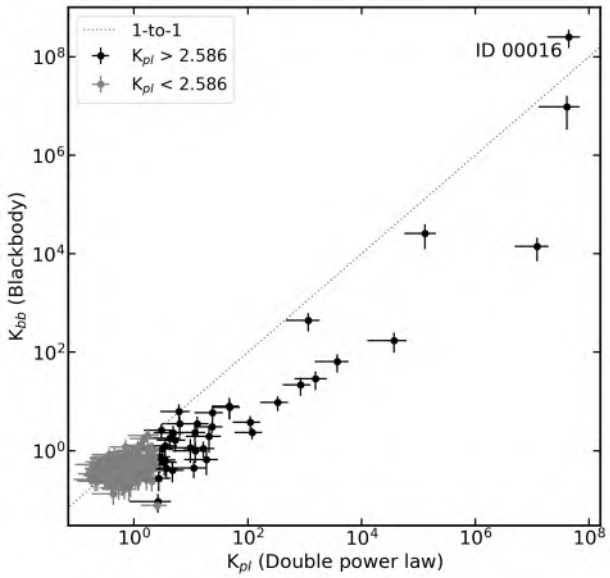

Abbildung 3.9: Comparison of Bayes factors between soft excess models. The Bayes factor for the double power law is shown on the horizontal axis, while the Bayes factor for the blackbody model is shown on the vertical axis. The dashed grey line shows the one-to-one relation, with sources lying below the line being better fit with the PL+PL model, and sources above the line being better fit with the PL+BB model. Sources which show evidence for a soft excess (K_{pl}>2.586, corresponding to a 97.5% significance) are shown in black, and sources which do not show evidence for a soft excess are shown in grey.

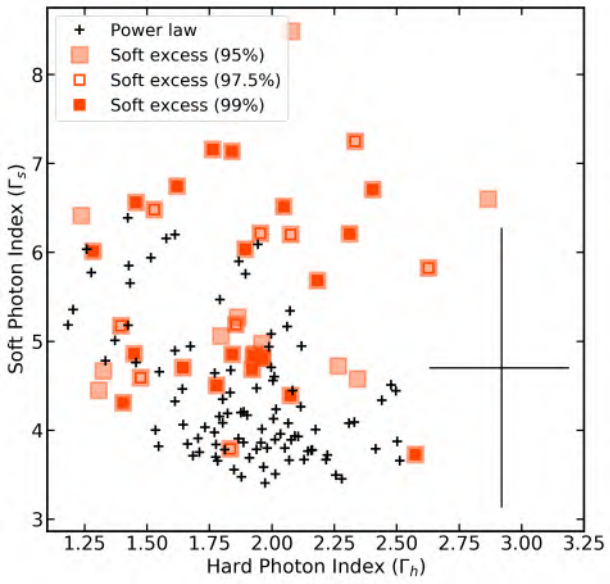

Abbildung 3.10: Soft excess (double power law) parameters for all AGN in the sample. Sources with soft excess components of various purity levels (95%, 97.5% and 99%) are indicated with red squares (translucent, unfilled and opaque, respectively). Typical error bars are indicated with a black cross.

shape of observed soft excesses, or may be competing with the host galaxy absorption to attempt to match the observed spectral shape.

An example of a source (ID 00039) best fit with the double power law model is shown in Fig. 3.11. As for the previous models, the background (black dashed line), the double power law soft excess model (red) and an absorbed power law model (grey) are shown over-plotted with the folded spectrum. The fit improvement by adding the second power law is visually apparent throughout the spectrum, and in particular in the hard band - the single power law model tends to try to approximate the softer end of the spectrum where the instrument is more sensitive, while the second power law provides a better fit across all energies. The corresponding corner plot for this source is shown in Fig. 3.12, and the component labels are explained in the caption. Here there are some additional degeneracies between parameters, including between the photon indices and normalisations, as well as between the photon indices and the covering fraction. While some parameters are less well independently constrained, the double power law model is still highly informative in characterising the shape of the soft excess.

3.4 Simulations to assess Bayes factor thresholds

In order to determine which values of the Bayes factors for each model correspond to a significant improvement to the fit, simulations are performed. Here, 1000 spectra are generated using the fakeit command in XSPEC. Spectral simulations are based on an

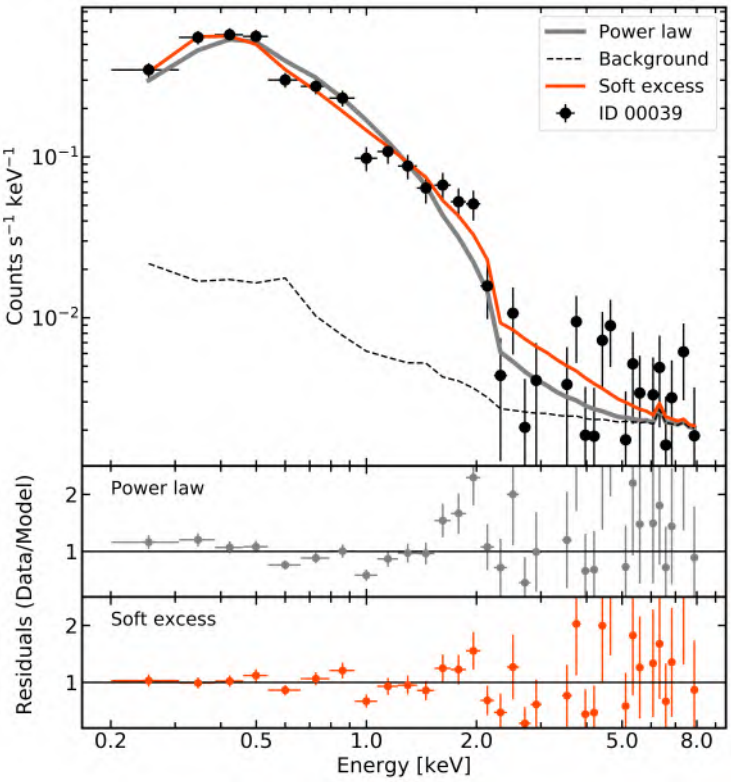

Abbildung 3.11: ID 00039 (z = 0.3893), a source best fit with a double power law soft excess model. The spectrum (re-binned for display) is shown in black, the background model is shown as a black dashed line, the power law model is shown as a grey line, and the double power law model is shown in red. The bottom two panels show the residuals for the power law, and a soft excess, respectively. Data have been re-binned for display purposes.

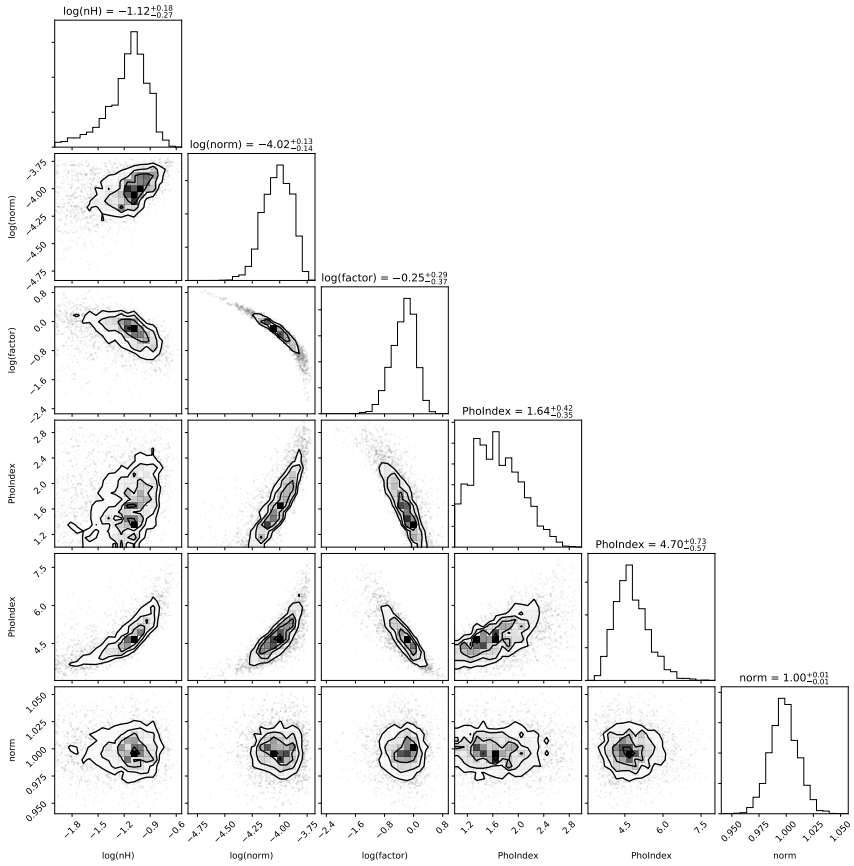

Abbildung 3.12: Corner plot for source ID 00039, best fit with a double power law soft excess. Here, the first instance of log(nH) is the host galaxy absorber column density (in units of $\times 10^{22}\,\mathrm{cm}^{-2}$), log(norm) is the power law normalisation, log(factor) is the relative normalisation of the soft power law with respect to the hard power law, the first instance of PhoIndex is the photon index of the hard power law, the second instance of PhoIndex is the photon index of the soft power law, and norm is the relative renormalisation of the background model with respect to the source model, which is consistent with unity.

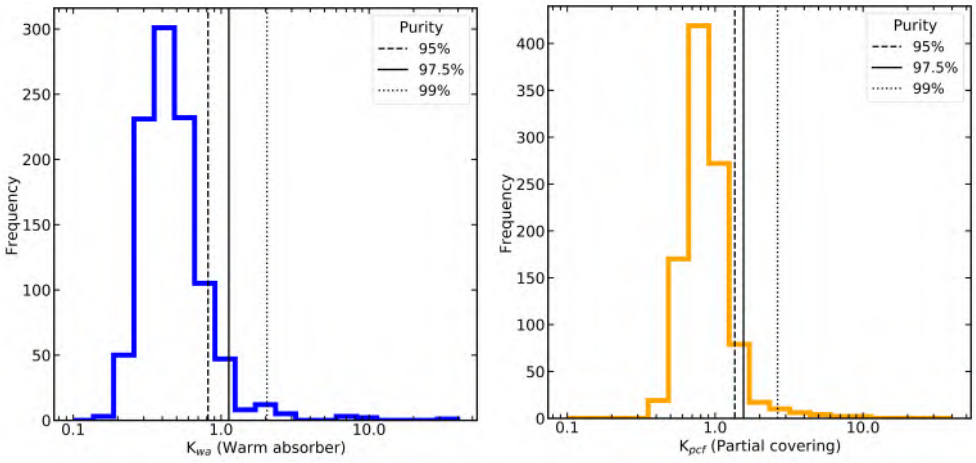

Abbildung 3.13: Distributions of Bayes factors obtained from simulations. Left: Computed warm absorber Bayes factors (K_{wa}). Right: computed partial covering Bayes factors (K_{pcf}). Confidence levels of 95%, 97.5% and 99% are indicated using vertical lines of different linestyles.

absorbed power law model (`tbabs` × `ztbabs` × `powerlaw`) using the average spectral properties from the absorbed power law (PL) model, with a photon index of $\Gamma \sim 2$ and a host galaxy column density of $\sim 2 \times 10^{20}\,\mathrm{cm}^{-2}$. In this way, the simulated spectra are representative of the average eFEDS spectrum and can be used to calibrate the Bayes factor threshholds for the entire sample.

First, to determine the significance of the absorber models, each simulated spectrum is then fit with the baseline (PL), warm absorber (PL+WA), and partial covering (PL+PCF) models, and the Bayes factor is computed for each model using equation 3.1. Formulae for the Bayes factor for all models are given in Table 3.3. The resulting histograms are shown in Fig. 3.13, with results for warm absorption modelling shown in the left-hand panel and results for neutral partial covering shown on the right. Purity thresholds (the confidence per object that the absorption component is real) levels of 95%, 97.5% and 99% are indicated using vertical lines with dashed, solid and dotted linestyles, respectively. These Bayes factor purity thresholds are selected such that cases where if the model both has the lowest Bayesian evidence of all models, and the Bayes factor exceeds the threshold, it is considered a false detection.

By applying these purity thresholds to the eFEDS hard X-ray sample, the number of expected false detections of an absorption component at each confidence level can be predicted. These results are summarised in Table 3.4. The percentile gives the percentage of simulated sources with Bayes factors below the quoted value, and the false detection rate gives the corresponding number of false detections. The number of sources is the number of real sources with Bayes factors above the stated value and where the given model provides the best fit, and the number of true detections is calculated by subtracting the expected number of false detections in a sample of 200 sources. Since evidence comparison is used

Tabelle 3.4: Summary of detection thresholds for complex absorbers. Column (1) gives the percentile of sources, and column (2) gives the corresponding Bayes factor purity threshhold. Column (3) gives the resulting assumed number of false detections, obtained from multiplying the purity by the total number of sources (200). Column (4) gives the number of sources in the sample with K_{wa} values greater than the value given in column (2) and where the model provides the best fit to the data. Column (5) gives the assumed number of true detections in the eFEDS hard X-ray–selected sample of 200 sources.

(1) Percentile	(2) K	(3) Expected false detections	(4) Expected Sources	(5) True detections
Warm absorption				
95%	0.815	10	36	26
97.5%	1.126	5	29	24
99%	2.040	2	15	13
Partial covering				
95%	1.360	10	28	18
97.5%	1.555	5	25	20
99%	2.646	2	12	10

to determine which of the complex models provides the best fit to the data, scenarios where a source appears in multiple sub-samples are excluded. The 97.5% purity thresholds are taken as the samples of sources with warm absorbers or partial covering components respectively.

The Bayes factors obtained from fitting the eFEDS spectra with the two complex absorption models are compared in Fig. 3.14, with error bars are shown for each source. One source, ID 00011, is excluded as the K_{wa} and K_{pcf} values are many orders of magnitude larger than the rest of the sample. There is clear evidence both for sources which are very well fit with neutral partial covering and with ionised warm absorption, confirming that both models should be included in the analysis in order to properly characterise the shape of the observed X-ray spectra.

From the computed PL+PL Bayes factors, to determine which sources have significant soft excesses, the same method employed to determine the significance of warm absorbers and partial covering components is used. The same 1000 simulated spectra are used, but are now fit with the PL+PL model, and Bayes factors are again computed. The resulting histogram of K_{pl} values is shown in Fig. 3.15. Confidence levels of 95%, 97.5% and 99% are indicated using vertical lines with solid, dashed, dotted and dash-dot linestyles, respectively. In general, a higher threshold is required to distinguish a soft excess than to distinguish complex absorption. The false detection rates are again calculated and are summarised in Table 3.5.

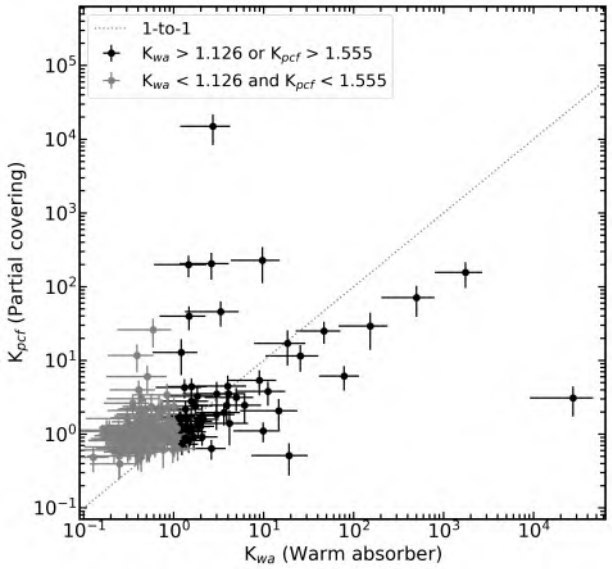

Abbildung 3.14: Comparison of Bayes factors between warm absorber models. The Bayes factor for the warm absorption model is shown on the horizontal axis, while the Bayes factor for the neutral partial covering model is shown on the vertical axis. The dashed grey line shows the one-to-one relation, with sources lying below the line being better fit with the PL+WA model, and sources above the line being better fit with the PL+PCF model. Sources which do not show evidence for any complex absorption are shown in grey.The values of $K_{wa} = 1.126$ and $K_{pcf} = 1.555$ correspond to purities of 97.5%.

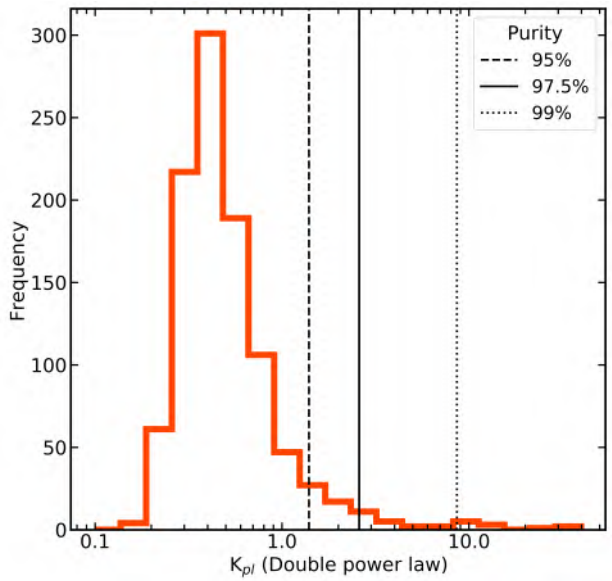

Abbildung 3.15: Distribution of computed K_{pl} values obtained from simulations. Confidence levels of 95%, 97.5% and 99% are indicated using vertical lines of different linestyles.

Tabelle 3.5: Summary of detection thresholds for soft excesses. Column (1) gives the percentile of sources, and column (2) gives the corresponding K_{pl} values. Column (3) gives expected number of false detections. Column (4) gives the number of sources in the sample with K_{pl} values greater than the value given in column (2) and where the model provides the best fit to the data. Column (5) gives the assumed number of true detections in the eFEDS hard X-ray–selected sample of 200 sources.

(1) Percentile	(2) K	(3) Expected false detections	(4) Expected Sources	(5) True detections
95%	1.392	10	40	30
97.5%	2.586	5	29	24
99%	8.613	2	20	18

3.4.1 The soft excess, warm absorber and partial covering samples

After analysing these phenomenological models, based on selecting samples with 97.5% purity, this chapter finds 29 sources with true soft excesses, 29 with warm absorbers, and 25 sources with partial covering absorbers, where by definition there is no overlap between each of these samples. This is because the model must provide the best fit to the data as well as satisfying the Bayes factor criteria. With these defined samples, it is possible to search for possible sources of bias in the relatively limited parent sample used in this chapter. Figure 3.16 shows the detection likelihood in the $2.3 - 5\,\mathrm{keV}$ (DET_LIKE_3) band for sources best fit with each model. To assess any potential differences in the distributions, a Kolmogorov-Smirnov test (KS-test) is used, which compares two samples to compute the likelihood that they are drawn from the same parent sample. No differences between distributions of hard band detection likelihoods are found, with KS-test p-values > 0.1 when comparing each sub-sample, and soft excesses and warm absorbers are found in sources with the lowest and highest computed DET_LIKE_3 values alike. This shows that the selection of the hard X-ray–selected sample does not heavily bias the detection of soft excesses or complex absorbers, and indeed, likely facilitates these measurements as the hard power law can better be constrained. This also motivates using hard X-ray–selected samples of AGN for further investigations of future eROSITA samples.

Considering the true soft excess, it is also of interest to examine the energy at which the two power laws have the same flux. This point marks where the soft excess begins to dominate over the hard corona power law. This value (E_{cross}) has been computed, and the resulting histogram is shown in Fig. 3.17, using the 97.5% purity samples. Regardless of which model is the best fit, the PL+PL model is necessarily used to compute the E_{cross} value. All results are shown in the rest-frame, and the typical error bar is also shown in black in the top right-hand corner. Most sources have E_{cross} values of less than 1 keV, and all are below 2 keV. The median value for the soft excess sample of $E_{cross} = 0.55\,\mathrm{keV}$

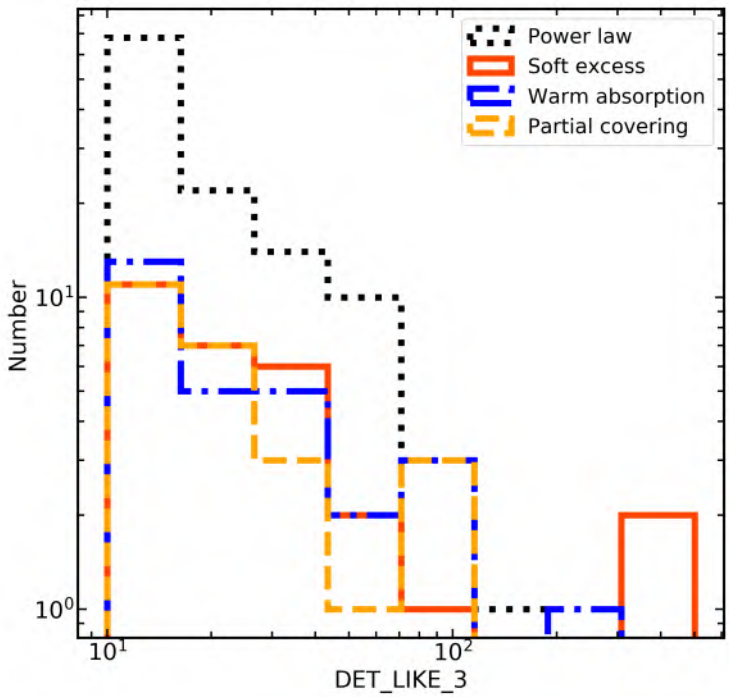

Abbildung 3.16: Distributions of detection likelihood in the 2.3 – 5 keV band (DET_LIKE_3) shown for sources best fit by each model. The vertical axis is given in log-space to highlight the true distribution of the sample. Sources best fit with a power law are shown with a black dotted line, sources best fit with a soft excess are shown as a red solid line, sources best fit with a warm absorber are shown as a blue dash-dot line, and sources best fit with partial covering absorption are shown with an orange dotted line.

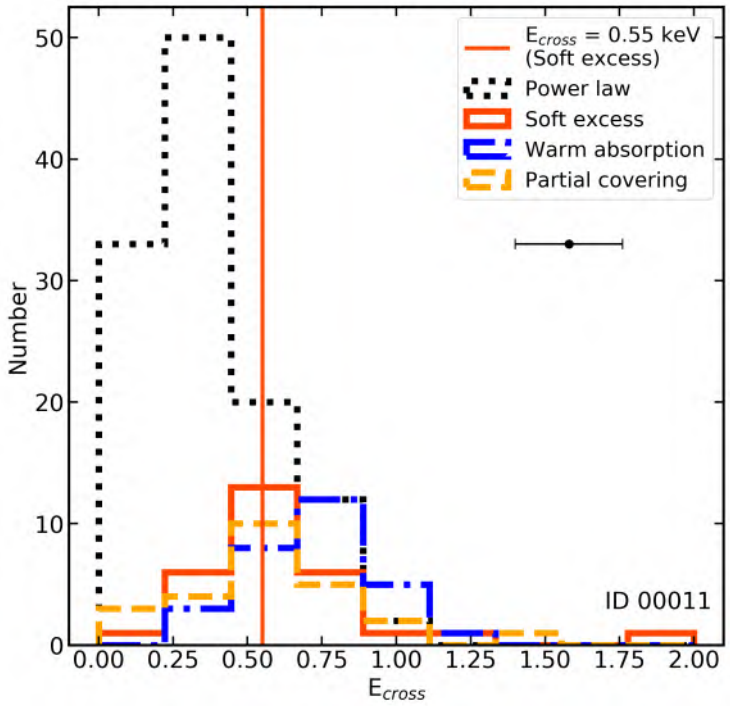

Abbildung 3.17: Distribution of computed rest-frame E_{cross} values. Sources best fit with a power law are shown with a black dotted line, sources best fit with a soft excess are shown as a red solid line, sources best fit with a warm absorber are shown as a blue dash-dot line, and sources best fit with partial covering absorption are shown with an orange dotted line. The median of $E_{cross} = 0.55\,keV$ for the soft excess sample is indicated with a solid red line. The typical error bar is shown in black.

is indicated with a solid red vertical line. This demonstrates the importance of having a good fit in the softest energy X-rays in order to properly characterise the soft excess; fits performed only above $0.5 - 1\,keV$ will likely not be able to fit the true soft X-ray shape.

In addition to examining the Bayes factors and crossing energies, the soft excess can also be defined in terms of the soft excess strength (SE). In this chapter, this is defined as;

$$SE = \frac{F_{SE}}{F_{PL}} \tag{3.2}$$

where F_{SE} is the unabsorbed, rest-frame flux of the soft power law in the $0.2 - 1\,keV$ band, and F_{PL} is the unabsorbed, rest-frame flux of the hard power law in the $0.2 - 1\,keV$ band. The soft excess strength therefore indicates how much excess flux is provided by the soft excess component in the $0.2 - 1\,keV$ band. Separately, the soft flux fraction (SFF) can also be defined;

$$SFF = \frac{F_{0.2-1}}{F_{0.2-10}} \tag{3.3}$$

Tabelle 3.6: Physical soft excess model abbreviation, and corresponding XSPEC implementations.

Name	XSPEC implementation
NTH	`tbabs` × `ztbabs` × (`nthComp` + `constant` × `nthComp`)
REL	`tbabs` × `ztbabs` × `relxill`

where $F_{0.2-1}$ is the unabsorbed, rest-frame flux in the $0.2 - 1\,\mathrm{keV}$ band, and $F_{0.2-10}$ is the unabsorbed, rest-frame flux in the $0.2 - 10\,\mathrm{keV}$ band. Fluxes are measured using the PL+PL model. Therefore, this flux ratio indicates the fraction of the broad-band flux which is emitted in the soft X-ray band.

Histograms for both of these are shown in Fig. 3.18, with the soft excess strength shown in the top panel and the soft flux fraction in the bottom panel. In both panels, it is apparent that the SE and SFF values for sources with soft excesses are higher on average than those which do not. This seems reasonable, as it would be expected that sources with statistically significant soft excesses would have stronger soft emission and would therefore emit a higher fraction of their total flux in the soft X-ray. Furthermore, it is shown that sources with warm absorbers also appear to have very strong soft excesses and soft flux fractions with this model, likely due to the fact that very steep soft photon indices are preferred for these sources in order to approximate the shape of the absorption features, which highlights the important of using the correct model for characterising the soft excess.

To better quantify these differences, a KS-test is performed comparing the distribution of soft excess sources to those best fit with a power law. For the soft excess strength and for the soft flux fraction, the KS-test returns a significance level of 4×10^{-6} and 6×10^{-9}, respectively. This implies that for both qualifications of the soft excess, the null hypothesis that both distributions are drawn from the same parent sample can be rejected with $> 99.999\%$ confidence. As would be expected, the soft excess strength and soft flux fractions are significantly higher for sources with soft excesses are higher on average than for those which do not. This may suggest that sources with and without soft excesses are two distinct populations of AGN. This conclusion, however, still cannot confirm the physical origin of the soft excess; to do this, physically motivated models must be fit to each spectrum, and the evidence compared.

3.5 Physical interpretation for the soft excess

3.5.1 Soft Comptonisation

One physical interpretation for the soft excess is that it is produced via Comptonisation of disc blackbody photons in a secondary warm corona, which is cooler than the hot corona responsible for the primary hard power law Γ_h. The warm corona is hypothesised to have a higher optical depth than the hot corona (e.g. Done et al. 2012; Petrucci et al. 2018, 2020), but as the temperature is lower, the resulting X-ray emission will be a steeper power law

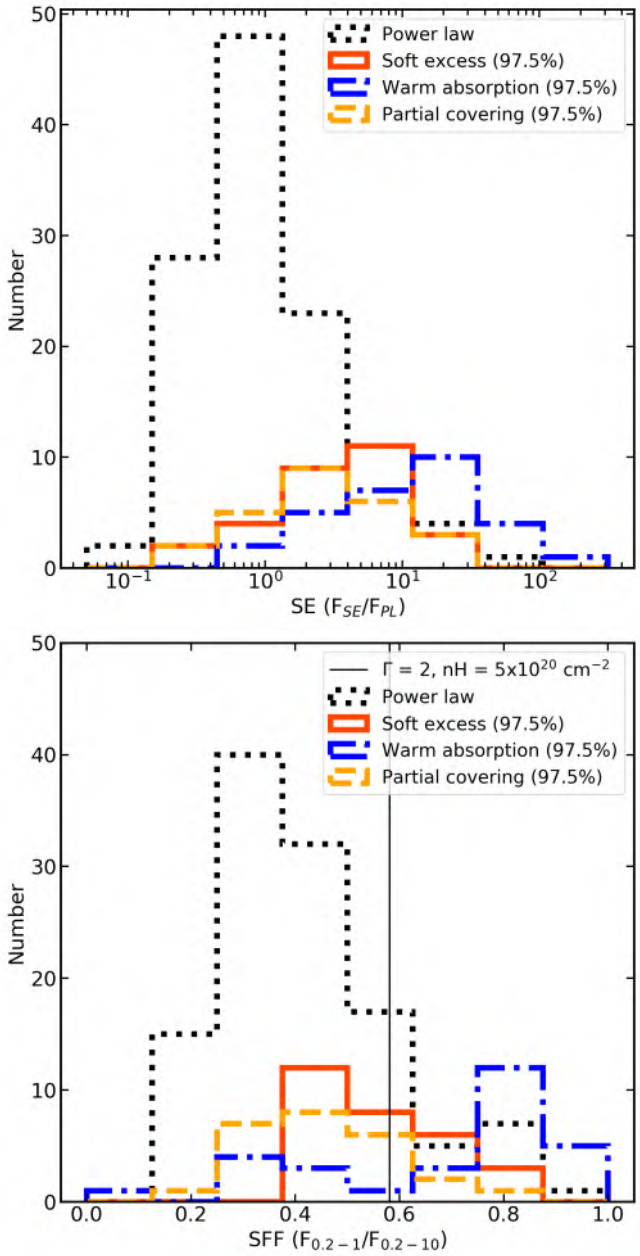

Abbildung 3.18: Distributions of soft excess strength and soft flux fraction. Top: Distribution of soft excess strengths (F_{SE}/F_{PL}) shown for sources best fit by each model. Sources best fit with only with a single absorbed power law are shown with a black dotted line, sources best fit with a soft excess (PL+PL) are shown as a red solid line, sources best fit with a warm absorber are shown as a blue dash-dot line, and sources best fit with partial covering absorption are shown with an orange dotted line. Bottom: same as top, but shown for the soft flux fraction, $F_{0.2-1}/F_{0.2-10}$. The vertical black line shows the expected SFF for a source with $\Gamma = 2$ and $\mathrm{nH} = 5 \times 10^{22}\,\mathrm{cm}^{-2}$.

Tabelle 3.7: Priors used for the physical true soft excess models. Column (1) gives the shortened name of the model. Column (2) shows the lower and upper limit of the host galaxy absorption. Column (3) shows the lower and upper limit of the hard (coronal) power law component. Columns (4) through (7) show the upper and lower limits of the warm Comptonisation parameters, and columns (8) through (10) shows the constraints placed on the blurred reflection parameters. Column (11) gives the constraint placed on the power law normalisation.

(1) Model name	(2) $\log(NH_z)$ [cm^{-2}]	(3) Γ_h	(4) Γ_s	(5) log(c)	(6) kT_e [keV]	(7) q_1	(8) Inclination (degrees)	(9) $\log(\xi)$	(10) Reflection fraction	(11) norm
Prior	log-uniform	uniform	uniform	log-uniform	log-uniform	uniform	uniform	log-uniform	log-uniform	log-uniform
NTH	20 – 25	1 – 3	2 – 3.5	-3 – 1	0.1 – 1	-	-	-	-	-10 – 1
REL	20 – 25	1 – 3	-	-	-	3 – 10	10 – 80	0 – 4	0.1 – 10	-10 – 1

which dominates at low energies. This interpretation has been used successfully to model steep but very smooth soft excesses in type-1 AGN, as the soft Comptonisation will not produce emission or absorption features. To model this in XSPEC, `nthComp` (Zdziarski et al. 1996; Życki et al. 1999) is used, with one `nthComp` component modelling the optically thin hot corona, and another modelling the optically thick, warm corona. The blackbody seed temperature, $kT_{\rm bb}$, is fixed at 1 eV and is linked between the two coronae (e.g. Petrucci et al. 2018, 2020). The X-ray spectral shape is not dependant on this parameter so long as it remains at a reasonable disc temperature of a few eV (Petrucci et al. 2018).

For the hot corona, the photon index is allowed to vary uniformly between one and three in order to capture the likely parameter space. The electron temperature is frozen at 100 keV, well outside the eROSITA bandpass and in agreement with the assumptions from other works (e.g. Fabian et al. 2015). For the warm corona, the photon index is allowed to vary uniformly between two and 3.5, and the electron temperature is allowed to vary uniformly between $0.1 - 1$ keV. These parameter ranges are based on fits and simulations performed by, for example, Petrucci et al. (2018, 2020), where it is demonstrated that these parameters are reasonable when assuming an optical depth of $\sim 10 - 20$. Finally, as in the PL+PL model, the normalisations of the two `nthComp` components are linked, and the flux of the soft corona relative to the hard is set using a cross-normalisation constant. The normalisation component is given a log-uniform prior between -10 and 1, and the cross-normalisation is given a log-uniform prior between -3 and 1. All priors of free parameters are listed in Table 3.7.

3.5.2 Blurred reflection

In a blurred reflection scenario, some of the X-ray photons emitted from the corona are reflected from the innermost regions of an accretion disc, producing a reflection spectrum (e.g. Ross et al. 1999; Ross & Fabian 2005; Dauser et al. 2012; García et al. 2013). As photons strike the disc, they are absorbed, producing deep absorption features and edges. As the atoms de-excite, they produce emission features mostly. For an ionised disc, the

reflected emission is concentrated in the soft X-ray, but also includes a prominent Fe Kα emission line and hard reflection continuum. Due to the fast rotation of the disc and gravitational redshift due to the central black hole, the features in the reflection spectrum are relativistically blurred, producing a soft excess (e.g. Crummy et al. 2006; Jiang et al. 2019). Understanding the reflection spectrum can reveal many properties of the innermost regions of the AGN, including the height and structure of the corona, the ionisation and abundances in the accretion disc, and changes in these parameters over time. This model has been successfully used to probe the geometry of the corona as well as successfully explain the variability and spectral shape of many type-1 AGN (e.g. Zoghbi et al. 2008; Dauser et al. 2012; Gallo et al. 2019; Waddell et al. 2019; Boller et al. 2021).

Here, the reflection spectrum and power law are both modelled using the `relxill` model (Dauser et al. 2012, 2014; García et al. 2013). There are many free parameters in this model, including the inner emissivity index q_1, which describes the illumination pattern of the corona onto the accretion disc. This parameter is allowed to vary uniformly between three and ten, while the outer emissivity index (q_2) is fixed to 3. The inclination, or viewing angle, is allowed to vary uniformly between ten to 80 degrees. This parameter should actually be evenly distributed in cosine space, however, the inclination is typically poorly constrained and difficult to measure correctly, so the uniform prior is acceptable. The black hole is assumed to have maximum spin, in part due to selection effects which make maximum spin AGN brighter and thus easier to detect in flux limited samples (e.g. Vasudevan et al. 2016; Baronchelli et al. 2018; Arcodia et al. 2019), and also due to the fact that the spin is difficult to constrain without high signal-to-noise data in the $4 - 10$ keV band, where the iron Kα line can be modelled (e.g. Bonson & Gallo 2016). The inner radius of the disc is fixed at the innermost stable circular orbit (ISCO; $1.235r_g$ for a maximum spin black hole with $a = 0.998$), and the outer radius is fixed somewhat arbitrarily to $400r_g$, beyond where significant reflection of X-ray photons is possible for moderate coronal heights. The iron abundance in the disc is fixed to solar, and the disc ionisation ($\xi = 4\pi F/n$, where F is the illuminating flux and n is the hydrogen number density of the disc) is allowed to vary between log(ξ) of zero and four.

Since the coronal power law component is also included in `relxill`, no separate power-law model is included. The photon index, Γ, is allowed to vary between one and three, to account for a very broad range of possible indices. The reflection fraction, which describes the fraction of flux from the corona which is reflected from the accretion disc, is allowed to vary uniformly between 0.1 and 10. Here, a reflection fraction (R) of 0.1 would indicate strong beaming (e.g. the corona is outflowing or forms the base of the jet), a reflection fraction of one suggests that half the flux from the corona is reflected off the disc, while the other half is observed directly, and a reflection fraction of 10 is a strong indicator for light bending (e.g. the corona is close to the disc such that the gravitational pull of the black hole bends the path of the light towards the disc). Finally, the normalisation component is given a log-uniform prior between -10 and one so that AGN of an extreme range of fluxes can be modelled. All priors of free parameters are listed in Table 3.7.

There are several different flavours of `relxill`, all intended to model different physical properties of the innermost regions of the AGN (e.g. Dauser et al. 2016; Jiang et al. 2019).

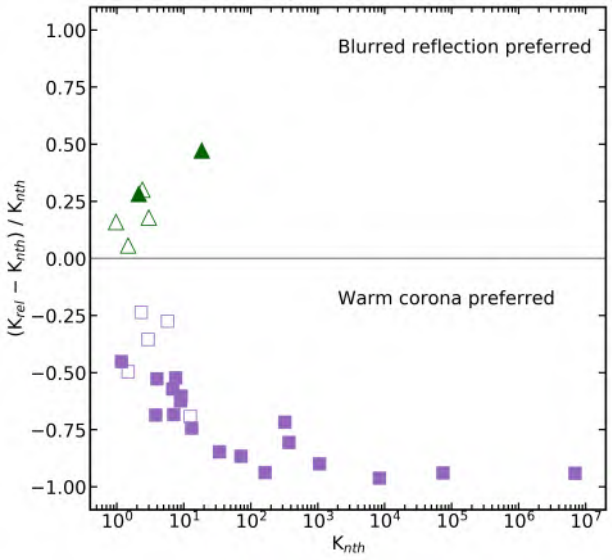

Abbildung 3.19: Comparison of Bayes factors for the warm corona and relativistic blurred reflection models. Bayes factors for the warm corona model are shown on the horizontal axis, and the difference between Bayes factors normalised by the warm corona Bayes factor is shown on the vertical axis. Open shapes indicate soft excesses with 97.5% confidence, and filled shapes indicate 99% confidence. Sources best fit with a warm corona are shown as purple squares, and sources best fit with blurred reflection are shown as green triangles.

Users can choose to assume a lamp-post geometry (`relxilllp`), a varying disc density (`relxillD`), among other changes. I also freeze many parameters in this analysis (e.g. the iron abundance, outer emissivity index, and black hole spin, although these have been shown to vary, in some cases dramatically, between AGN (Zoghbi et al. 2008; Fabian et al. 2009; Daly & Sprinkle 2014; Reynolds 2019). In particular, many parameters are best constrained using the iron Kα line profile, as the iron line is broadened due to the strong relativistic effects in the central region. Given the limited eROSITA sensitivity and high background levels above $\sim 5\,\mathrm{keV}$, this is very difficult for most sources, even those in the hard sample presented in this chapter. Nevertheless, this simplified treatment of relativistic reflection still has the potential to capture sources which display the typical characteristics of a blurred reflection spectrum.

3.5.3 Model selection

Having fit both of the models described above to the 29 sources in the soft excess sample, the evidence for each model can be computed and compared in order to determine which model is preferred, with Bayes factors computed as given in Table 3.3. The results of this comparison are presented in Table 3.8. The final column in the table indicates which is the preferred model for each source. Out of the 29 sources, six are better fit with blurred reflection and the remaining 23 are best fit with soft Comptonisation.

Tabelle 3.8: Evidence comparison for the 29 sources in the soft excess sample. The source ID is listed in column (1). Columns (2) and (3) show the Bayesian evidence, where the values are all normalised by subtracting the highest fit value. A negative number in the column therefore indicates the worse fitting model, while a value of zero shows the preferred model. In the final row, the same exercise is performed for the full sample. Column (4) lists the name of the best fitting model.

(1)	(2)	(3)	(4)
eROID	$\ln(Z_{nth}) - \ln(Z_{BEST})$	$\ln(Z_{rel}) - \ln(Z_{BEST})$	Best model
00001	0.0	-0.69	NTHCOMP
00004	0.0	-3.28	NTHCOMP
00007	0.0	-1.26	NTHCOMP
00011	0.0	-2.83	NTHCOMP
00029	0.0	-0.60	NTHCOMP
00034	0.0	-2.77	NTHCOMP
00035	-0.25	0.0	RELXILL
00038	0.0	-0.85	NTHCOMP
00039	0.0	-2.80	NTHCOMP
00045	0.0	-2.30	NTHCOMP
00054	0.0	-1.88	NTHCOMP
00057	-0.16	0.0	RELXILL
00076	0.0	-0.32	NTHCOMP
00121	0.0	-0.92	NTHCOMP
00122	-0.15	0.0	RELXILL
00153	0.0	-0.98	NTHCOMP
00176	-0.05	0.0	RELXILL
00200	0.0	-0.75	NTHCOMP
00204	-0.39	0.0	RELXILL
00216	0.0	-2.01	NTHCOMP
00237	0.0	-0.74	NTHCOMP
00288	0.0	-1.15	NTHCOMP
00340	0.0	-1.18	NTHCOMP
00358	0.0	-1.36	NTHCOMP
00426	0.0	-1.64	NTHCOMP
00760	-0.26	0.0	RELXILL
00784	0.0	-0.27	NTHCOMP
01136	0.0	-0.44	NTHCOMP
01736	0.0	-1.16	NTHCOMP
ALL	0.0	-30.93	NTHCOMP

More closely examining the sample, it is apparent that many of the sources which are best fit with blurred reflection have very small differences in Bayes factors between models. This is not the case for sources best fit with soft Comptonisation, where some sources have much larger differences in Bayesian evidence than with blurred reflection. This indicates that while not all sources are fit well with all models, all sources are relatively well fit with soft Comptonisation. This effect can also be seen in Fig. 3.19, which shows the Bayes factor (K_{nth}) for the warm corona for each source plotted with the normalised difference in Bayes factor values, (K_{rel} - K_{nth})/K_{nth}. Sources which lie above zero on the y-axis (shown with green triangles) are best fit with blurred reflection, and sources which lie below zero (shown in purple squares) are best with with the warm corona model. Many sources best fit with the warm corona model are much better fit with this model. Furthermore, sources which have more statistically significant soft excesses (shown with filled in symbols) are far more likely to be best fit with a warm corona, with only 2/20 preferring a blurred reflection model.

The last line of Table 3.8 shows the evidence comparison for the full sample, following, e.g. Baronchelli et al. (2018). Unsurprisingly, the best fitting model for the full sample is the soft Comptonisation model, however it should be noted that many parameter values are poorly constrained and have large errors for these complex models.

3.5.4 Model parameters

It is also of interest to compare the properties derived from the phenomenological double power law model of the soft excess sources, shown in Fig. 3.20. Sources best fit with blurred reflection are again shown in dark green, and sources best fit with a warm corona model are shown in purple. Median values for each sub-sample are shown with vertical lines. While the distributions of hard X-ray photon indices are very similar, the distributions of soft photon indices differ, with the median value being much higher for sources best fit with blurred reflection. While this result is not significant when using e.g. the Anderson-Darling test, there are very few sources best fit with blurred reflection so it is difficult to make firm conclusions. Nevertheless, the result is intriguing, as it may present a diagnostic tool to differentiate between a warm corona and a blurred reflection soft excess, and will be discussed further in Sect. 3.8.

Regarding the parameters of the warm Comptonisation models, the median hot corona photon index is $\Gamma = 1.61$, and the median warm corona photon index is $\Gamma = 3.15$. The median warm corona temperature is kT$= 0.45$ keV, which is consistent with other studies (Petrucci et al. 2018, 2020). The distributions of the best-fit parameters for all soft excess sources are shown in Fig. 3.21, where the warm corona values and errors are indicated with black squares. The red lines show different values of the optical depth, where the warm corona optical depths are between $\tau = 5$ and $\tau = 20$ and the optical depths for the hot corona are $\simeq 1$. While the warm corona photon index is also consistent with previous works, the hot corona has a much flatter spectral index than expected (e.g. $\Gamma \sim 1.8 - 1.9$ in previous studies, and $\Gamma \sim 2.0$ in typical eROSITA sources). This result is unexpected and is further discussed in Sect. 3.8.

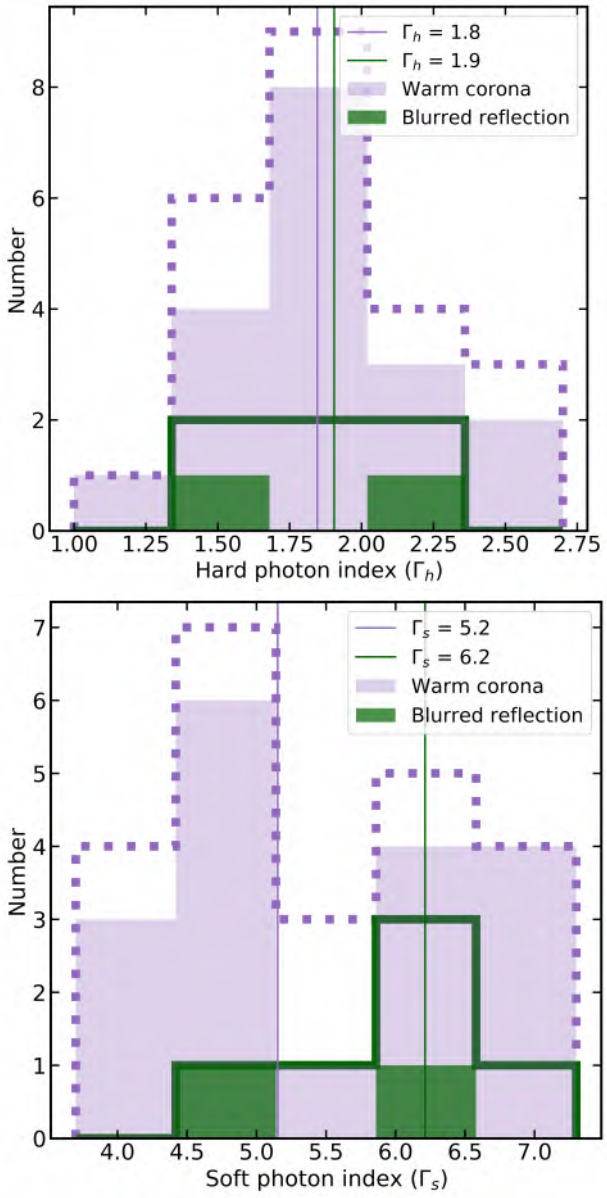

Abbildung 3.20: Distributions of soft and hard photon indices separated by best soft excess model. Top: Distributions of hard photon index obtained from the PL+PL modelling for sources in the soft excess sample. Sources which are best fit with soft Comptonisation are shown with a purple dotted line and sources best fit with blurred reflection are shown with a solid dark green line. Median values for each sample are shown with vertical lines in the corresponding colours. The shaded histograms indicate the sources with 99% significance on the soft excess. Bottom: As top, but showing the soft X-ray photon index.

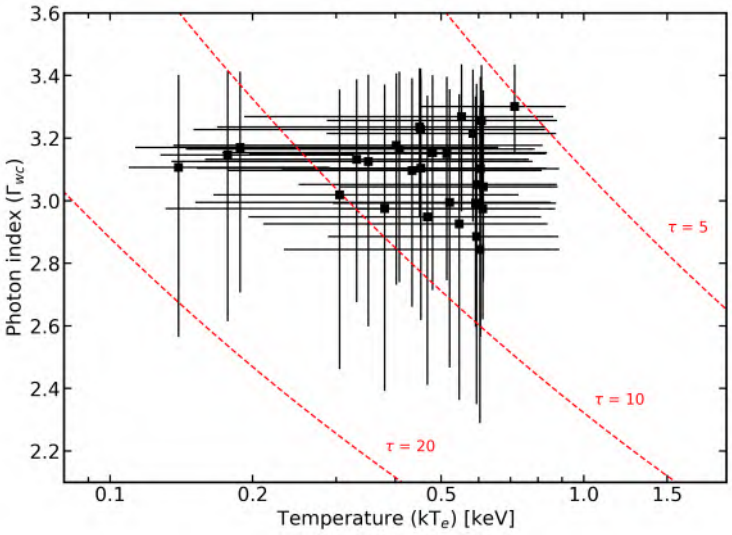

Abbildung 3.21: Warm corona photon indices and temperatures derived from the soft Comptonisation modelling. Lines of constant optical depth are shown in red.

Moving to the parameters of the blurred reflection modelling, it is apparent that many parameters are not well constrained, likely due to the lower quality of many eFEDS spectra as well as the absence of a high signal-to-noise iron line. Examining the best-fitting parameters for the two blurred reflection sources with soft excesses at >99% significance (corresponding to the filled green triangles in Fig. 3.19), both have intermediate disc ionisations of $\xi \sim 100$, intermediate inclinations of ~ 40 degrees (which are expected for type-1 AGN), and of particular note, high reflection fractions $R >> 1$. In fact, examining all sources best fit with blurred reflection, it is found that all have best fit $R > 1$, although not all are constrained to be > 1. This suggests that the spectral fitting method presented in this chapter is preferentially identifying sources with very strong reflection components. Indeed, in some cases of sources best fit with blurred reflection, it seems there is more excess emission around $0.7 - 0.9$ keV, which may correspond to iron-L which is present in the reflection spectrum but not the Comptonisation spectrum, which may explain why the blurred reflection model is preferred. Spectral modelling with brighter sources with more counts in the $4 - 7$ keV band, or including data above 8 keV, would help to identify the iron line and Compton hump if present and thus better constrain blurred reflection parameters.

To visually demonstrate differences between the spectral models, Fig. 3.22 shows an example source, ID 00034, which is best fit with a warm corona model. The data are shown in black along with the background model in a black dashed line, an absorbed power law in a grey dotted line, the blurred reflection model in a green dash-dot line, and the best fit warm corona model in purple. From the spectrum, it can be seen that the blurred reflection model under-estimates the flux in the $0.2 - 0.3$ keV energy band, and over-estimates the flux around 0.5 keV where the blurred reflection model features strong iron emission. Both models however clearly provide a much better fit than the absorbed power law model,

Tabelle 3.9: Summary of blurred reflection parameters for the two sources with highly significant soft excess ($> 99\%$) which are best fit with blurred reflection. Column (1) gives the eROSITA ID, column (2) gives the emissivity index, column (3) gives the disc inclination, column (4) gives the photon index, column (5) gives the disc ionisation and column (6) gives the reflection fraction.

(1)	(2)	(3)	(4)	(5)	(6)
eROID	q_1	Inclination	Γ	$\log(\xi)$	$\log(R)$
00035	$7.1^{+2.0}_{-2.3}$	36^{+21}_{-18}	$2.45^{+0.10}_{-0.11}$	$1.8^{+1.4}_{-1.2}$	$0.4^{+0.5}_{-0.8}$
00204	$4.3^{+2.6}_{-1.0}$	43^{+19}_{-21}	$1.86^{+0.22}_{-0.25}$	$2.0^{+0.8}_{-0.9}$	$0.8^{+0.1}_{-0.3}$

which fails to reproduce the spectral shape at most energies.

3.6 Spectral properties

3.6.1 Luminosity-redshift plane

To study the distribution of soft excesses and warm absorbers in a parameter space which can easily be compared to other surveys, sources are plotted in the L-z plane in Fig. 3.23, where most redshifts are spectroscopic. The $2-10\,\mathrm{keV}$ luminosity is estimated from the baseline absorbed power law model. Sources best fit with an absorbed power law are shown as grey crosses, sources best fit with warm absorbers are shown as blue circles, sources best fit with partial covering are shown as orange pentagons, and sources best fit with soft excesses are shown as red squares. In this way, I differentiate from the luminosity redshift plane already presented in Nandra et al. (2024). Different opacities and fill-styles indicate different purity thresholds for the best fit model, as previously defined. The rest-frame, absorption corrected X-ray luminosities of sources with soft excesses and warm absorbers follow those of sources best fit with an absorbed power law. Most sources with soft excesses can be detected up to about $z = 0.5-0.6$, above which the soft excess is presumably shifted out of the observed band, while complex absorption can be detected up to about $z = 1$, with a few sources at higher redshifts. Interestingly, the highest redshift source in the sample, with a spectroscopic redshift of $z = 3.277$ (see Nandra et al. 2024), shows complex absorption best fit by a warm absorber. However, because this source has very few counts below $\sim 1\,\mathrm{keV}$, it is hard to determine definitively the nature of the soft X-ray complexity without data with a higher signal-to-noise ratio.

3.6.2 Characterising the soft excess and complex absorption

Having completed the modelling for all 200 sources in the hard X-ray–selected sample of AGN in eFEDS, the soft excess, warm absorption and partial covering sub-samples can be examined in more detail to search for distinctive characteristics. When considering the full

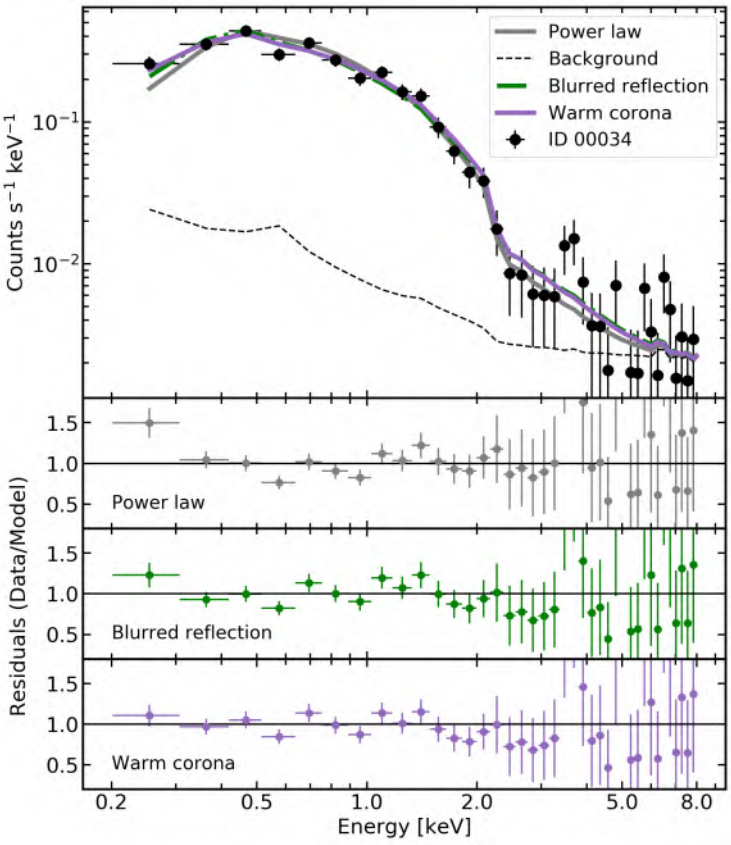

Abbildung 3.22: ID 00034 (Z=0.1027), a source best fit with a warm corona model. The spectrum (re-binned for display) is shown in black, the background model is shown as a black dashed line, the power law model is shown as a grey dotted line, a blurred reflection model is shown in green dash-dot line, and the best fitting warm corona model is shown in purple. The bottom three panels show the residuals for the power law, blurred reflection, and a warm corona, respectively. Data have been re-binned for display purposes.

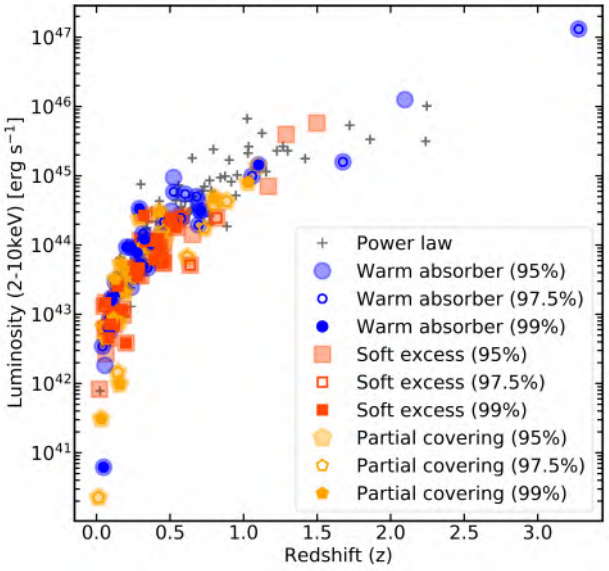

Abbildung 3.23: Distributions of redshifts and $2 - 10\,\mathrm{keV}$ un-absorbed X-ray luminosity for each source. Sources best fit with an absorbed power law are shown as grey crosses, sources with soft excesses are shown as red squares, and sources with warm absorbers are shown with blue circles.

sample of hard X-ray–selected AGN, only $\sim 15\%$ of sources show strong statistical evidence for a soft excess. However, this fraction can increase significantly when only considering a small parameter space. Figure 3.24 shows the distribution of photon index and host-galaxy column density. These values are obtained from the `ztbabs` component of the baseline PL model (and thus not including any additional spectral components), regardless as to the best-fit model for each source. In this way, the properties of sources can be compared for a naive approach wherein it is assumed that all spectra can only be fit with a power law.

Here, the sources with soft excesses are heavily clustered at large photon indices and small host galaxy column densities. This is sensible, as the single power law attempts to explain both the high energy component and the steep soft excess with a single power law, increasing the slope. Selecting only sources with photon indices larger than two and host galaxy column densities less than $2 \times 10^{20}\,\mathrm{cm}^{-2}$, 42% of the sources have soft excesses. These are not intrinsic properties of these sources; indeed, it is found that when the soft excess is modelled correctly, the measured mean photon index decreases by 0.35, from a mean of 2.15 when using only one power law to a mean of 1.8 for the hard photon index when the second power law is added. These values are similar to the mean value of the sample of sources best fit with only a power law.

Examining now the parameter space region populated by the complex absorption sources, many of the sources with extremely low photon indices ($\Gamma < 1.4$) and a range of column densities are best fit with warm absorption or partial covering, which likely explains why these photon indices appear so flat compared to the more typical values of $\Gamma \sim 1.8$. When the correct absorption model is applied, the photon index increases to more

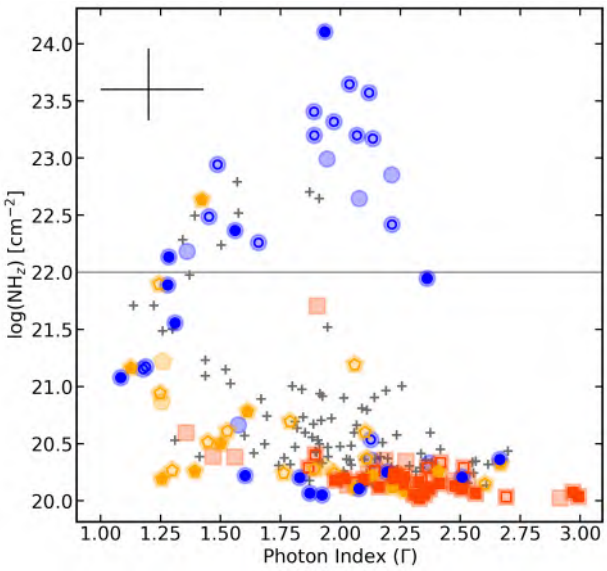

Abbildung 3.24: Distributions of photon indices and the host-galaxy absorption column densities, These values are always measured using the baseline PL model, irrespective of the true best-fit model for each source. Sources with soft excesses are shown as red squares, sources with warm absorbers are shown with blue circles, sources best fit with partial covering are shown with orange pentagons, and sources best fit with the baseline power law model are shown as black crosses. Marker styles represent samples of different purities, as described in sections 3.3 and 3.4. The typical error bars are shown in the top right corner, and the horizontal grey line indicates a column density of $10^{22}\,\mathrm{cm}^{-2}$.

reasonable values for many of these sources. However, there also appears to be a large cluster of sources with column densities of $> 10^{22}$ and photon indices around $\Gamma \sim 2$ (see upper middle of Fig. 3.24). These sources may be of particular interest, as they suggest the presence of Compton-thin AGN in eFEDS, which may show absorption and scattered emission from the torus.

Almost all sources with column densities $> 10^{22}\,\mathrm{cm}^{-2}$ (18/26), and all eight sources with column densities $> 10^{23}\,\mathrm{cm}^{-2}$ have evidence for a warm absorber. This raises the very interesting possibility that many of the AGN that might have been classified as Compton-thin obscured AGN are actually better described with a warm absorber model, and suggests that eROSITA is more likely to probe these complex absorption sources as opposed to AGN obscured by neutral, distant gas. Such column densities are likely too large to be associated with absorption on host-galaxy scales, and must instead originate in the torus. However, these large column densities would not be expected in type-1 AGN, which are typically unobscured. Searching for obscured ($> 10^{22}\,\mathrm{cm}^{-2}$) sources which also have SDSS spectra, 17 sources were observed.

Examining each of these sources individually, many show deep absorption edges in the $\sim 0.5-1\,\mathrm{keV}$ band. When modelled using a single absorbed power law, the absorption edge is well fit with the single absorber, but the emission is significantly under-fit at low energies (e.g. $\sim 0.2-0.5\,\mathrm{keV}$). An example of this is shown in Fig. 3.25, which shows a spectrum (ID 00439, also presented in Brusa et al. 2022) folded with the background model, an absorbed power law model, a partial covering absorber, and a warm absorber. Without a warm absorber, the flux is significantly under-estimated in the softest X-ray energies (below $0.4-0.5\,\mathrm{keV}$). When the warm absorber is added, the host galaxy column density modelled using `ztbabs` is consistent with $10^{20}\,\mathrm{cm}^{-2}$, and the absorption edge better describes the emission below $\sim 1\,\mathrm{keV}$. The partial covering absorber also fails to produce the correct spectral shape. Furthermore, when the warm absorber is added to these sources, the host-galaxy column densities are all consistent with $10^{20}\,\mathrm{cm}^{-2}$, which is consistent with values found in the Milky Way and is consistent with other host galaxies. This can also be seen in the corner plot for the warm absorption model of source ID 00439, shown in Fig. 3.26. Interestingly, the optical spectrum for source was found to show evidence for an ionised outflow Brusa et al. (2022), further supporting the warm absorber X-ray model. This source has relatively fewer X-ray counts as it is heavily absorbed in the soft band, and key parameters are less well constrained. There is some degeneracy between the power law index and normalisation, and the ionisation appears to be low, but the marginal posterior probability distribution has a small secondary peak at a higher ionisation. Most strikingly, the column densities of the host galaxy and warm absorber cannot be independently constrained, though the host galaxy absorption is consistent with $10^{20}\,\mathrm{cm}^{-2}$, and the warm absorber component significantly improves the fit. The constraining of the absorbing column densities will be discussed in further detail in Sect. 3.8.

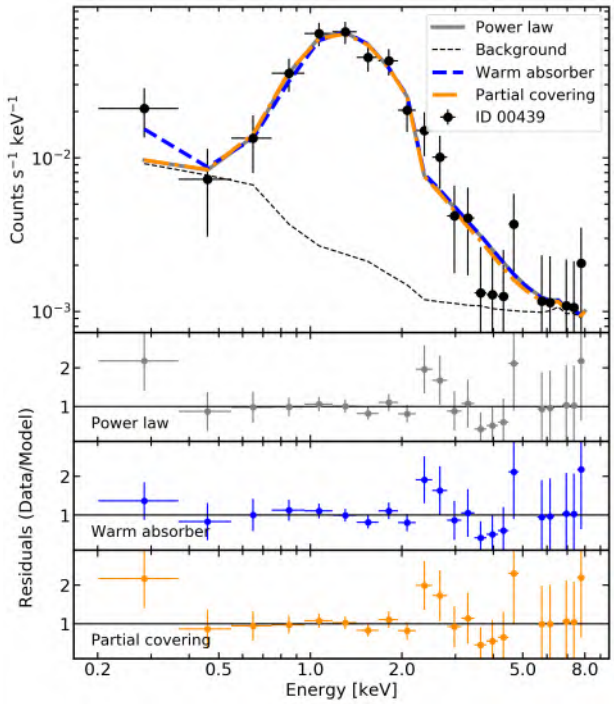

Abbildung 3.25: Example spectrum of ID 00439 (z = 0.6027), which is best fit fit a warm absorber model. The spectrum is shown along with the background model (black dashed line), the absorbed power law model (grey), a partial covering absorption model (orange), and the best-fit warm absorber model (blue). The bottom three panels show the residuals for the power law, warm absorber and partial covering models, respectively. Data have been re-binned for display purposes.

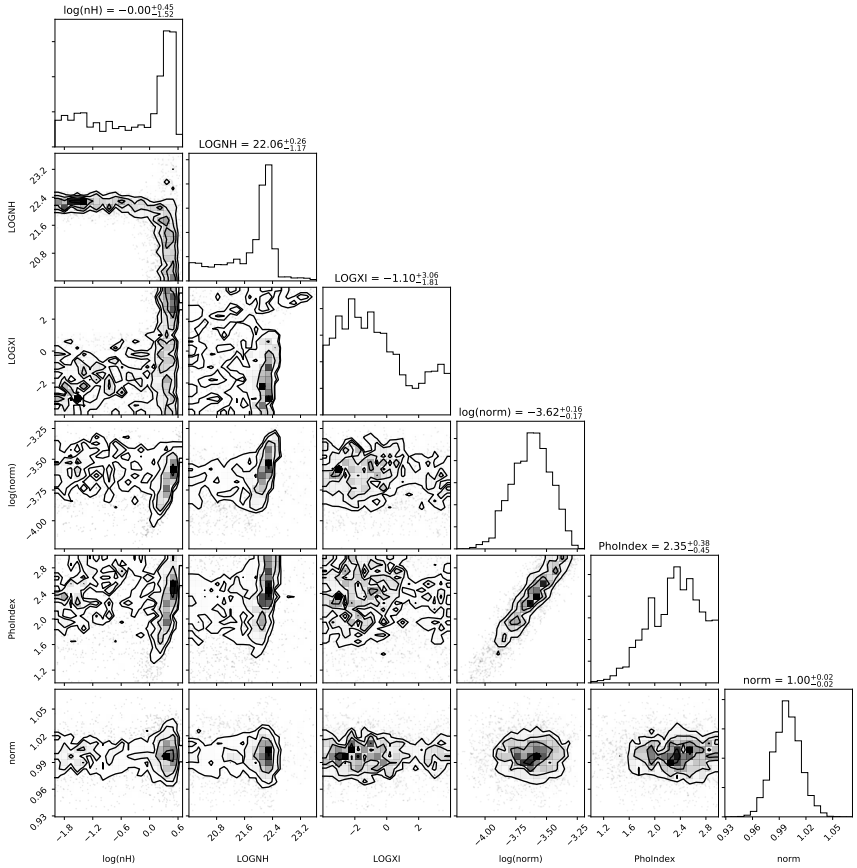

Abbildung 3.26: Corner plot for source ID 00439, which shows evidence for Compton-thin absorption when modelled with an absorber power law but is best fit with a warm absorption model. The diagonal panels show the marginal posterior probability distribution for each parameter, while the other panels show the conditional probability distribution functions for each pair of parameters. Here, log(nH) is the host galaxy absorber column density (in units of $\times 10^{22}\,\mathrm{cm}^{-2}$), LOGNH is the column density of the warm absorber (cm^{-2}), LOGXI is the ionisation of the warm absorber (ergs cm s^{-1}), log(norm) is the power law normalisation, PhoIndex is the photon index of the power law, and norm is the relative renormalisation of the background model with respect to the source model, which is consistent with unity.

3.7 Trends with luminosity, black hole mass and accretion rate

3.7.1 Differences according to the best-fit model

Black hole masses and accretion rates have been presented in Chapter 2, for the visually inspected broad-line objects in the eFEDS hard X-ray selected sample (Nandra et al. 2024). Here we examine the relationship between the soft X-ray best fit models with the properties derived from the optical spectroscopy, specifically the black hole masses and bolometric luminosities (derived from the optical luminosity) and the Eddington ratio $\lambda_{\rm Edd}$.

Fig. 3.27 shows the distribution of black hole masses and bolometric luminosities for the sources in the visually inspected broad line sample. Lines of constant Eddington ratio are shown in black, with different linestyles indicating different Eddington ratios. A number of the highest Eddington ratio sources also show very low black hole masses and low/moderate luminosities, and appear best fit with a soft excess model. Most of the sources with warm absorbers tend to be found at lower Eddington ratio, and three sources with warm absorbers show large black hole masses on the order of few $\times 10^9$ solar masses.

In Nandra et al. (2024), the distributions of redshift, black hole mass, bolometric luminosity and Eddington ratio for the eFEDS hard sample of AGN are discussed in detail. In this section, these distributions are re-examined, but separating the sources based on the best fitting soft X-ray models. These distributions are shown in Fig. 3.28, with median values for each sub-sample indicated with vertical dashed lines in the corresponding colours. A lower median Eddington ratio is recovered for sources best fit with warm absorbers as compared to those best fit with a soft excess, but the difference is not statistically significant.

In general, it is also seen that sources with soft excesses tend to have lower than average $M_{\rm BH}$ and FWHM values (as compared to the full sample), while those with warm absorbers tend to have larger $M_{\rm BH}$ and FWHM. Also of interest is to examine the sources with FWHM values consistent with those found in narrow-line Seyfert 1 (NLS1) galaxies (e.g. Boller et al. 1996), which are classified based on their Hβ line widths and other optical line properties (Osterbrock & Pogge 1985; Goodrich 1989) and are believed to host younger, lower mass black holes accreting at a high fraction of the Eddington limit (e.g. Pounds et al. 1995; Grupe 2004; Gallo 2018; Waddell & Gallo 2020). For sources with FWHM < 2000km s^{-1}, four out of five sources show evidence for a soft excess, which is consistent with previous findings that NLS1s typically have strong and steep soft excesses (e.g. Boller et al. 1996; Waddell & Gallo 2020). These result are statistically significant; when we compare the KS-test p-value of sources best fit with a power law to those which require a soft excess, the p-value is 0.015 for the $M_{\rm BH}$ distributions and 0.031 for the FWHM distributions. This will be discussed further in Sect. 3.8.5.

To look at this in more detail, it is useful to examine the fraction of sources best fit by each model as a function of accretion rates. This is shown in Fig. 3.29, where the fraction of sources best fit with a soft excess, warm absorber or partial covering component are

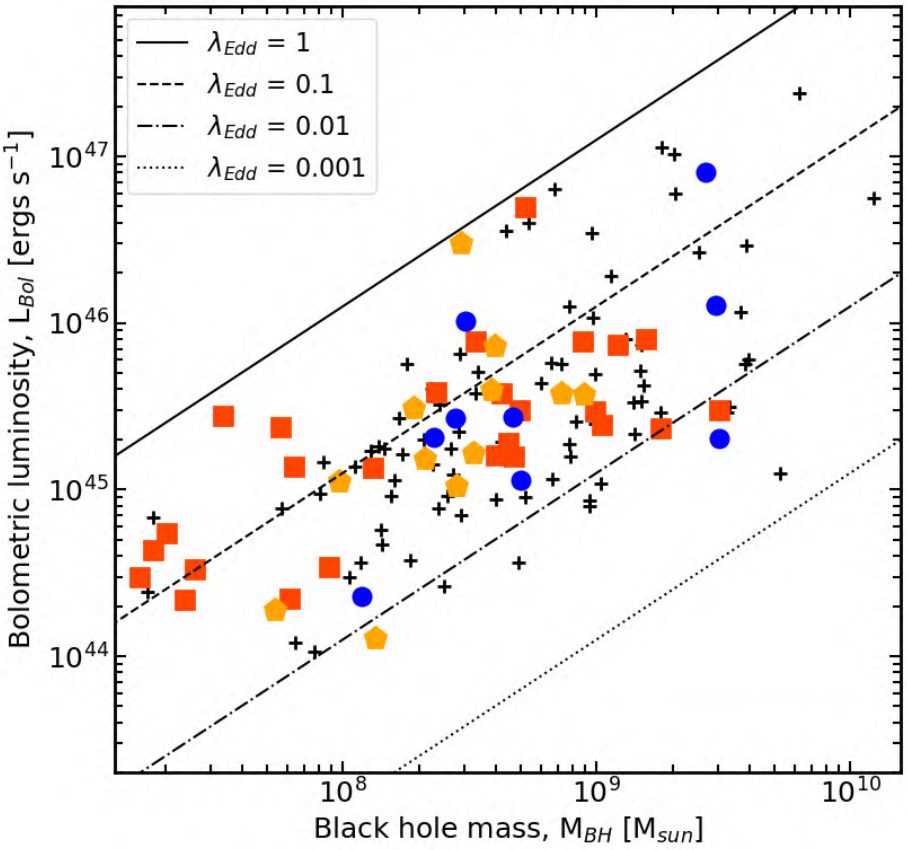

Abbildung 3.27: Distribution of black hole masses and bolometric luminosities calculated from the optical spectra. Sources best fit with an absorbed power law are shown as grey crosses, sources with soft excesses (97.5%) are shown as red squares, sources with warm absorbers (97.5%) are shown with blue circles, and sources with partial covering (97.5%) absorbers are shown with orange pentagons.

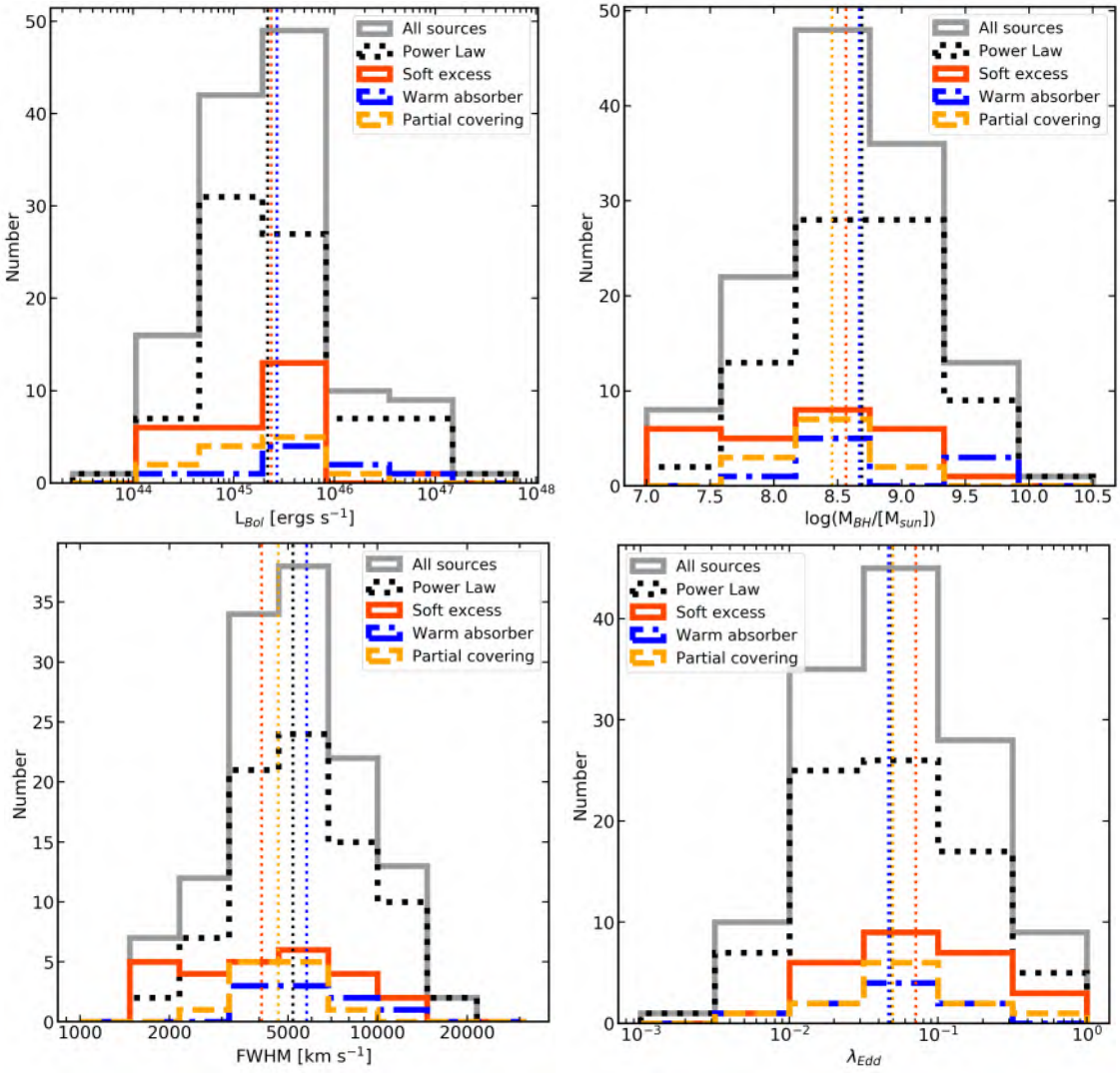

Abbildung 3.28: Histograms of key optical-derived properties, shown as: all sources (grey), sources best fit with a power law (black dotted line), sources best fit with a soft excess (red solid line), sources best fit with a warm absorber (blue dot-dash line) and sources best with with partial covering (orange dashed line). Median values for each sub-sample are indicated with vertical lines in corresponding colours. Top left: distributions of bolometric luminosities. Top right: distributions of black hole masses. Bottom left: distributions of FWHM of the Hβ, Mg II, or C IV optical lines. Bottom right: distributions of the Eddington ratios ($\lambda_{\rm Edd}$).

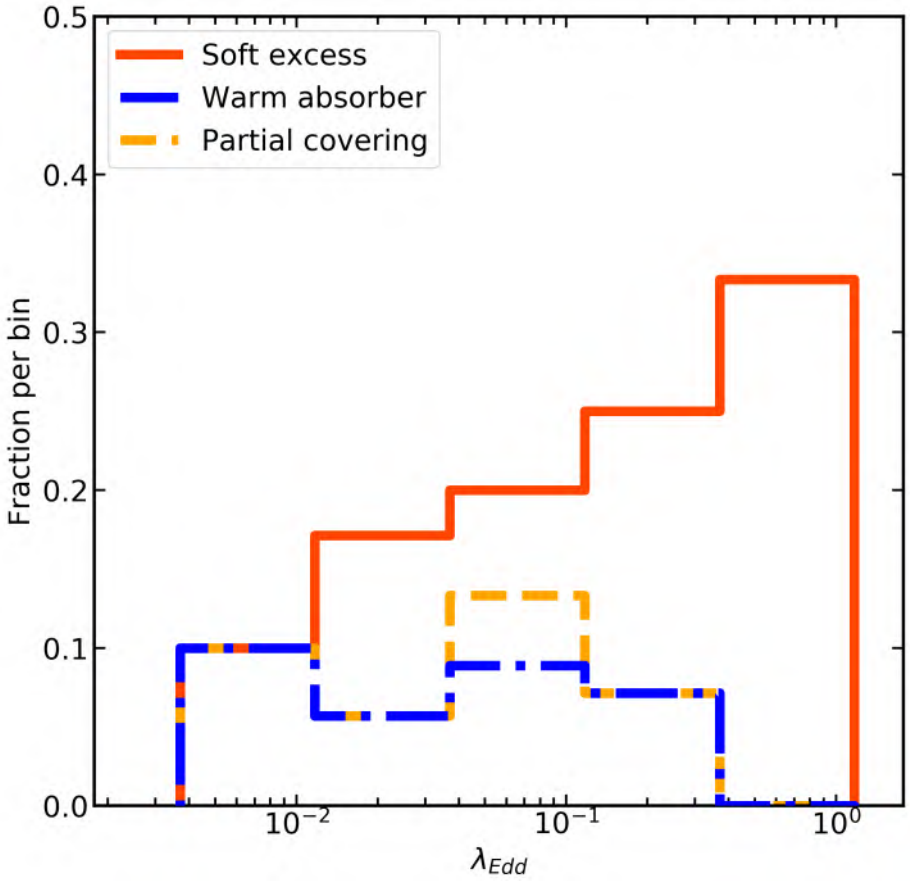

Abbildung 3.29: Fraction of the total sources best fit with each model (soft excess, warm absorber, partial covering), shown per bin using the same binning as the bottom-right panel of Fig. 3.28. Sources best fit with soft excesses are shown in red, sources best fit with warm absorbers are shown as blue dashed lines, and sources best fit with partial covering are shown as orange dash-dot lines.

shown, using the same binning as in the bottom-right panel of Fig. 3.28. Sources best fit with soft excess tend to have higher Eddington ratios, and the fraction of sources with soft excess increases with increasing Eddington ratio. By contrast, the sources with warm absorbers and neutral partial covering absorbers typically have lower Eddington ratios of $\lambda_{\rm Edd} \sim 0.005 - 0.03$, with no absorption found in the higher Eddington ratio regime.

Some studies (e.g. Shemmer et al. 2008; Risaliti et al. 2009; Brightman et al. 2013; Trakhtenbrot et al. 2017) have shown a shallow correlation between the X-ray photon index and the Eddington ratio, although this has been debated or refuted in other works (e.g. Laurenti et al. 2022). This correlation is often explained in the context of the eigenvector 1 (EV1) space, wherein the main variance in samples of quasars is shown to take the form of an anti-correlation between the strength of the FeII and [OIII] optical emission, which has been confirmed to be driven by the Eddington ratio (Boroson & Green 1992; Shen & Ho

2014; Wolf et al. 2020). To examine this in the context of eFEDS, the (hard) photon index from the best fitting spectral model is shown in Fig. 3.30, plotted with the Eddington ratios, with histograms also shown for each parameter. We do not recover a significant correlation between Γ and $\lambda_{\rm Edd}$. Performing a KS test on the photon indices for the sources best fit with warm absorbers and soft excesses, a p-value of 0.4 is obtained, suggesting that there is no difference in photon index. This also suggests that some additional intrinsic properties of these systems is responsible for the observed differences. Separately, the Eddington ratio was also checked along with the soft excess strength (F_{SE}/F_{PL}), however no significant correlation between these parameters was found. These findings will be discussed in more detail in Sect. 3.8 and Chapter 6.

3.7.2 Comparison to Waddell et al. (2024a)

In the published work of Waddell et al. (2024), in particular sections 5.3 and 6.4 from the published manuscript, an analysis is presented of properties derived from the optical spectra, with particular reference to the Eddington ratios. While the majority of the analysis in this Chapter corresponds exactly to that presented in (Waddell et al. 2024), the sample of objects used in the bolometric luminosity, black hole mass and accretion rate analysis differs between this thesis and the published paper. In particular, the values presented in Chapter 2, in Nandra et al. (2024) and in this chapter are only considered reliable for objects which show evidence for broad Hb, MgII or CIV based on visual inspection. These are fewer in number than the black hole mass sample of (Waddell et al. 2024), which did not rely on visual inspection and instead followed Wu & Shen (2022), selecting sources where the flux of the line divided by the error on the flux must be > 2, and where the Eddington ratio divided by the error must be > 1. This affects the conclusions somewhat, in that the reduced sample size is insufficient for performing some statistical test and does not show the same clear distinctions between Eddington ratio distributions as a function of spectral model. This section presents a more detailed description of reasons behind this change, an explanation of the steps taken, and a comparison of the new results with those from Waddell et al. (2024).

The optical spectral analysis was performed using methodology developed for the analysis of bright quasars (Wu & Shen 2022). The method uses a multi-order polynomial in order to approximate the AGN host galaxy continuum, and did not include host galaxy subtraction. Upon visual inspection of the spectra, it became apparent that some sources showed strong host galaxy contamination or other unusual phenomena. In these cases, the width of the broad lines was often overestimated, leading to erroneously high black hole mass measurements and low Eddington ratios. An example of such a source is shown in Fig. 3.31, which shows an example of an optical spectrum which met the criteria for a good black hole mass measurement in Waddell et al. (2024). It was also measured to have a Hβ FWHM of $> 10,000$ km s^{-1}. Upon more careful visual inspection of the source, it is apparent that there is some contribution from the host galaxy in the spectrum, as can be seen from the significant reddening at lower wavelengths as well as from the curvature in the continuum. The broad curvature feature can be modelled with a very broad Hβ com-

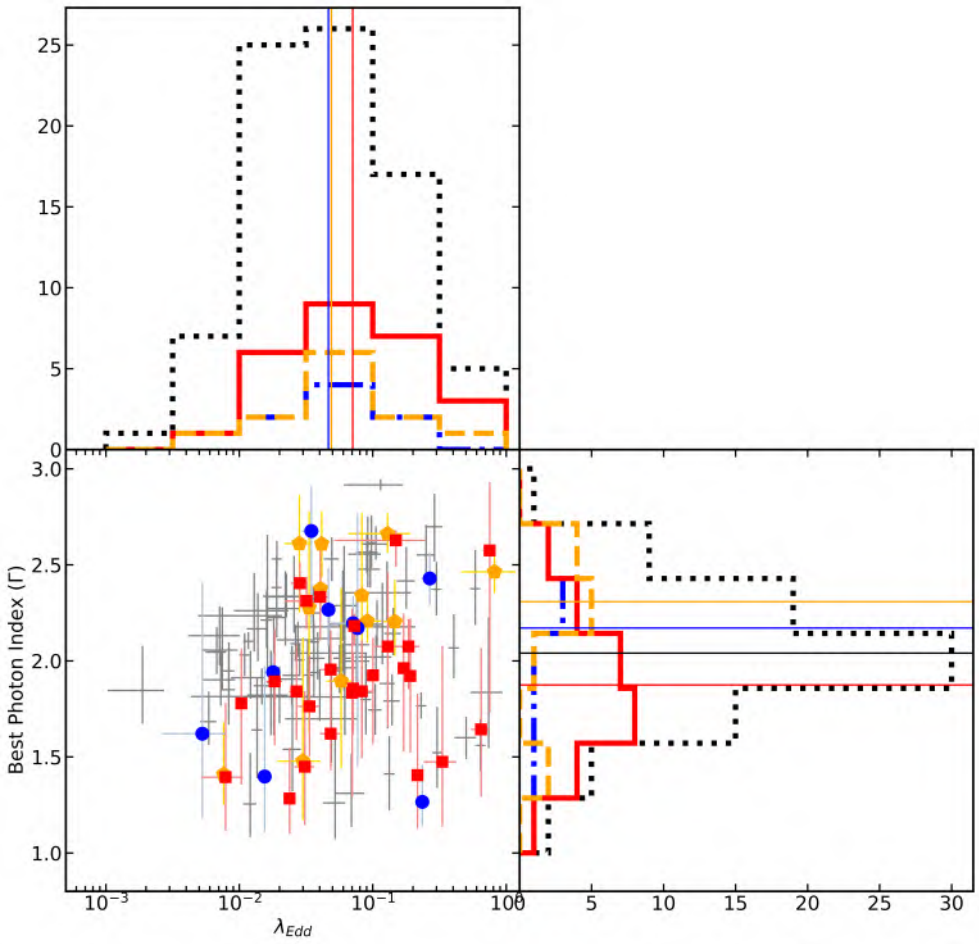

Abbildung 3.30: Distribution of Eddington ratios and photon indices measured from the best fitting spectral model. Sources shown in grey are best fit with a power law, sources shown with red squares are best fit with a soft excess, sources shown in blue circles are best fit with warm absorbers, and sources shown as orange pentagons are best fit with partial covering. Histograms for each parameter are also shown, and median values are indicated with solid lines in the corresponding colours.

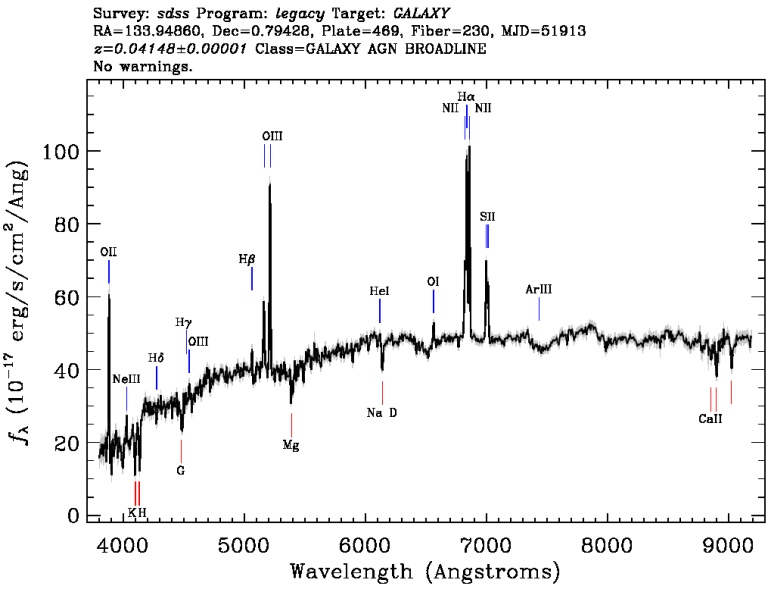

Abbildung 3.31: An example of an AGN (ID 04444) with strong host galaxy contamination. A blind fit of this spectrum met the criteria for a well measured black hole mass using the Hβ line, but visual inspection reveals no evidence for a broad Hβ component.

ponent. Such FWHM measurements are erroneously large, resulting in a large BH mass and low Eddington ratio.

One class of AGN for which the masses can be improperly measured is in Seyfert 1.9 galaxies (e.g. Osterbrock 1981), sources which show some or weak evidence for broad Hα but no evidence for other broad lines including Hβ. A physical explanation for this is that the source is moderately obscured. As a result, the optical spectrum appears reddened, including the broad wings of the Hβ line, while broad Hα remains visible. In the X-rays, Seyfert 1.9 galaxies show evidence for higher obscuring column densities than other type 1 AGN, more similar to type 2 (Koss et al. 2017), further supporting this physical picture. Because of the reddening leading to a more clearly visible host galaxy spectrum and a lack of broad Hβ, these sources require more careful treatment in order to estimate their masses, which is beyond the scope of this chapter.

In order to establish which sources have reliable black hole mass measurements, we perform the visual inspection as described in Sect. 2.4.3 of this thesis, removing all sources where the broad Hβ, MgII or CIV was not visibly apparent. We are then left with a sample of 119 sources with good black hole masses and well constrained Eddington ratios. Dividing the sources based on the best fit soft X-ray model, we are left with 8 sources best fit with a warm absorber model, 26 best with with a soft excess, 10 best fit with a neutral partial covering absorber, and the remaining 75 which prefer a power law model. A majority of the sources removed from the analysis were best fit with either a power law or a warm absorber model.

In Waddell et al. (2024), using a KS-test, no significant differences were found in the bolometric luminosities or black hole masses between sources best models. However, the distributions of FWHM of the optical broad lines (bottom left), as well as the distributions of Eddington ratios (bottom right), were found to differ significantly. Sources best fit with a warm absorber tended to show lower Eddington ratios than sources best fit with a soft excess model, with high statistical significance (Anderson-Darling test p-value 0.0059). This result was found to be robust to a variety of selection effects. However, after the removal of some sources following visual inspection, there is no longer a sufficiently large number of sources to perform statistical test comparing the warm absorber and soft excess subsamples, and the results are not significant. Nevertheless median value of the Eddington ratio remains lower for sources best with a warm absorber. Another difference is that the analysis in this chapter shows that sources with a soft excess may show evidence for lower FWHM and black hole masses than sources best fit with a power law; this result is of interest and is discussed in Section 3.8.5.

Fig. 3.29 shows an increase in the fraction of sources which are best fit with a soft excess model with increasing Eddington ratio, while warm absorbers are not found in the highest Eddington ratio bin. This is only partly in agreement with the analysis of Waddell et al. (2024), which reports the same behaviour and trends for sources best fit with a soft excess, but finds that the fraction of sources best with a warm absorber tended to decrease with increasing Eddington ratio. The difference observed is again due to several sources with the lowest Eddington ratios and preferring the warm absorber model being removed as a result of the updated visual inspections. Further observations and black hole mass measurements, particularly for sources which show strong evidence for warm absorbers, would be needed to verify if indeed sources best fit with warm absorbers are more likely to be found in lower Eddington ratio sources.

3.8 Discussion

3.8.1 Soft X-ray spectral properties of eFEDS AGN

The eFEDS survey covering 140 deg^2 of the sky was designed to demonstrate the capabilities of eROSITA for extragalactic survey science, in anticipation of the all sky survey. Even with a very small fraction of the final survey area, eFEDS has demonstrated the power of eROSITA not only to detect new sources and source populations, but to characterise their properties and provide physical insights via their X-ray spectra. This chapter uses those capabilities to identify AGN with a soft excess, and to differentiate between various models for the excess, which may be an artefact of complex absorption or a true emission component. More specifically, this chapter uses Bayesian fitting methods and reliable model selection to characterise the X-ray spectral of a sample of 200 hard X-ray–selected AGN, finding the following results: .

- In addition to the underlying continuum modelled with a power law, an additional power law describes the shape of the soft excess better than a blackbody component.

This suggests a non-thermal origin for the soft excess is preferred over a blackbody originating in the inner disc.

- Making use of simulations and spectral fitting, 29 (14.5%) sources show evidence for a soft excess, 29 (14.5%) sources show evidence for a warm absorber, and 25 (12.5%) sources show evidence for a partial covering absorber, all with 2.5% (~ 5) spurious detections estimated from simulations. By design, there is no overlap between these samples.

- Examining these sources in colour/flux space reveals differences between sources best fit with soft excesses, complex absorbers, and only power laws, which can be used for comparison with soft X-ray–selected samples and the eROSITA all-sky survey.

- Of the 29 sources in the selected soft excess sample, 23 appear best fit with the soft Comptonisation interpretation, and six appear best fit with blurred reflection.

- Many sources which display evidence for Compton-thin absorbers ($> 10^{22}\,\mathrm{cm}^{-2}$) from the baseline absorbed power law model are actually warm absorbers, including all sources with column densities $> 10^{23}\,\mathrm{cm}^{-2}$, suggesting that a population of apparently absorbed type-1 AGN actually display evidence for warm absorption from e.g. a disc wind rather than absorption in the distant torus.

These findings and the consequences for the all-sky survey are further discussed in the following sections.

3.8.2 Soft excesses in eFEDS

The sub-sample of 29/200 sources with soft excesses identified in this chapter prove a powerful tool for studying soft excesses in AGN. First, it should be noted that this fraction of sources with soft excesses is low, with less than half (83 sources, or 41.5%) of sources showing evidence for some form of additional component (soft excess or complex absorption). Furthermore, only 29 sources (14.5% of the full sample) show evidence for a soft excess which cannot be explained with complex absorption alone. Previous works studying samples of soft excesses typically assume that all sources have a soft excess and apply the models accordingly, and indeed some works have claimed ubiquitous or near-ubiquitous detections (Piconcelli et al. 2005; Done et al. 2012; Scott et al. 2012), while in this chapter it is demonstrated that little or no information can be gained from fitting a soft excess component to sources without a soft excess. Soft excess measurements are not biased by the selection of the hard X-ray sample (Fig. 3.16), and that while sources with soft excess are typically higher fluxes, there are also very bright sources with no evidence for a soft component. It remains unclear if this chapter finds a lower number of AGN with soft excess due to low signal-to-noise for some data sets or overly stringent selection criteria, or if this is reflective of the true fraction of sources with soft excesses. Further investigation is needed to properly estimate this fraction.

While this chapter has found that most sources are best fit or at least well fit with soft Comptonisation, there are six sources which show some preference for blurred reflection. Sources best fit with blurred reflection have higher soft Γ values when fitting with the PL+PL model (see Fig. 3.20). Part of this may be due to modelling biases or uncertainties, however, this may imply that sources best fit with blurred reflection have steeper soft excesses. Furthermore, it is found that sources best fit with blurred reflection have a higher median soft excess strength, and high reflection fraction values. This suggests that the sources with extremely strong reflection spectra are still identified as best fit blurred reflection, while sources with weaker measured reflection spectra are in fact unlikely to be best fit with blurred reflection, and the soft excess is more likely produced with a warm corona. This may be a limitation of the eROSITA bandpass, where emission above 5 keV is very difficult to detect due to the combination of low effective area and high background. While the soft components can have a similar shape, reflection models predict a more prominent and relativistically broadened line at ~ 6 keV due to fluorescent emission from Fe Kα, which may differ from the more limited reprocessing in the warm corona model.

Strong and steep soft excesses are seen in a majority of X-ray observed NLS1 galaxies (e.g. Boller et al. 1996). Individually, many NLS1s have been shown to be well fit with blurred reflection, absorption, or a combination of the two (e.g. Tanaka et al. 2004; Fabian et al. 2009; Gallo et al. 2015, 2019; Jiang et al. 2019; Waddell et al. 2019; Boller et al. 2021). On aggregate, NLS1s also appear to have stronger soft excess strengths, a strong correlation between the soft excess strength and the strength of the Compton hump, and a weak anti-correlation between the photon index and soft excess strength (Waddell & Gallo 2020). Together, this may suggest that some AGN display NLS1-type spectra, and that these sources are more likely fit with blurred reflection models compared to more typical broad-line Seyfert 1 (BLS1) sources with flatter, smoother and weaker soft excesses, which are well fit with the soft Comptonisation model (Waddell & Gallo 2020).

It is also noteworthy that while the warm corona model fits most sources in this sample, other works have clearly demonstrated the presence of broad lines and deep absorption edges associated with Fe Kα emission and absorption. These features cannot easily be explained with the soft Comptonisation model; rather, this requires an absorption component, blurred reflection, or some combination of models. In this chapter, sources with a very strong reflection component (e.g. $R \gg 1$) were typically best fit with a relativistic blurred reflection model, suggesting that sources with very strong reflection components are preferentially detected. Further investigation with a larger sample (e.g. eRASS:1) as well as follow-ups with high resolution spectroscopy (e.g. XRISM; Tashiro et al. 2020) will help to identify more sources with high energy spectral curvature and verify if there are sub-samples of AGN which require blurred reflection to explain the spectral shape.

One caveat to this analysis is that this analysis does not attempt to characterise properties of potential warm absorbers and disc winds simultaneously. To address this, all sources which showed evidence for a soft excess or warm absorber were also fit with a model including both a disc wind (warmabs) and a secondary power law component (in XSPEC, this corresponds to `tbabs` $\times$ `ztbabs` $\times$ cwa18 $\times$ (`powerlaw` + `powerlaw`)). The same priors as in Sect. 3.3 were used, and Bayes factors were again calculated for each

source. For most sources, the inclusion of both components did not result in a better fit to the spectrum, and many spectral parameters could not be well constrained. Therefore, it appears that for the eFEDS analysis, the components can only be treated separately, which may lead to some modelling errors. Laha et al. (2013) discuss the complications of simultaneously modelling a soft excess and warm absorption component in NLS1 IRAS 13349+2438, and demonstrate that different ionising continua produce different ionisation structures in the warm absorption component, whereas including a soft excess tends to decrease the ionisation while the column density remains consistent. More recently, Parker et al. (2022) discuss the degeneracies between X-ray winds and relativistic blurred reflection, where systematic biases are found in the derived outflow parameters of the wind when the emission from the disc is not properly characterised. These are particularly crucial in the case of broad features, where even microcalorimeter resolution does not help to break some degeneracies (Parker et al. 2022). Care should therefore be taken in interpreting the ionisation parameters in this analysis, and further analysis should be done to better understand any potential superposition between these components.

3.8.3 The warm corona model

Despite the caveats discussed above, the evidence comparison for individual sources shows that most are well fit or best fit with a warm corona model (Table 3.8). The combined evidence comparison also shows that if one assumes that the underlying model for all sources is the same, then the warm corona model is preferred over blurred reflection.

The warm corona model tends to produce fairly flat hot corona photon indices, with a median value of ~ 1.6. This is flatter than the typical population values of $\Gamma \sim 1.8 - 1.9$. Interestingly, these flatter photon indices are not found when examining the PL+PL model (see the left-hand panel of Fig. 3.20). The hard X-ray photon index is correlated with the detection likelihood in the $2.3 - 5\,\mathrm{keV}$ band such that sources with very flat photon indices are also very close to the background in the hard X-ray, which may explain why flatter slopes are measured. However, the fact that the PL+PL modelling returns more reasonable slope values of $\Gamma \sim 1.85$ may suggest that the stricter priors placed on the soft photon index may also be artificially flattening the hard X-ray photon index. This may also be a result of the combined shape of the two `nthComp` components, in particular when considering the steep cut-off present in the soft power law: indeed, Xu et al. (2021) also find a very flat slope when fitting the type-1 AGN ESO 362-G18 with a double `nthComp` model as compared to modelling with blurred reflection. These findings suggest that the model parameters obtained in this chapter are acceptable for this analysis.

One limitation to modelling the warm corona using the method outlined in this chapter is that it fails to characterise the effect one corona has on the emission from the other. In a more realistic scenario, photons from the warm corona would be incident on the hot corona, and vice versa. It could even be plausible that the hard and soft components are part of the same, multi-temperature and multi-density cloud of electrons. Therefore, it is interesting to investigate evidence for interplay between the two coronae. One way to do this is by understanding the soft excess strength, that is the ratio of the fluxes emitted

by each of the two coronae in the same energy band. Examining the distribution of soft excesses (Fig. 3.18), most of the soft excesses span a relatively small parameter space of factor ~ 10, with few extreme values. Interestingly, using a similar definition for the soft excess strength and modelling broad-line Seyfert 1 galaxies observed with *Suzaku*, Waddell & Gallo (2020) found that soft excess strengths span about factor ~ 10. By contrast, the observed soft excess strengths in NLS1 galaxies spanned a much larger parameter space (factor ~ 100), extending to very strong soft excess strengths (~ 10), similar to the extreme values found for some sources in this chapter. This may suggest some coupling between the two coronae which only permits certain ratios of fluxes between the two components, while this restriction does not exist for blurred reflection dominated sources, where the flux ratio between components is more dependant on the height of the corona above the accretion disc.

3.8.4 Dense absorption in eFEDS may be warm and complex

When modelling eFEDS sources with the baseline absorbed power law model, many sources appear to show evidence for Compton-thin absorption with column densities $> 10^{22}\,\mathrm{cm}^{-2}$. However, examining these sources more closely, a majority (19/26) show evidence for complex absorption (18 warm absorbers and one partial covering absorber), as seen in Fig. 3.24. This includes all eight sources with column densities above $10^{23}\,\mathrm{cm}^{-2}$, as measured by the baseline model, which are all best fit by a warm absorber model. This raises an intriguing possibility that some, if not most, of the apparently Compton-thin absorbers observed in the Universe are associated with ionised gas hosted in disc winds, and not necessarily in a more distant obscurer such as the torus.

In sources with complex absorbers, while the cold and warm absorption column densities are not individually well constrained, the sums of the simple (`ztbabs`) and complex (cwa18 or `zpcfabs`) absorbers are constrained (see also Fig. 3.5, Fig. 3.8, Fig. 3.12 and Fig. 3.26.) Fig. 3.32 shows the distributions of total column density for each model. Contrary to Fig. 3.24, the definition of the column density changes depending on the best-fit model for each source; for sources best fit with a single power law or soft excess, this corresponds to the column density measured in the `ztbabs` component. For sources best fit with a warm absorber, this corresponds to the sum of the column densities from the `ztbabs` model and the cwa18 grid, and for sources best fit with neutral partial covering, this corresponds to the sum of the column densities measured from the `ztbabs` and `zpcfabs` components. A majority of sources best fit with a power law have low column densities ($< 10^{21}\,\mathrm{cm}^{-2}$). This can be associated with absorption in the host galaxy, with a small minority showing evidence for some form of Compton-thin obscuration. However, the majority of absorbed sources are obscured by a partial covering absorber or a warm absorber, and thus not a distant torus. Indeed, no significant difference is found between soft excess and a power law, demonstrating that most of the absorption originates in the partial covering of warm absorbers.

To investigate this in more detail, the 19 apparently obscured sources best fit with complex absorption are re-fit with a more physical torus model, `UXClumpy` (Buchner et al.

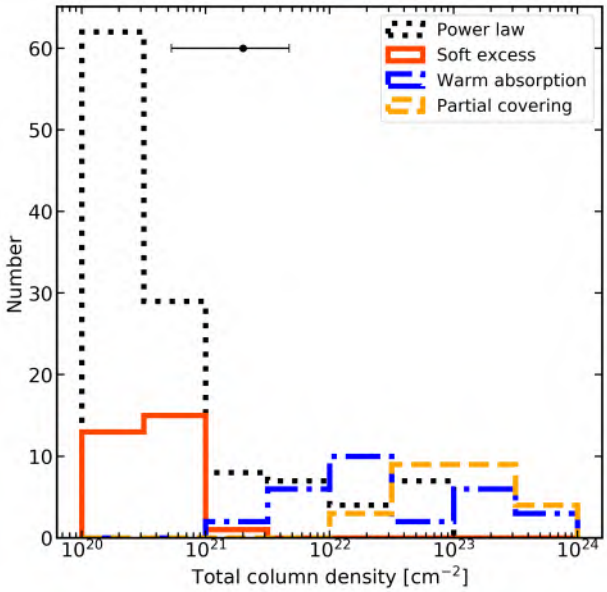

Abbildung 3.32: Distribution of total absorbing column density for each model. For sources best fit with a single power law or soft excess, this corresponds to the column density measured in the `ztbabs` component. For sources best fit with a warm absorber, this corresponds to the sum of the column densities from the `ztbabs` model and the XSTAR grid, and for sources best fit with neutral partial covering, this corresponds to the sum of the column densities measured from the `ztbabs` and zpcfabs components. Sources best fit with soft excesses are shown in red, sources best fit with warm absorbers are shown as blue dash-dot lines, and sources best fit with partial covering are shown as orange dashed lines. The typical error bar is shown in black.

2019), which describes the X-ray spectrum of a clumpy absorber illuminated by the corona. The Galactic column density is also included in the model, as before. Free parameters include the coronal photon index (given a uniform prior between one and three, as in all other models in this chapter), the line-of-sight column density (given a log-uniform prior between $10^{20}\,\mathrm{cm}^{-2}$ and $10^{25}\,\mathrm{cm}^{-2}$), and the torus inclination, vertical extent of the clouds, and covering factor, all of which were given uniform priors between the minimum and maximum values allowed by the model and none of which were well constrained in spectral fits. The energy cut-off is fixed at 400 keV, but does not influence the spectral shape for large cut-off values given the limited hard X-ray coverage of eROSITA. The Bayes factor is then computed for each source to ease comparison with other models.

Of these 19 sources, only six are better fit with `UXClumpy`. In all of these cases the column density in the torus is $\sim 10^{23}\,\mathrm{cm}^{-2}$, and other torus parameters cannot be constrained. Furthermore, very steep photon indices of $\Gamma \sim 2-3$, with many in agreement with the upper limit of $\Gamma = 3$, are required to explain the spectral shape. This seems highly unlikely, as many of these sources are expected to be type-1 AGN do not typically show evidence for absorption with column densities $> 10^{22}\,\mathrm{cm}^{-2}$ (e.g. Shimizu et al. 2018), and this is at odds with the AGN unification model wherein type-1 AGN offer a direct view of the central region and only type-2 AGN are viewed through the torus (e.g. Antonucci 1993; Urry & Padovani 1995). Furthermore, the high photon indices are likely un-physical and are atypical for obscured AGN studied with high quality spectra (e.g. Ricci et al. 2017). Therefore, the warm absorber model, which yields more reasonable photon indices of $\Gamma \sim 1.8-2$ is preferred for these sources. Future modelling with higher signal to noise spectra or expanded energy coverage (e.g. with *NuSTAR*) would be required to confirm this result (Waddell et al. in prep.).

If indeed the majority of Compton-thin obscuration in this sample is caused by warm absorption, this may change our view of the way obscuration is considered and treated in AGN. If these warm absorbers are indeed physically produced by winds launched from the accretion disc, these winds are believed to play a significant role in the evolution of the system. In a comprehensive work on absorption in AGN, Buchner et al. (2015) studied the evolution of AGN which appear unobscured (column densities $< 10^{22}\,\mathrm{cm}^{-2}$), Compton-thin (column densities $10^{22}-10^{24}\,\mathrm{cm}^{-2}$), and Compton-thick (column densities $> 10^{24}\,\mathrm{cm}^{-2}$), and found some evidence that Compton-thin AGN are evolving faster at redshifts $z = 0.5-4$ (specifically, their space density rises more rapidly and peaks earlier than unobscured and Compton-thick AGN). This is most apparent for sources with column densities $10^{22}-10^{23.5}\,\mathrm{cm}^{-2}$, where a majority of the sources in this chapter (15/23) show evidence for warm absorption. This might suggest that disc winds are playing a role in the more rapid space density evolution of apparently Compton-thin sources in this redshift range. A more in-depth analysis incorporating more sources spanning a larger redshift range would be required to more fully understand this result.

3.8.5 The Eddington ratio distinction between spectral models

A major difference between this chapter and the analysis published in Waddell et al. (2024) is in optical spectral analysis, where this chapter includes an update reliable black hole mass sample based on visual inspections presented in (Nandra et al. 2024). Since a number of the sources which were removed from the analysis showed strong evidence for warm absorbers, this lead to a statistically significant difference between the measured Eddington ratios of sources best fit with warm absorbers versus soft excesses, where sources best fit with a warm absorber tended to be hosted in lower Eddington ratio AGN, while sources best fit with a soft excess tended to be found in higher Eddington ratio sources. Repeating this analysis in this chapter while removing sources with dubious black hole mass estimates lead to a drop in the statistical significance of this result, and therefore the result remains uncertain and further testing is needed to confirm if the differences are physical.

Using the updated subsample of sources with visually inspected black hole masses, we find the same increase of the fraction of sources which are best fit with a soft excess with increasing Eddington ratio, in agreement with Waddell et al. (2024) However, in this Chapter we find a relatively constant fraction of sources which show evidence for warm absorbers across low Eddington ratios, with a sharp drop and then cut off above $\lambda_{\rm Edd} \sim 0.01 - 0.03$, with no sources having higher $\lambda_{\rm Edd}$ values of $\sim 0.03 - 1$. The analysis also reveals a statistically significant (KS-test p-value 0.015) difference between the black hole mass distributions of sources best fit with a power law and those best with with a soft excess, where sources best fit with a soft excess were also found in lower mass black holes. Examining the FWHM distributions shown in Fig. 3.28, it is revealed that a number of these sources have low FWHM, including 5 which have FWHM$<$ 2000 km s^{-1}, consistent with NLS1 classification. Indeed, 4/5 sources with FWHM$<$ 2000 km s^{-1} and half of sources with FWHM$<$ 3000 km s^{-1} show evidence for a soft excess. As discussed in Sect. 3.8.2 of this thesis, NLS1 galaxies also tend to show strong and steep soft excesses. Further testing would be necessary to determine if disc winds are more frequently over-ionised in NLS1s, perhaps due to their extreme variability, thus more frequently collapsing to form a soft excess, or whether another mechanism is responsible for this effect.

3.9 Lessons learned from spectral fitting with eFEDS

In this Chapter, the 200 sources classified as AGN from the eFEDS hard X-ray–selected sample (Chapter 2) are modelled with a variety of phenomenological and physically motivated models in order to investigate the nature of the X-ray soft excess. X-ray spectra are fit using BXA so that the Bayesian evidence can be compared to select the best fitting models. This Chapter demonstrates that eROSITA can be used to identify signatures of both complex absorption and soft excesses, and using simulations, the significance of these features can be evaluated. This analysis identifies a total of 29 sources that have warm absorbers (14.5% of the sample), 25 sources that have neutral partial covering absorbers (12.5% of the sample), and 29 sources (14.5% of the sample) with soft excesses (all

with 97.5% purity), which clearly shows that soft excesses and complex absorbers are key features for understanding the properties of large samples of AGN.

It is shown that most sources with true soft excesses are best explained by a warm corona model as opposed to a relativistic blurred reflection scenario. Follow-up observations of these sources with better sensitivity in the hard X-ray (e.g. with *XMM* or *NuSTAR*) can help to search for the presence of a broad iron line or Compton hump which are strong signatures of blurred reflection (e.g. Fabian et al. 1989; Ross & Fabian 2005), and a timing analysis incorporating reverberation mapping (e.g. following the prescription of Uttley et al. 2014) can also help to distinguish between soft excess and absorption models. Using the results from this Chapter and repeating this analysis using the all-sky survey (e.g. eRASS:1, eRASS:4) will provide a large sample of $\sim$ 1000s of AGN for enhanced analysis (see Chapter 4). Furthermore, using data with very high spectral resolution (e.g. XRISM and Athena) will help to confirm these results, and to create a more complete picture of the nature of warm absorbers and soft excesses in AGN in the local Universe.

Kapitel 4

eRASS:1 Hard X-ray selected Active Galactic Nuclei

The eROSITA instrument aboard the SRG satellite performed its first all-sky survey between December 2019 and June 2020. This Chapter presents the resulting 5466 sources in the hard X-ray ($2.3 - 5\,\mathrm{keV}$) sample, the first created from an all-sky imaging survey at these X-ray energies, for sources within western galactic sky (eROSITA_DE). I apply the lessons learned from our detailed analysis of the eFEDS data, presented in Chapters 2 and 3, to find counterparts for the X-ray sources, to produce a large uniform sample of hard-X-ray selected AGN, and to characterise them with supporting multi-wavelength astrometry, photometry and spectroscopy. From here, the primary objective is to create a rich catalog of hard X-ray selected AGN with spectroscopic redshifts, which is presented in chapter 4.8. I also perform comparisons with the Swift BAT sample and HEAO-1 AGN sample to attempt to better understand the effectiveness and sensitivity of eROSITA in the hard band. Comparing with other hard X-ray selected surveys, the eROSITA hard sample covers a larger redshift range and probes dimmer sources, providing a complementary and expanded sample as compared to Swift-BAT. A sample of hard-only sources, many of which are likely to be heavily obscured AGN, is also presented and discussed. X-ray spectral fitting reveals that these sources have extremely faint soft X-ray emission and their optical images suggest that they are found in more edge-on galaxies with lower b/a; these sources will be explored with more detail in Chapter 5.

This work has been re-submitted to A&A (Waddell et al. resubmitted), accompanied by catalogs of counterparts in different sky areas. Scripts, methodology and analysis tools to clean the X-ray catalog were developed by co-authors, and the catalog was first presented in Merloni et al. (2024). As the lead author, I was responsible for application of the tools to the hard X-ray selected sample, as well as testing and verification of the results.

4.1 Hard X-ray selected samples in the eROSITA era

The first eROSITA all-sky survey (eRASS1; Merloni et al. 2024) yielded a large sample of $\sim$ 1 million X-ray selected sources, $\sim$ 80% of which are expected to be AGN. Although eROSITA lacks sensitivity above 8 keV, it is highly sensitive to soft X-ray emission, enabling accurate modelling of the obscuring column densities. As we learned from eFEDS (see Chapter 2), by studying the hard X-ray selected sample, we are able to constrain the coronal photon index which dominates the X-ray emission above 2 keV. Furthermore, with excellent soft X-ray spectral resolution, eROSITA spectra can be fit with physically motivated models to better understand the nature of soft X-ray emission and absorption components (Chapter 3). Selecting hard-X-ray detected sources which lack soft X-ray emission can also allow for the identification of obscured AGN.

Numerous previous surveys making use of X-ray telescopes with sensitivity above 2 keV have been conducted to detect obscured AGN, as outlined further in Chapter 1. First, the non-imaging telescope HEAO-1 (Rothschild et al. 1979) surveyed the entire sky in the $0.2 - 25$ keV band, creating a bright sample ($> 3 \times 10^{-11}$ ergs s^{-1} cm^{-2}) of hard X-ray AGN presented in Piccinotti et al. (1982), which have been re-examined in many later works. Subsequent hard X-ray surveys have used imaging telescopes covering smaller and/or non-contiguous areas of the sky; for example, using *ASCA* (Ueda et al. 1999), *BeppoSAX* (Della Ceca et al. 2008), *Chandra* (e.g. Nandra et al. 2015; Civano et al. 2016; Liu et al. 2017) and *XMM-Newton* (e.g. Fiore et al. 2003; Cocchia et al. 2007; Hasinger et al. 2007b). These missions are mostly sensitive below 10 keV; complementary, higher energy missions, including NuSTAR ($3 - 79$ keV; Harrison et al. 2013) and *Swift* BAT ($14 - 195$ keV; Gehrels et al. 2004), have been employed to detected heavily obscured, Compton-thick AGN (Akylas et al. 2016; Marchesi et al. 2016). *Swift* BAT has performed an all-sky survey in the ultra-hard X-ray band, with the most recent 105 month release (Oh et al. 2018) detecting 1632 X-ray sources, including at least 1105 AGN. Many of these sources have been followed up with NuSTAR, whose improved energy sensitivity and softer X-ray coverage has enabled spectral modelling to compare torus models (Ikeda et al. 2009; Murphy & Yaqoob 2009; Brightman & Nandra 2011; Baloković et al. 2018, 2019; Buchner et al. 2019). As a precursor to the eROSITA all-sky surveys, eFEDS (Brunner et al. 2022) provided key insights as to the science capabilites of eROSITA surveys; this sample was analysed in detail in Chapters 2 and 3, and the lessons learned from that analysis will be crucial to the understanding of the eRASS1 data.

In this Chapter, I present the counterparts, classifications and an analysis of the X-ray and optical properties of the 5466 hard X-ray selected sources in the first all-sky survey. I focus on X-ray point sources which have not been flagged as potentially spurious or problematic, with a particular focus on the identification of AGN. In Sect. 4.2, I provide an expansion of the X-ray properties of the catalog described in Merloni et al. (2024). In Sect. 4.3, information on the counterparts within the DESI imaging legacy surveys data release 10 (Dey et al. 2019) is described, spectroscopic redshifts are compiled, where available, sources are classified, and beamed AGN are flagged. In Sect. 4.4, I briefly describe AGN matched from outside the Legacy survey. Section 4.5 presents a comparison to the *Swift*

BAT survey and chapter 4.6 compares to the HEAO-1 AGN sample. Section 4.7 briefly discusses variable sources, and in particular, variable AGN. The spectroscopic AGN sample is presented in sect. 4.8, with a particular focus on hard-only AGN. Further discussion is given in sect. 4.9, and conclusions are given in sect. 4.10.

4.2 Properties of the eRASS1 hard sample

4.2.1 Population information

The hard (2.3−5 keV) X-ray selected catalog from the eRASS1 all sky survey in the western galactic sky (eROSITA_DE) is presented in Merloni et al. (2024), and the X-ray positions of all sources are shown in Fig 4.1. Relevant additional sub-samples are also shown in this plot, which will be described in the following chapters.

The catalog contains a total of 5466 sources, selected to have hard-band detection likelihoods of DET_LIKE_3> 12, where band 3 is 2.3 − 5 keV (Merloni et al. 2024). Based on simulations performed in Seppi et al. (2022), this threshold implies a spurious fraction of 10%, as can be seen from Fig. 4.2 (top). The background or spurious sources are shown in grey, the point-like simulated sources are shown as blue dashed lines, and extended simulated sources are shown as light blue dash-dot lines. The spurious fraction is strongly dependent on the detection likelihood; it is around 10% at DET_LIKE_3 > 12, dropping to 2.5% around DET_LIKE_3 > 15, and nearly 0 at a DET_LIKE_3 > 18. For point-like sources, Fig. 4.2 (bottom) shows the same for hard-only sources; these will be further discussed in section 4.2.2.

Out of the 5466 sources, 5087 sources are considered to be point-like in the X-ray, having a EXT_LIKE (extended likelihood) of 0. These are sources where the extend of the source does not exceed the PSF of eROSITA. The other 379 sources are candidate extended sources in the X-ray. These are discussed in more detail in chapter 4.2.3.

A number of flags indicating spurious or otherwise problematic sources are presented in Merloni et al. (2024), Table 5. These include five flags for potentially spurious sources, identified due to their proximity to SNR, bright point sources, stellar clusters, large local galaxies, known galaxy clusters, or sources potentially contaminated with optical loading. A further three flags identify sources for which no error bars were provided on the coordinate errors, extent errors, or source count errors. While these sources are not necessarily spurious, they are removed for analysis as these parameters are necessary for sample selection and counterpart identification, and cannot be trusted for flagged sources. When these flagged sources are removed, a total of 5020 sources remain, 4895 of which are point-like.

A key difference between the main and hard X-ray samples is the distribution of the positional uncertainty (at the 1σ level), contained in the column POS_ERR. These can be seen in Fig. 4.3, where sources in the main sample are shown in grey, all sources (including those flagged as spurious) in this hard sample are shown in black, non-flagged hard-only sources are shown in orange, and non-flagged sources in the hard sample with soft detections are shown in blue. The astrometric calibration (Merloni et al. 2024) imposes a lower bound

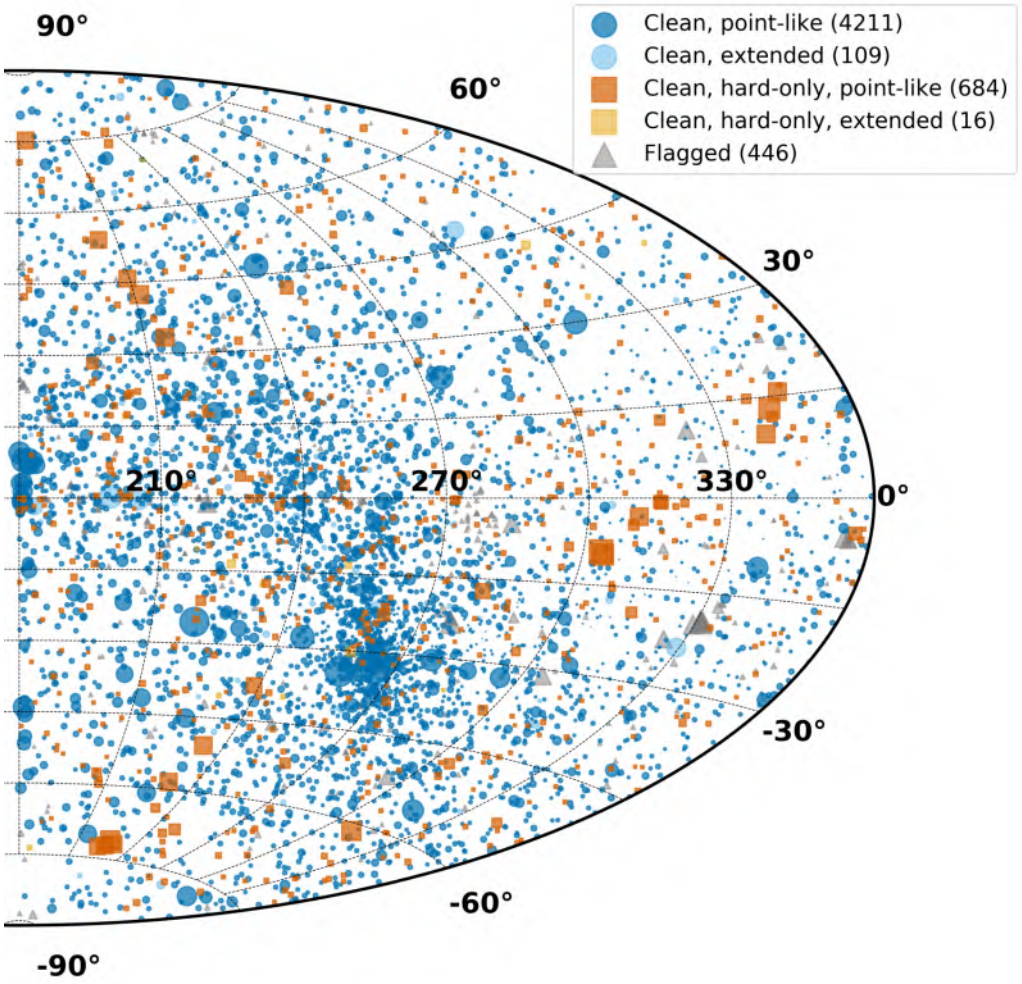

Abbildung 4.1: Hard X-ray detected sources in the first all-sky survey. The non-flagged, point-like sample is shown as blue circles, and non-flagged extended sources are shown as pale blue circles. The non-flagged, point-like, hard-only sample is shown with orange squares, and the hard-only extended sources are shown with light orange squares. Sources flagged as problematic are shown as grey triangles. The size of the shapes corresponds to the hard band ($2.3-5\,\mathrm{keV}$) flux, with larger shapes corresponding to brighter X-ray fluxes. There are a total of 5466 sources.

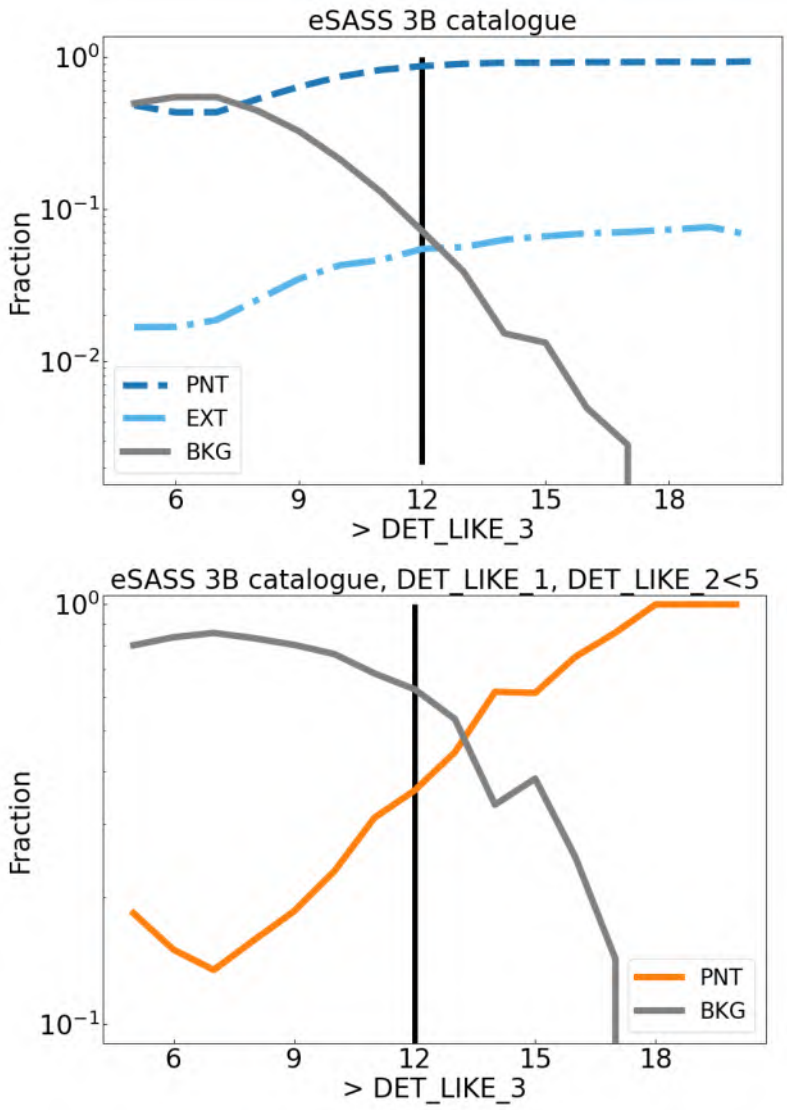

Abbildung 4.2: Top: Distribution of population fractions for simulated sources versus `DET_LIKE_3`. Simulated point sources are shown as blue dashed lines and labelled PNT, simulated extended sources are shown as light blue dash-dot lines and labelled EXT, and spurious sources (e.g. not real sources in the simulations) are shown as a solid grey line, labelled BKG. Bottom: Distribution of population fractions for simulated hard-only sources with respect to `DET_LIKE_3`. Simulated point sources are shown in orange, and background sources are shown as a solid grey line, labelled BKG. A solid black line at `DET_LIKE_3` of 12, indicating the threshold for a significant hard band detection, is shown.

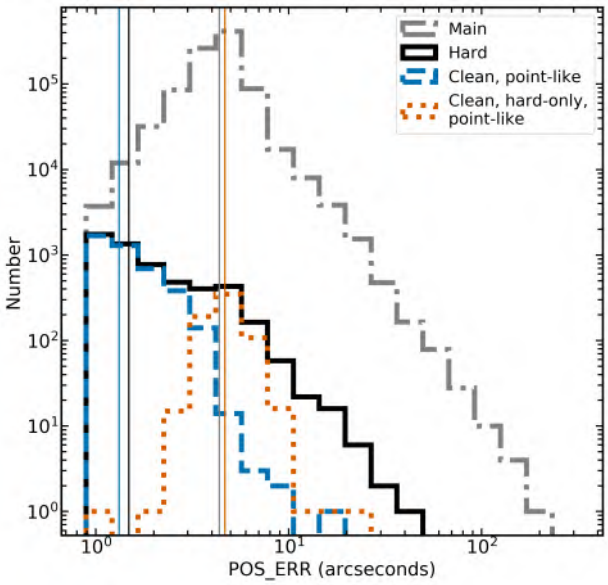

Abbildung 4.3: Distribution of errors on the coordinates (`POS_ERR`) for individual sources. The eRASS1 main sample is shown with a grey, dash-dot line. The hard sample is shown as a black solid line, and is further sub-divided into non-flagged, point-like hard sources, and point-like hard-only sources, in a blue dashed line and an orange dotted line, respectively. Median values are indicated with vertical lines in corresponding colours.

of 0.9" in the calculation of POS_ERR. In the main sample, the peak of the distribution is found at $\sim$ 4.5", whereas for the hard sample it is found around 1.5". This is because the hard sources are typically brighter in the soft band, and both the soft and hard X-ray positional information is used to localize the sources, reducing the positional errors compared to sources only detected in one X-ray band. This also affects the counterpart identification procedure, since preference will be given to counterparts within the positional error compared to those at a larger radius. There is also a difference observed for the hard-only sources; since their positions are only determined by the hard X-rays, the positional errors are larger and the distribution more closely follows that of the main sample. Many sources flagged as spurious also have large positional errors, and are only presented in the black curve; some sources are also missing positional errors and are not shown here. In contrast, most of the sources with very small positional errors are real, bright sources with excellent positional accuracy. The reason the histogram drops off sharply at low `POS_ERR` is that many sources have very small positional errors, so the peak of the distribution should be at or below the imposed minimum of 0.9".

4.2.2 Hard-only sources

In this chapter, hard-only sources are defined as those that are undetected in the softer bands, specifically defined by:

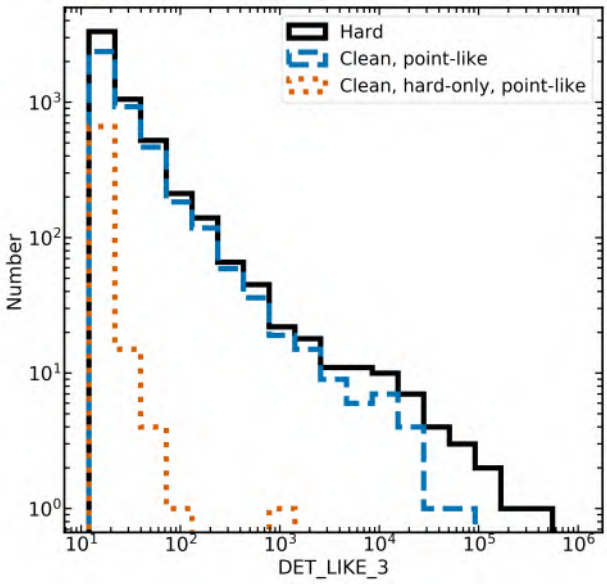

Abbildung 4.4: Distribution of DET_LIKE_3 ($2.3-5$ keV) values. The hard sample is shown as a black solid line, and is further sub-divided into point-like sources, and point-like hard-only sources, in a blue dashed line and an orange dotted line, respectively.

$$\text{DET\_LIKE\_1} < 5 \;\&\; \text{DET\_LIKE\_2} < 5 \tag{4.1}$$

where 1 indicates the $0.2-0.6$ keV band and 2 indicates the $0.6-2.3$ keV band. The same threshold of DET_LIKE_3>12 is also applied. In total, 738 sources meet these criteria. The hard only sources are distributed across the sky, with many being in or near the galactic plane, as distant soft X-ray emission in this region is partly obscured by the Milky Way disc. They have higher positional errors than other sources in the hard sample, complicating counterpart identification. There is also a higher expected spurious fraction among the hard-only sample, as shown in Fig. 4.2 (bottom), closer to $\sim 65\%$ at DET_LIKE_3 of 12, but falling off to about 30% at a DET_LIKE_3 of 15 and to 0 above a DET_LIKE_3 of 17. This would imply that $\sim 480/738$ hard-only sources are spurious, leaving ~ 260 which are expected to be real astrophysical sources. This is similar to the result found in Chapter 2, where it was also found that a larger fraction of hard only sources were spurious, and also that the hard-only sources had poorer counterpart quality. Stricter cuts could also be made on DET_LIKE_3 in order to select a purer, but possibly less complete, sample. Fig. 4.4 shows the distribution of DET_LIKE_3. In general, the hard-only sources have lower DET_LIKE_3 values, which is sensible as a large fraction of these source are believed to be spurious, based on the simulations shown in Fig. 4.2, which is more likely at lower DET_LIKE_3.

Examining only the point-like hard-only sources, there are 684 meeting the definition in Eq. 1. This definition differs from that presented in Merloni et al. (2024), which relies on spatial coincidence of sources in the main and hard catalogs (criteria for a weak match) and a tolerance placed on the ratio of their counts (criteria for a strong match). Sources

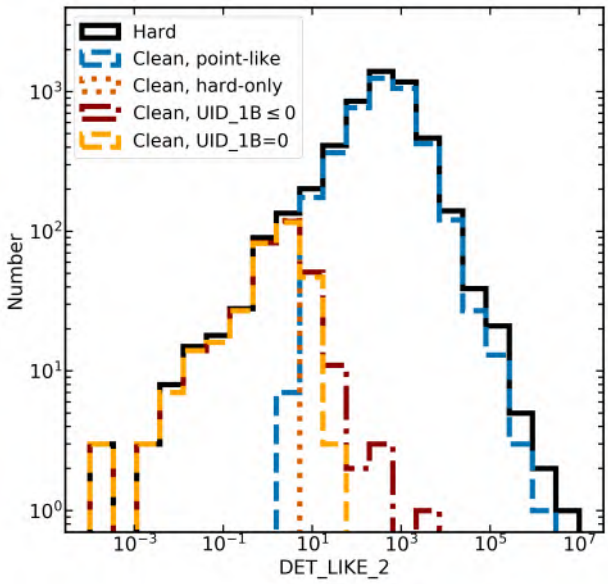

Abbildung 4.5: Distribution of DET_LIKE_2 ($0.6 - 2.3\,\mathrm{keV}$) values. The hard sample is shown as a black solid line, and is further sub-divided into non-flagged, point-like hard sources, and point-like hard-only sources, as the blue dashed line and the orange dotted line, respectively. The hard-only samples as defined in Merloni et al. (2024) are also shown, with the higher fidelity, UID_1B=0 sample shown as a pale orange dashed line, and the less pure, UID_1B$\leq$ 0 sample shown as a dark red, dash-dot line.

which are weak matches have a negative value in the UID_1B column in the hard X-ray catalog, and those no matches have a 0 in the UID_1B column. Selecting the X-ray point-like sources, 729 have no matches, and 752 have no or weak matches. However, examining the distribution of fluxes and detection likelihoods in band 2 ($0.6 - 2.3\,\mathrm{keV}$; see Fig. 4.5), there are some sources with high detection likelihoods of tens to thousands. It is unlikely that these sources are good candidates for heavily obscured AGN, and they are not comparable to the hard-only AGN presented and discussed in eFEDS in Chapters 2 and 3. By using the modified definition of Eq. 1, all of these sources are removed. Out of the 684 hard-only sources from this work, 673 are also hard-only using the definition in Merloni et al. (2024), demonstrating very high overlap between the two selections.

Also of interest are the sources for which the softest band has a significant detection likelihood, but the intermediate band does not; that is, having DET_LIKE_1 > 5 & DET_LIKE_2 < 5. There are five such sources, all with DET_LIKE_1 < 9. Four of these sources are in the LS10 area, and none have any flag in the X-ray. Such sources may be candidate obscured AGN, with significant soft X-ray leakage due to Compton scattering or an ionised absorption medium.

4.2.3 Extended sources

There are a total of 379 sources in the hard extended sample, peaking at an extent likelihood of EXT_LIKE $\sim 140 - 150$. This is very different from the extent likelihood distribution of the main sample, which peaks at EXT_LIKE $\sim 3 - 4$. This is because the hard sample primarily selects the brightest sources in the eROSITA X-ray sky, and such bright extended sources are more readily identified. Of these 379 candidate extended sources, many are flagged as lying near a known cluster or SNR, and are thus likely spurious. Removing these, a total of 125 sources remain. Of these, 16 have been identified as being hard-only sources which satisfy equation 1. All but one of the remaining sources are found in the main catalog and will be described in Bulbul et al. (2024) and Kluge et al. (2024). Of the 16 hard-only extended sources, one is positionally coincident with a source in the main catalog, and only three have sufficiently high extent likelihoods (EXT_LIKE > 9) to make it into the secure cluster catalog (Bulbul et al. 2024). Two sources are in the galactic plane, and the remaining source is spatially coincident with a known galaxy, LEDA 483209. It is therefore likely that, aside from LEDA 483209, these sources are most likely spurious rather than being true heavily obscured clusters. For the remainder of this work, the focus will be restricted to the 5087 point-like X-ray sources.

4.3 Counterparts in the legacy survey footprint

4.3.1 Counterpart identification

To attempt to identify the counterparts of the eRASS1 Hard sample, the photometric and other data provided in the DESI imaging legacy surveys, specifically the Legacy Survey Data Release 10 (LS10; Dey et al. 2019) are used. In addition to g-, r-, i- and z- band photometry, LS10 also includes WISE fluxes, obtained from forced-flux photometry on the unWISE data at the positions of the Legacy Survey optical positions. LS10 is the first of the Legacy Survey data releases to include i-band photometry, and coverage is improved in the Southern Hemisphere, coincident with a large portion of the eROSITA-DE sky. In all, 15342 deg^2 of the sky are covered by at least one pass in all four photometry bands. This is defined using a multi-order coverage map (MOC), and is the sky area covered by all those 0.25x0.25deg legacy survey 'bricks' which have at least 1000 detections, and which have estimated 5 sigma point-source depths of at least g<23.5, r<23.3, i<22.8 and z<22.3, in AB magnitudes. Of the 4895 X-ray point-like sources without flags, 2863 are in the legacy survey area. Out of these, 363 are hard-only sources. These numbers and other key values of interest are listed in Table 4.1.

Counterpart identification is performed using `NWAY` (Salvato et al. 2018, 2022), a Bayesian framework which has been specifically designed to identify counterparts for astronomical sources. This method uses positional information, positional errors, and the local X-ray source density as well as the multiwavelength properties to assess the likelihood of an optical source being an X-ray emitter as well as being the correct counterpart. The training is performed on a 4XMM sample cut at a $0.5 - 2$ keV flux of 2×10^{-15} ergs s^{-1} cm^{-2},

Tabelle 4.1: Useful reference numbers for the different samples listed throughout this work. Good `p_any` is defined as `p_any`>0.033 for sources with both soft and hard X-ray detections, and `p_any`>0.061 for the hard-only sources. We caution that the hard-only sub-samples suffer from higher contamination from spurious X-ray sources, as shown in Fig. 4.2; further cuts on `DET_LIKE_3` or `p_any` could be made to improve upon this.

Category	No. of sources
Hard sample (`DET_LIKE_3` > 12)	5466
Point-like sources	5087
Clean, point-like sources (CPL)	4895
Hard-only CPL sources	684
CPL sources in the LS10 area	2863
Hard-only CPL in LS10	363
CPL soft sources outside the LS10 area	1710
Hard-only outside LS10	321
CPL extragalactic (spec-z>0.002)	1538
CPL Galactic	325
CPL in LS10 with good `p_any`	2547
Hard-only good `p_any`	111
Good `p_any` with spec-z>0.002 (non-beamed)	1243
Hard-only good `p_any` with spec-z>0.002 (non-beamed)	23
Stars	325
Beamed AGN	319
BAT matches	487
CPL BAT matches	456
Other AGN outside LS10	692
Final spec-z sample (Sect. 4.8.1)	1328
Final spec-z hard-only	29

which corresponds to the minimum flux of the hard sample. More details of the training are described in Salvato et al. (2022, Salvato et al. (in prep.)), in which the same method and training priors are applied to the main eRASS1 catalog. Some caveats to this will be discussed in chapter 4.9.1.

Starting from the entire hard X-ray catalog, the extended sources were removed. `NWAY` was then run on the sample to find the X-ray counterparts. LS10 sources within a 60 arcsecond radius of the X-ray source were considered, in order to account for the few X-ray sources with very large positional error (maximum $\sim$ 50 arcseconds), some of which are hard-only sources which are of particular interest in this chapter. Out of the 2863 point-like, non-problematic X-ray sources in the LS10 area, 2850 sources have a primary counterpart (match_flag=1).

4.3.2 Sources without a counterpart

Since 2850 out of 2863 sources have at least one counterpart, this leaves a mere 13 sources without a LS10 counterpart. These sources are listed in Table 4.2. To attempt to identify these sources, I examined the images in the legacy survey viewer tool [1], and query their positions in SIMBAD (Wenger et al. 2000) using a two arc minute match radius. This extremely large search radius allows us to check if there are any known sources in the general area of the X-ray detections, but this is not used to assign counterparts. Notes on the results of these searches can be found in Table 4.2.

In total, three sources are clearly associated with bright AGN which may have issues with their photometry in LS10, one source is coincident with a bright star, one source is coincident with a cluster of galaxies, one source appears in a gap in LS10 which is not reflected in the MOC, and one source has no previously named nearby sources. A further six sources are all found in the South Ecliptic Pole (SEP) area (Merloni et al. 2024). There are issues with the shallow depth or missing photometric bands in LS10 in this area, as well as issues with source confusion given the very high density of X-ray sources due to the repeated exposures. The repeated exposures also mean that lower flux sources are probed. Therefore, it is reasonable that some sources are missing counterparts there. The `DET_LIKE_3` values are high for most sources without LS10 counterparts, so it is likely that most are real X-ray sources.

4.3.3 Calibration of `p_any` thresholds

In order to assess the quality of the counterpart match, the `p_any` values produced by `NWAY` are used. For each entry in the X-ray catalogue, the `p_any` is related to the probability that there is a counterpart. To assign thresholds to these values, a mock X-ray catalog is produced, where the RA and Dec of the X-ray sources have each been shifted by two arcminutes. This was done by adding 2 arcminutes to the RA and Dec of each source, except where this would push the source outside the eROSITALS10 area, where 2 arcminutes

[1] https://www.legacysurvey.org/viewer

Tabelle 4.2: Non-flagged, point sources without counterparts in LS10. The eROSITA name is given in column (1), and the X-ray RA and Dec are given in columns (2) and (3). Column (4) gives the hard band detection likelihood (DET_LIKE_3), and column (5) lists notes for each source. Coincident indicates < 5 arcsecond match radius, while nearby is defined as < 2 arcminutes.

(1) eROSITA Name	(2) RA	(3) Dec	(4) DET_LIKE_3	(5) Notes
1eRASS J015513.1-450612	28.8050	-45.1035	17.56	Coincident with bright quasar QSO J0155-4506 (z=0.451)
1eRASS J033336.3-360825	53.4013	-36.1403	969.98	Coincident with Sy1 NGC 1365 (z=0.00554)
1eRASS J041524.0-232234	63.8501	-23.3763	15.20	Coincident with QSO [VV2006] J041524.0-232234 (z=0.6163)
1eRASS J041619.1-480412	64.0796	-48.0700	18.69	No known nearby objects in SIMBAD
1eRASS J041722.9-474838	64.3452	-47.8107	19.87	Coincident with galaxy cluster SPT-CL J0417-4748
1eRASS J053232,8-655141	83.1373	-65.8614	75.67	Hole in Legacy Survey coverage
1eRASS J055909.1-663206	89.7878	-66.5351	20.27	In SEP
1eRASS J055927.9-662925	89.8662	-66.4904	45.51	In SEP
1eRASS J055934.2-653831	89.8927	-65.6421	17.59	In SEP
1eRASS J060008.9-662505	90.0371	-66.4182	40.40	In SEP
1eRASS J060012.4-662731	90.0517	-66.4587	19.42	In SEP
1eRASS J060015.5-663231	90.0645	-66.5421	119.46	In SEP
1eRASS J185305.9-501051	283.2746	-50.1810	20.52	Coincident with star V* PZ Tel

was instead subtracted. Sources at these new shifted positions which are coincident with an original X-ray source were then removed. In this way, it is possible to evaluate the incidence of chance alignment in the NWAY cross-matching procedure by comparing the p_any distributions of real matches and those from the mock catalog.

The (normalized) distributions of the p_any values obtained from the real and shifted samples are shown in Fig. 4.6. Only point sources with no flags in the X-ray are considered. For the shifted sources, only a very small minority of sources have large p_any, showing that NWAY rarely finds a secure counterpart where none should be found. There is a very high peak at ~ 0, with most sources having very low p_any values. In contrast, the p_any values of the real sources are more broadly distributed, with the largest peak at a p_any of 1.

To define a p_any threshold to use to define a secure counterpart, purity and completeness curves are constructed. The completeness is defined as the fraction of sources which have a p_any derived from matching the true X-ray positions greater than a specified value, whereas purity is defined as the fraction of mock sources which have p_any greater than a given threshold. These are shown in Fig. 4.7, with, the hard-only sources shown in orange, and the sources which also have soft detections (e.g. those which are not hard-only) are shown in blue. Completeness curves are shown as solid lines, and purity curves are shown with dashed lines. To achieve a purity of 90%, a threshold of p_any > 0.033 should be used for the hard sample, and p_any > 0.061 for the hard-only sample, indicated with horizontal lines in corresponding colours. The hard sample has a completeness of $\sim 98\%$ at this purity, and the hard-only sample has a much lower completeness of $\sim 30\%$. This is

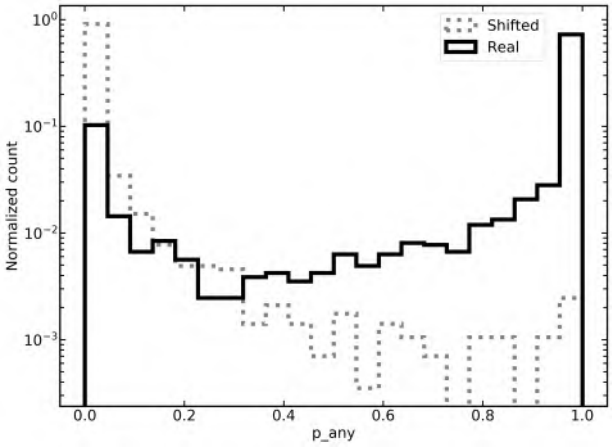

Abbildung 4.6: Distribution of `p_any` thresholds for the real (black, solid line) and shifted (gray, dotted line) samples.

consistent with the much higher expected spurious fraction of the hard-only sample, where up to 65% of sources are expected to be spurious (see chapter 4.2.2).

The dependence on the counterpart quality as a function of detection likelihood (`DET_LIKE_3`) is shown in Fig. 4.8. We include hard-only sources in all curves provided they meet the detection thresholds. Three different thresholds are shown: the percentage of X-ray detected sources are expected to be spurious is 10% for `DET_LIKE_3`>12, 2.5% for `DET_LIKE_3`>15, and << 1% and for `DET_LIKE_3`>18 (see Seppi et al. 2022; Merloni et al. 2024, chapter 4.2.1). The purity and threshold curves are much higher when a higher `DET_LIKE_3` cut is applied. These cuts also effectively remove all the hard-only sources, which typically have lower `DET_LIKE_3` (e.g. Fig. 4.4). A highly pure, complete and very securely associated sample could be selected using higher `DET_LIKE_3` and `p_any` cuts.

The `p_any` values with respect to the `DET_LIKE_3` values for the hard-only (orange squares) sources are shown in Fig. 4.9, alongside the entire sample of sources with counterparts. A `DET_LIKE_3` of 18 is indicated with a black solid line, above which X-ray detections have a very low (much less than 1%) chance of being spurious. A `p_any` of 0.061 is also shown with a black dashed line, which indicates the 90% purity cut on `p_any` for the hard-only sources. Predictably, many of the low `DET_LIKE_3` sources also have low `p_any`, probably because many are spurious. Most sources with `DET_LIKE_3` larger than 18 have secure counterparts. Of those which do not, a majority are located in the SEP, where both X-ray and LS10 counterparts are problematic due to the patchy coverage and low depth of the LS10 survey and the potential source confusion and dim fluxes probed in X-rays, and as such likely the results should not be trusted (see Liu et al. in prep).

Applying these findings, we apply the 90% purity cuts to our samples, meaning that 10% of sources are spurious associations between X-ray and optical counterparts. Therefore, we define sources with a good counterpart as having `p_any`>0.033, and hard-only sources with good counterparts as having `p_any`>0.061, selecting a total of 2547 sources (of which 111 are hard-only sources) with good counterparts. These definitions are used in the following

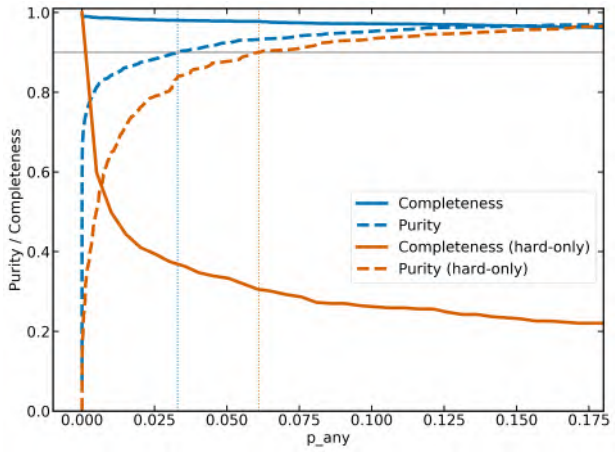

Abbildung 4.7: Purity and completeness as a function of `p_any` values, defined as in Salvato et al. (2022). Purity and completeness for the sample with soft X-ray detections are shown as blue solid and dashed lines, respectively, and purity and completeness for the hard-only sample are shown as orange solid and dashed lines, respectively. The horizontal grey line shows a purity of 0.9, and vertical dotted lines in corresponding colours indicate the `p_any` values for the 90% purity cuts.

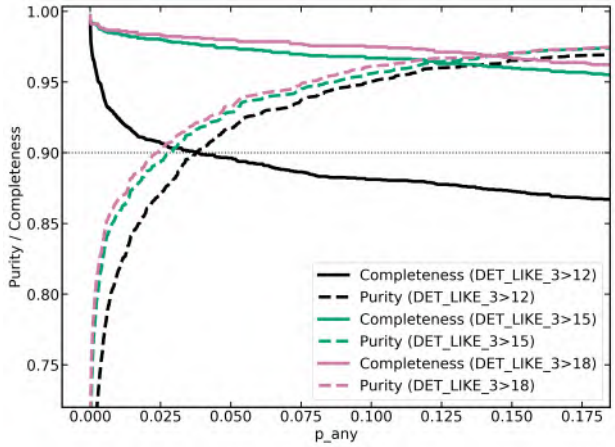

Abbildung 4.8: Purity (dashed lines) and completeness (solid lines) as a function of `p_any` values, defined as in Salvato et al. (2022). The black, green and pink lines show sub-samples of sources with `DET_LIKE_3`>12, `DET_LIKE_3`>15 and `DET_LIKE_3`>18, respectively.

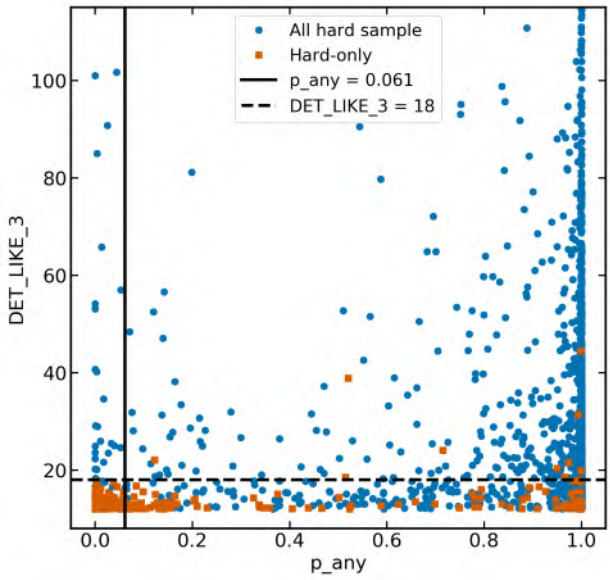

Abbildung 4.9: Distribution of `DET_LIKE_3` values as a function of `p_any`, shown for both the hard-only sample (orange) and the remainder of the hard sample (blue). The vertical black line indicates the 90% purity cut on `p_any` for the hard-only sample, and the horizontal dashed line shows a high-fidelity cut of `DET_LIKE_3`>18.

sections, Chapter 5 and Table 4.1.

If the errors on the X-ray coordinates and the matching algorithm are accurate, then the the angular separations of the LS10 counterparts from their eROSITA X-ray counterparts, normalised by the position errors, are expected to be modelled by a Rayleigh distribution. This is tested in Fig. 4.10, and indeed, the true distribution is in close agreement with the Rayleigh distribution.

4.3.4 Spectroscopic redshifts

Following the method of Ider Chitham et al. (2020), a compilation of spectroscopic redshifts was constructed from a large variety of literature sources. Specific catalogs, queries and selection criteria are further described in Appendix D of Kluge et al. (2024).

The spectroscopic catalog includes the RA and Dec taken from the Legacy Survey sources, which were matched within one arcsecond to the LS10 position of the counterpart to the hard X-ray source. The separations are typically less than 0.3 arcseconds, suggesting that the spec-z are indeed being reported for the LS10 sources associated with the X-ray counterparts. 711 sources with good `p_any` have a spec-z which is available in this catalog, 706 of which have $z > 0.002$ and can be considered extragalactic sources. Then, to further increase the number of spec-z, we also match our optical counterparts to the spec-z in NED using a one arcsecond match radius. This match yields an additional 335 NED sources, 327 with extragalactic spec-z values of $z > 0.002$.

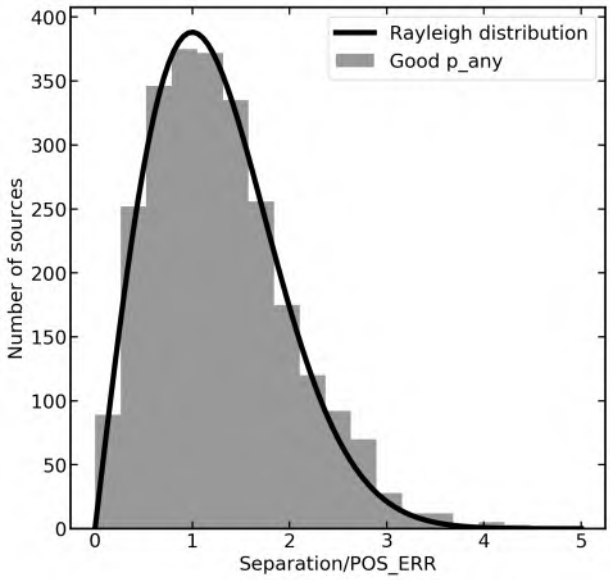

Abbildung 4.10: Distribution of separations between the optical and X-ray positions, divided by the X-ray positional errors. The sources with good `p_any` are shown in grey, and the expected Rayleigh distribution is shown with a solid black curve.

In addition to these spec-z values, we take advantage of the inclusion of the Gaia DR3 information in the LS10 catalog to also include spec-z values from Quaia (Storey-Fisher et al. 2023), a publicly available Gaia–unWISE Quasar Catalog. Spec-z are measured using the low-resolution Gaia spectrometer, and are improved using WISE selection criteria. The final sample has magnitudes $G > 20.5$ and shows good agreement with SDSS spec-z measurements (Storey-Fisher et al. 2023). Matching directly by Gaia DR3 ID and considering sources which do not have a spec-z from the compilation or NED, 505 sources with good `p_any` have a spec-z, all of which have $z > 0.002$ are can be considered extragalactic sources, as expected from a quasar catalog. In total, this gives 1538 sources, including 23 hard-only sources, with a good counterpart and a spec-z of $z > 0.002$. The redshift distribution from each spectroscopic catalog, as well as the total distribution, is shown in Fig. 4.11.

The agreement between the redshifts from the three catalogues used has been verified by computing the outlier fraction, where an extreme outlier is defined as:

$$|z_2 - z_1|/(1 + z_1) > 0.15 \tag{4.2}$$

where z_1 is the better-constrained redshift and z_2 is the lesser-constrained value. This definition is commonly used in order to assess the goodness of redshifts as well as to identify outliers, including catastrophic outliers (e.g. Salvato et al. 2009, 2011). Comparing NED to the spectrocopic redshift catalog, the outlier fraction is very small, at 1.5%. Comparing the spectroscopic redshift catalog with Quaia, the outlier fraction is higher at $\sim$ 17%. This is likely related to the low resolution of the Gaia spectrograph, and the apparent lack of sources in their low redshift ($z < 0.1$) SDSS comparison sample. Some of these

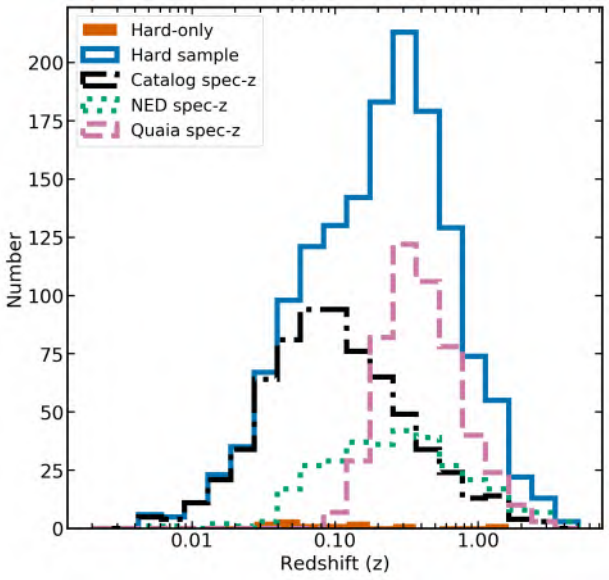

Abbildung 4.11: Distributions of spectroscopic redshifts for the eROSITA hard sample. The hard sample is shown as a blue solid line, and the hard-only sample is shown with an orange filled histogram. Spec-z from the spectroscopic redshift compilation of Kluge et al. (2024) are shown as black dash-dot lines, and NED spec-z are shown as green dotted line. The Quaia redshifts are shown with pink dashed lines.

extreme outliers are removed when beamed sources removed (see description in chapter 4.3.6, decreasing the outlier fraction to 13%). A majority of the outliers are at $z > 0.3$, so over interpretation of these sources is cautioned.

4.3.5 Source classification

Learning from eFEDS, a similar source classification methodology is followed to Salvato et al. (2022). First, all sources which have a spectroscopic redshift of $z > 0.002$ are classified as extragalactic (1536 sources), while those with a spec-z of $z < 0.002$ are classified as Galactic. Additionally, sources which have a parallax which is significant at the $> 5\sigma$ level are classified as stars, selecting a total of 325 unique Galactic sources.

Where no spectroscopic redshift is available, or the parallax is not significant, the optical morphology of the source is examined to see if it is extended. This is done by looking at the TYPE column in the LS10 catalog. If the type is not equal to PSF, then the source is believed to be optically extended, and is thus likely to be a galaxy rather than a star. This allows us to classify 660 sources as likely extragalactic. Indeed, $\sim 50\%$ of the spectroscopic extragalactic sample is optically extended, while none of the Galactic sources are, further justifying this classification.

For remaining sources (339) without a spec-z, significant parallax, and which are not flagged as optically extended, we apply colour-colour cuts similar to Salvato et al. (2022).

The relevant colour-colour plots are shown in Fig. 4.12, with different sub-samples indicated with different colours and symbols, and equations used for cuts are listed in the legend. Only sources with type PSF in LS10 and sources where the signal-to-noise ratio > 3 in all photometry bands are plotted.

The top-left panel of Fig. 4.12 shows the $z - W1$ and $g - r$ colours for the hard sample. The line of separation has the same slope as prescribed by Salvato et al. (2022), but with an intercept of -0.2, rather than -1.2 reported in that work. This is because Salvato et al. (2022) included many local galaxies and clusters in this diagnostic, which are removed in our work by selecting only point-like sources in the X-ray and optical. We investigate the three outliers with large $z - w1$ values.Two are located < 10" from known stars (V* HQ Eri and G 41-14B), one is ~ 5" from a known galaxy NGC 1097 at a redshift of 0.004, and all have low or intermediate `p_any` values, suggesting that there may be issues with the counterparts for these sources. Despite these outliers, the new selection is very pure, with 99% of spectroscopic extragalactic sources lying in the upper region of the plot. The $W1$ vs. soft X-ray diagnostic from Salvato et al. (2022) shown in the right-hand corner of Fig. 4.12 is similarly pure. Here, the soft X-ray flux is computed using the hard $(2.3 - 5\,\mathrm{keV})$ X-ray flux which is available even for the hard-only sources, extrapolated using an absorbed power law with a photon index of 1.9 (e.g. Liu et al. 2022b; Nandra et al. 2024, Chapter 3) and a column density of $4 \times 10^{20}\,\mathrm{cm}^{-2}$ to represent a typical Galactic value. Here, we use the hard X-ray flux to estimate the intrinsic soft X-ray flux for all sources, since in our sample, the hard X-ray flux is consistently available, and again is estimated from the X-ray counts (see Merloni et al. 2024). Since we are not using this diagnostic plot to separate between obscured or unabsorbed AGN, but rather between X-ray bright AGN and X-ray faint stars, we should estimate the intrinsic flux of the AGN, rather than the observed soft X-ray flux. Since most sources have a photon index close to 1.9, using this to estimate intrinsic flux is likely to be appropriate for most sources, even obscured ones, since the underlying spectral shape is believed to be similar. We caution that these values are only rough estimates of the intrinsic soft X-ray flux for the hard-only sources, and suggest against over-interpreting the classification of hard-only sources without spectroscopic redshifts.

The bottom-left corner of Fig. 4.12 combines the two upper figures to create the final source classification. Sources which have both colours in this plot > 0 are classified as having extragalactic colours and plotted as unfilled, dark purple diamonds, and those with either colour < 0 are classified as having Galactic colours and are plotted as unfilled, dark red stars. In the same plot we also mark sources classified as Cataclysmic variables (CV) in SIMBAD (unfilled black squares). It can be seen that many stars which lie in the extragalactic colours region are in fact CV's, where the X-ray emission may arise from the accretion process onto the white dwarf.

Summarizing, to classify the hard band selected sources with optical counterparts, we first remove 220 sources where one or both colours could not be computed, mostly due to issues with the W1 photometry, especially in the SEP region, or where sources do not have sufficient signal-to-noise ratio in these photometric bands. Applying the colour cuts leaves 96 sources which are classified as likely extragalactic, and 23 classified as likely Galactic. Using all classification methods 2340 sources are classified as likely extragalactic or secure

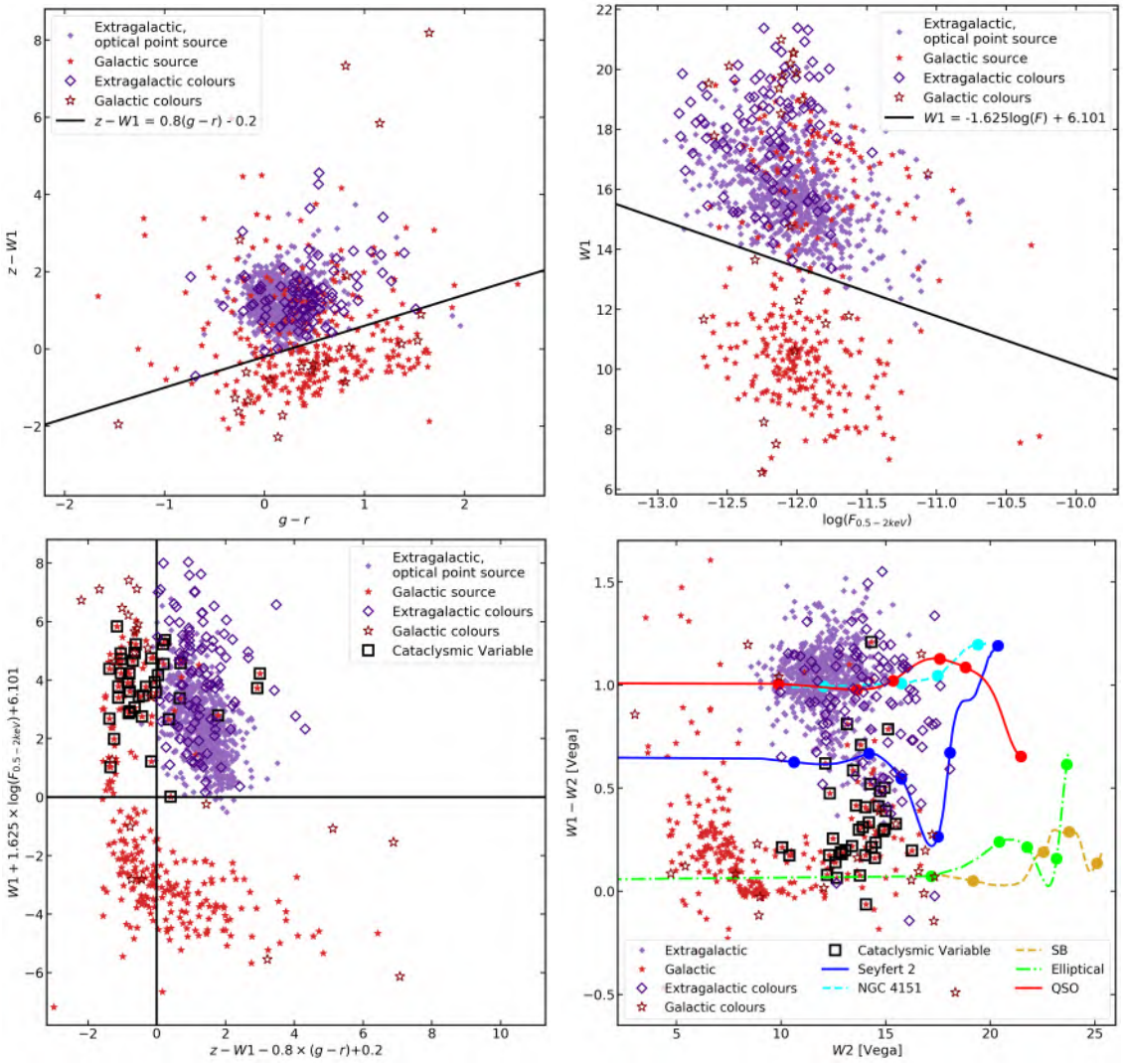

Abbildung 4.12: Colour-colour classification plots for the hard sample. Only sources with type PSF in LS10 and with a signal-to-noise ratio > 3 in all photometry bands are plotted. Sources classified as extragalactic (spec-z>0.002) are shown as filled purple diamonds, sources classified as Galactic (significant parallax at $> 5\sigma$ or spec-z<0.002) are shown as filled red stars, sources classified as extragalatic based on their colours are plotted as unfilled, dark purple diamonds, and sources classified as Galactic based on their colours are plotted as unfilled, dark red stars. These shapes and colours are labelled in the bottom-right panel. Cataclysmic variable (CV) systems are plotted with unfilled black squares on the bottom plots. The tracks on the bottom-right show the evolution of the colour-magnitude space occupied by different sources, with increasing redshift from left to right. Redshift ticks can be found at $z = 0.04, 0.2, 0.4, 0.92$, and where available, $z = 1.4, 3.8$.

extragalactic, and 362 are classified as likely or secure Galactic. Out of the 2547 sources with good `p_any`, 2114 are classified as likely extragalactic or extragalactic, and 337 are classified as likely or secure Galactic, indicating that $\sim 80\%$ of the sample is extragalactic.

We note that the classifications based only on colours should be interpreted with some caution, particularly for the stars, as some are located in a region of the parameter space (upper-left quadrant of Fig. 4.12) which is occupied by extended (and thus likely extragalactic) sources with Galactic-source-like colours. Furthermore, we again stress that sources with low `p_any` may have different astrophysical origins for the X-ray and W1 photons, making the W1-X-ray colour unreliable.

Finally, the bottom-right panel of Fig. 4.12 shows the $W1 - W2$ colour versus the W2 magnitude for these sources. Overplotted here are tracks corresponding to the colours of typical sources as they increase in redshift, akin to those presented in Salvato et al. (2022). A majority of extragalactic sources are shown to lie in the vicinity of the QSO and NGC 4151 (Seyfert 1 galaxy) tracks, with very few Galactic sources matching Elliptical and SB tracks. Also of interest is a smaller number of sources which appear to lie along the Seyfert 2 track, which may be obscured AGN. These sources will be further discussed in chapter 4.3.7.

4.3.6 Beamed AGN

Having defined a sample of secure and likely extragalactic sources, it is also of interest to identify which of these may be associated with blazars. The X-ray emission from these sources is likely to originate in the jet rather than in the X-ray corona, and may be relativistically boosted. This means that the total spectral shape and luminosity is unlikely to be representative of the corona, so removing these from an AGN sample will improve our understanding of the true coronal emission.

To identify candidate beamed AGN in our sample, we perform a cross-match with three catalogs/databases; CRATES, BZCAT and SIMBAD. Further investigation of all blazars in eRASS1, including a more detailed analysis of source types and X-ray spectral properties, will be presented in Haemmerich et al. (in prep); here we apply a simplistic approach to remove known beamed sources. CRATES (Healey et al. 2007) is an 8.4GHz selected flat-spectrum radio quasar (FSRQ) catalog constructed using multiple radio surveys and additional follow-up work to achieve nearly uniform sky coverage of latitudes $|b| > 10^{o}$. The Roma-BZCAT catalog is a multi-frequency blazar catalog, containing both FSRQ sources and BL Lac. The 5th edition was used here (Massaro et al. 2014), containing 3561 sources, all of which have radio detections. Finally, to identify other blazar candidates, we cross-match with SIMBAD. Sources which have a type of Blazar, BLLac, Blazar_Candidate, or BLLac_Candidate are flagged. The information is stored in the column class_beamed, where sources which appear in BZCAT have a class of 1, sources appearing in CRATES have a class 2, sources in SIMBAD have a class 3, and other sources have class 0. In total, 319 sources are flagged as beamed. These sources are removed from the main AGN sample, and will be presented in Haemmerich et al. (in prep).

4.3.7 Multiwavelength properties of the hard-only sources

Of particular interest in this thesis is the hard-only sample, which is likely to consist largely of obscured AGN which lack soft X-ray emission. However, it should also be noted that the hard-only sources may have properties which are not consistent with the remainder of the main sample. In particular, it is clear that many more hard-only sources are in fact spurious, and that even some sources which have a good counterpart may still be chance associations of galaxies and spurious hard X-ray detections.

One visualisation of this can be seen in Fig. 4.13, which shows the distribution of dereddened r magnitude (corrected only for Galactic absorption) and hard X-ray fluxes for the sample. The population of stars is evident, with $R < 12.5$ for typical X-ray fluxes. A very large population of hard-only sources lie in an extreme region of the parameter space, with relatively typical hard X-ray fluxes but extremely faint r band magnitudes. However, it can also be seen that many of these sources have low `p_any`. To further investigate this, we also plot a second sub-sample with a 99% purity cut on `p_any`, corresponding to `p_any` > 0.61, shown in dark red. Here it is clear that while some hard-only sources are fainter in r, other sources with $R > 22$ are likely wrong associations. Of the hard-only sources, 24 ($\sim 20\%$ of the hard-only sample) lie in the region of $R > 22$ and $\log($`ML_FLUX_3`$) \sim -12.25$. Given the high expected X-ray spurious fraction ($\sim 70\%$) of the hard-only sources, which is understood using simulations (see Fig. 4.2), this 20% fraction of incorrect identifications is not surprising. We do not make further cuts in`DET_LIKE_3` or p_any on this sample in order to void further reducing the completeness, however a purer sample could be obtained by making further cuts on `DET_LIKE_3`, p_any, or in optical magnitude (e.g. $R > 22$). Further investigation optical and X-ray flux ratios of these sources is left for future work, and combining data from future eROSITA all-sky surveys will help to reduce the fraction of spurious X-ray sources.

Also of interest is to examine the location of the hard-only sources in the $W1 - W2$ vs. $W2$ plane, and in particular in reference to the Seyfert 2 track. This is shown in Fig. 4.14. We note that this figure includes sources which have extended PSF types, and so includes additional sources as compared to the bottom-right panel of Fig. 4.12. Indeed, many hard-only sources are consistent with being Seyfert 2 AGN. Some are also consistent with stellar populations, and additional sources are also in agreement with being very bright QSOs, which is worthy of further exploration.

Additionally, the optical light profiles of sources which do not have an optical type of PSF are fitted in the LS10 catalog in order to measure the ellipticity of the source (Dey et al. 2019). For sources where this information is available, the ratio between the semi-major and semi-minor axes, a/b, is shown in Fig. 4.15. Hard-only sources are shown separately, and are binned up in order to better compare the distribution of sources to the rest of the hard sample. The median b/a for the sources with both hard and soft detections is 0.71, whereas for hard-only sources, it is 0.52. This could imply that the hard-only sources tend to preferentially found in galaxies which are edge-on. This interpretation will be discussed further in Sect. 4.8.2 and 4.9.2, and in greater detail in Chapter 5.

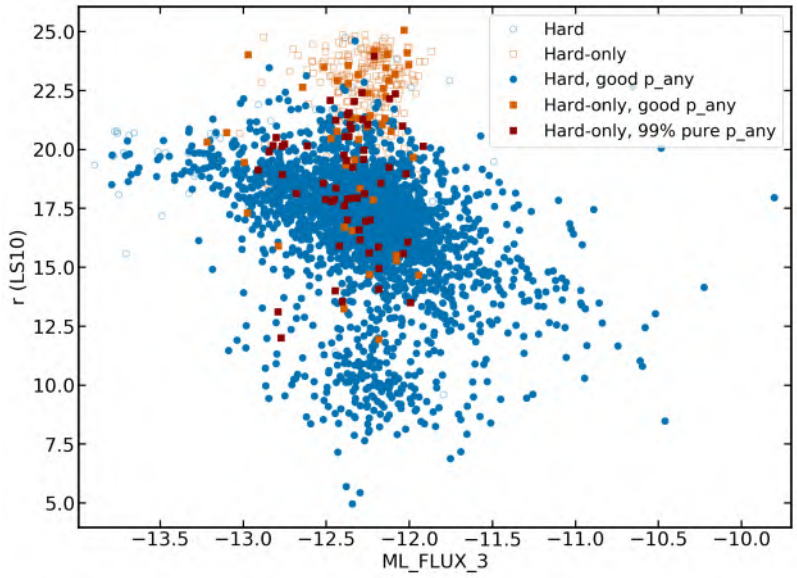

Abbildung 4.13: De-reddened r magnitudes and hard X-ray ($2.3-5\,\mathrm{keV}$) fluxes for the hard sample. All hard sources with `p_any` less than the threshold of 0.033 are shown as blue unfilled circles, and sources with good `p_any` are shown with blue filled circles. Hard-only sources with `p_any` less than the threshold of 0.061 are shown as orange unfilled squares, and hard-only sources with good `p_any` are shown with orange filled squares. A 99% purity cut on `p_any` at 0.61 is also applied, and these sources are shown as dark red filled squares.

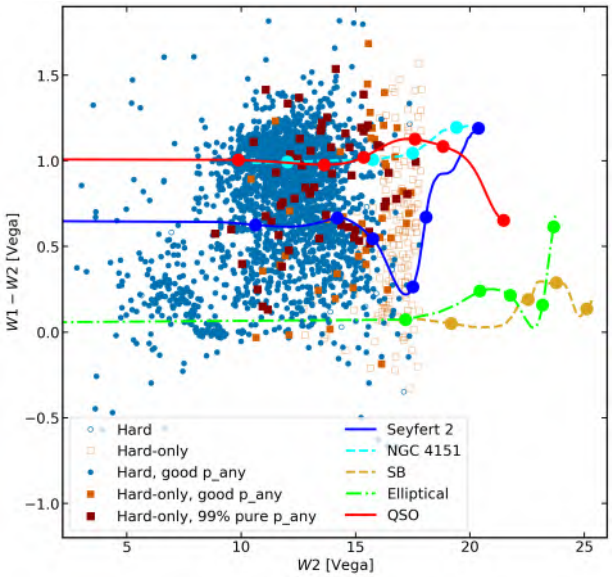

Abbildung 4.14: W1-W2 and W2 magnitudes shown for the eRASS1 hard sample. All hard sources with `p_any` less than the threshold of 0.033 are shown as blue unfilled circles, and sources with good `p_any` are shown with blue filled circles. Hard-only sources with `p_any` less than the threshold of 0.061 are shown as orange unfilled squares, and hard-only sources with good `p_any` are shown with orange filled squares. Tracks as in Salvato et al. (2022) and Fig. 4.12 of this work are shown in various linestyles and colours. Redshift ticks can be found at $z = 0.04, 0.2, 0.4, 0.92$, and where available, $z = 1.4, 3.8$.

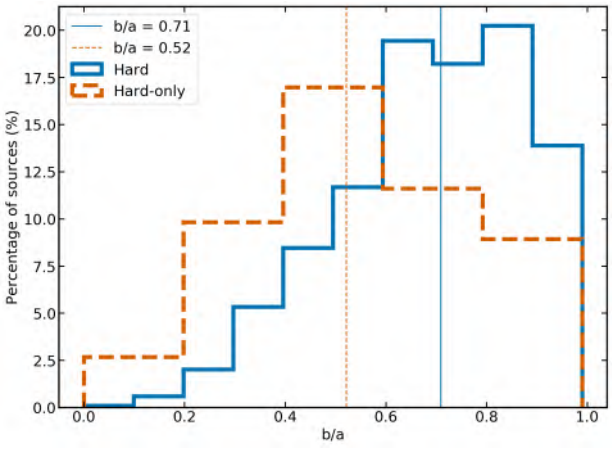

Abbildung 4.15: Ratio of the semi-major and semi-minor axes of the optical counterparts, where available. Sources are shown in blue solid lines, and sources in the hard-only sample are shown in orange dashed lines. The respective median values are indicated with vertical lines. Only sources with good `p_any` are shown.

4.4 Counterparts outside the legacy survey footprint

As an additional complement to the LS10 counterparts, it is also of interest to attempt to identify counterparts outside of the legacy survey area. For the hard-detected sources which also have soft emission, the positional uncertainties are very small (see Fig. 4.3). This enables a simple positional match to archival catalogs. The sources which are not covered in LS10 are primarily located at low Galactic latitudes, meaning that this sample consists primarily of stars and AGN located behind the Galactic plane. These AGN are of particular interest as it is more difficult to select AGN here. To do so, we first select 1710 sources which are outside of the LS10 MOC, have no flags, and are not hard-only. We then perform a 5" positional match to public QSO catalogs; the Gaia/Quaia AGN sample (Storey-Fisher et al. 2023), the Gaia/unWISE catalog (Shu et al. 2019) and the Million Quasar Catalog (MilliQSO; Flesch 2023). Although we expect some of these X-ray sources to be stars, given the small errors on the X-ray positions, we can still consider matches between X-ray positions and known AGN as interesting AGN candidates which can be explored more in the future. In all, we recover 692 matches out of 1710 sources, such that 40% of the sample outside LS10 has a QSO match. Out of these, 339 have spec-z from Quaia, and 349 have photo-z from the other catalogs (Flesch 2023; Storey-Fisher et al. 2023). These sources are also shown in Fig. 4.19.

4.5 Comparison with *Swift* BAT AGN

4.5.1 Match to *Swift* BAT

The *Swift* Burst Array Telescope (BAT) is sensitive in the $15-150$ keV energy range, and is thus highly complementary to many other contemporary X-ray missions. With a large field of view and rapid slew capability, the mission was designed to provide rapid follow-up for Gamma-ray bursts and other highly energetic events. While searching for these, the BAT instrument has also surveyed the sky in X-rays, with the first 105 months of operation providing a sample of 1632 hard X-ray detections (Oh et al. 2018) with a sensitivity of 8.40×10^{-12} ergs s^{-1} cm^{-2} in the $14-195$ keV band, over 90% of the sky. Counterparts have also been identified for these sources in Oh et al. (2018) and Ricci et al. (2017), and sources are classified based on their optical spectra or using other multiwavelength properties. This sample is of interest for comparison with eROSITA, given the highly complementary energy ranges, and the fact that both surveys cover bright, nearby sources, especially AGN.

To find matches with BAT, we first perform a 60 arcsecond cross-match between the eROSITA sky positions and the *Swift* BAT 105 month catalog counterpart positions provided by Oh et al. (2018). By examining the distribution of separations, we determined that a 10 arcsecond match radius was appropriate, as beyond this atch radius we found mostly secondary and tertiary matches. It should be noted here that while matching the X-ray positions directly to an optical catalog is not appropriate in this sample, it is extremely unlikely to have a spatially coincident BAT and hard X-ray selected eROSITA source which

do not have the same physical origin, justifying this matching procedure. Out of the 849 BAT sources in the eROSITA-DE area, we found 487 matches, of which 481 eROSITA are point sources and 456 are point sources with no eROSITA flags. There are no duplicates in this matching analysis. Additionally, 250 matched sources were inside the LS10 area. For the subsequent analysis, we consider all 487 matches.

It is then of interest to determine which types of BAT sources are detected by eROSITA. BAT sources are assigned a type by searching for publicly available optical soft X-ray and spectra, and the procedure is described in Oh et al. (2018). Fig. 4.16 (top) shows the numbers of matched and non-matched sources in the eROSITA-DE sky, separated by the different source classes in the BAT 105 month catalog (c.f. Table 1 of Oh et al. 2018). A large fraction of the sources detected with eROSITA are Seyfert 1 AGN with broad optical lines (142), followed by Seyfert 2 AGN (126) and beamed AGN (52). A large fraction of CVs, X-ray binaries, and unknown AGN (where the optical spectra and type are not available/known) are also detected, as well as sources of Unknown class II, which are sources with previous soft X-ray detection, but with unknown type. Of the 15 matches with the hard-only sample, seven are Seyfert 2, two are AGN of unknown class, two are unknown class II sources, and four are X-ray binaries.

Beyond type 1 and 2 Seyfert galaxies, examining the optical spectra can also allow further sub-division into the classes Seyfert 1, 1.2, 1.5, 1.8, 1.9 and 2.0, classified based on the strength of the broad optical lines. These are also taken directly from the BAT 105 month catalog (Oh et al. 2018). Fig. 4.16 (bottom) shows the numbers of matched and non-matched sources in the eROSITA-DE sky, according to these optical classes. It is clear that eROSITA is highly sensitive to detecting the relatively un-obscured, type 1-1.9 AGN, but is much less sensitive to detecting type-2 AGN, which is sensible as these sources are likely to be heavily obscured and thus very dim or absent in the eROSITA bandpass. The seven hard-only Seyfert 2s and two hard-only unknown AGN are also shown.

An earlier release of the BAT catalog (the BAT 70 month catalog) also reported information on the soft X-ray properties, including the column density, of BAT-detected AGN, where available (Ricci et al. 2017). We therefore match the eROSITA hard sample to this catalog, finding 273 matches. This match can also be justified by assuming that spatial coincidence of an eROSITA hard source with a *Swift* BAT source is extremely unlikely. For sources where the soft X-ray follow-up detection (with e.g. *Swift* XRT, *Chandra*, *XMM-Newton*) allows for the constraint of the column density (293 sources), we also identify matched sources in eROSITA (159). The histogram of these column densities is shown in Fig. 4.17. We note that Ricci et al. (2017) simply quote a column density of $20\,\mathrm{cm}^{-2}$ for sources which were not well constrained, and these have been removed from this plot.

The analysis shows that eROSITA appears to reliably detect sources up to column densities of $10^{23}\,\mathrm{cm}^{-2}$, at which point the detection fraction is $\sim 50\%$. At higher column densities the detection fraction decreases to $10-20\%$. Since the BAT instrument is relatively unbiased to obscuration, we can consider that these fractions represent the completeness of the eROSITA samples with respect to the absorption column density. In the Compton thick regime, we find one matched source which is nearby and very bright. This is the Circinus galaxy, ($z = 0.0015$) We may be detecting starburst emission with eROSITA (e.g.

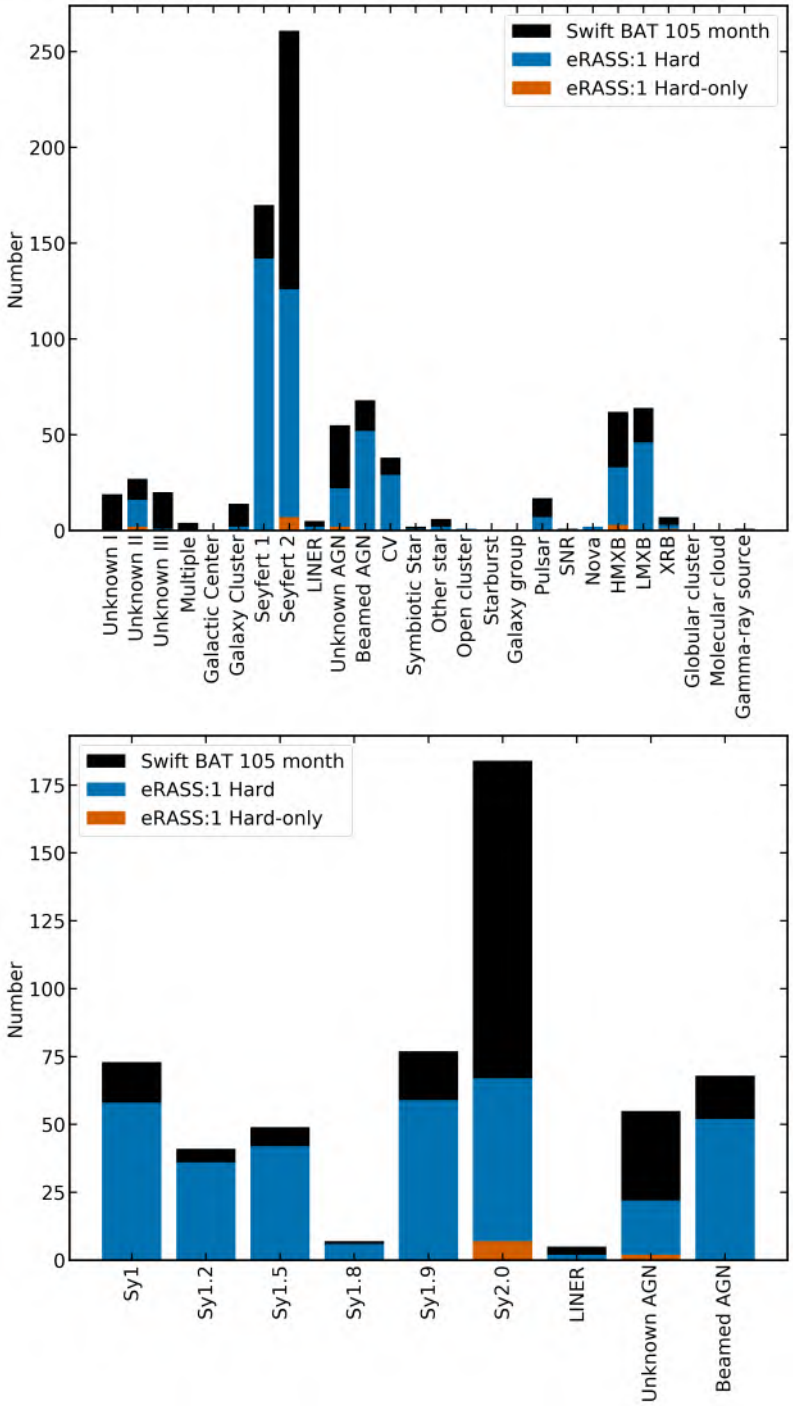

Abbildung 4.16: Top: Distribution of BAT 105 month catalog sources in the eROSITA-DE sky according to type (black), and of matched eROSITA sources according to BAT type (blue). Bottom: Distribution of BAT 105 month catalog sources according to BAT class (black), and of matched eROSITA sources according to BAT class (blue). Hard-only eROSITA sources are shown in orange.

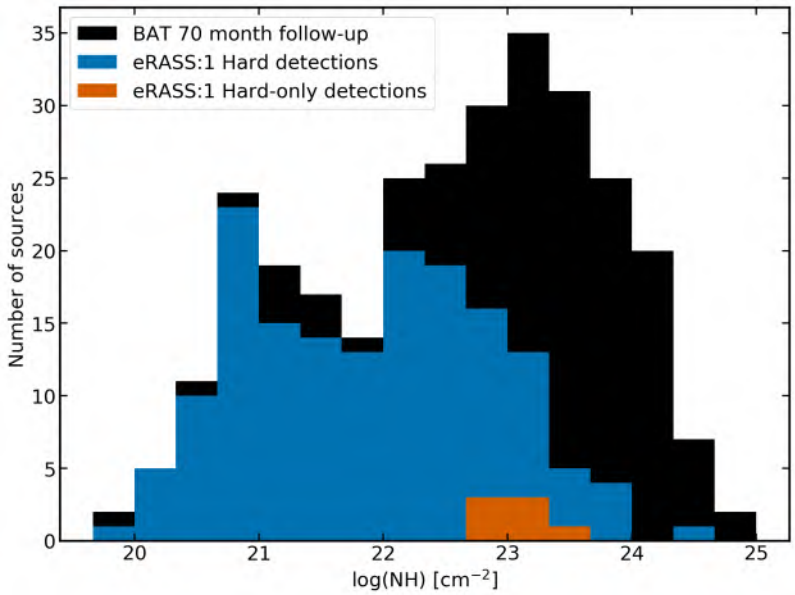

Abbildung 4.17: Distribution of column densities $\log(N_H)$ of sources with soft X-ray detections in the BAT 70 month catalog, shown for sources in the eROSITA-DE sky (black) and for those matched with eROSITA sources (blue). Hard-only eROSITA sources are shown in orange.

see Arévalo et al. 2014; Bauer et al. 2015) rather than directly detecting the transmitted component of the AGN. Of key interest is the distribution of the seven hard-only sources which also appear in this catalog. All seven have column densities of $> 10^{22.5}\,\mathrm{cm}^{-2}$, and a majority (4/7) are above $10^{23}\,\mathrm{cm}^{-2}$, further demonstrating that eROSITA can detect and select obscured AGN. The subsequent eROSITA all-sky surveys will permit further investigation of these sources, including identifying more sources and performing more detailed X-ray spectral analysis.

The distribution also shows a bimodality in column densities with a dip at $\sim 10^{22}\,\mathrm{cm}^{-2}$, which is also visible in the eROSITA matched sources and is also found in eFEDS (Nandra et al. 2024). This suggests the hard X-ray selected eROSITA sample will be able to give further constraints of unobscured to moderately obscured ($\sim 10^{22}\,\mathrm{cm}^{-2}$) column densities of AGN in the local Universe, where eROSITA is relatively unbiased to absorption effects.

4.5.2 Comparison of eROSITAand BAT counterparts

It is of interest to compare the counterparts of the eROSITA sources found in our NWAY match with those implied by the Swift-BAT matching. To do so, we select the 250 sources which had BAT 105 month catalog matches which are in the LS10 area, and remove three sources which were flagged as problematic in the X-ray catalog. We remove one additional source, NGC 1365, which has missing photometry in the LS10 catalog and is thus not found as a counterpart. We then match these to the X-ray sources, and measure the separation

between the BAT counterpart positions and the LS10 counterpart positions. Of these 246 sources, 227 have separations of less than two arcseconds, and are thus highly likely to represent the same optical source. For the remaining 19 sources, we visually inspect the LS10 image, and verify sources names and surrounding sources in SIMBAD. The maximum separation is ~ 8 arcseconds. All sources appear to be consistent, and every difference in position seems to be related to small differences between the coordinates in the BAT catalog versus other catalogs presented in SIMBAD. The analysis suggests that NWAY is performing well for the hard X-ray selected sample, and that the training sample used for NWAY (Salvato et al. in prep) is sufficient for counterpart matching for the hard sample. More discussion will be given in section 4.8.1.

4.6 Comparison with the Piccinotti AGN sample

The HEAO 1 experiment A2 (Rothschild et al. 1979) performed an X-ray survey in the $2-10$ keV band, covering 8.2sr of the sky at galactic latitudes $|b| > 20^{o}$ with a limiting sensitivity of 3.1×10^{-11} ergs s^{-1} cm^{-2}. A total of 85 sources were detected and identified, 31 of which are AGN, forming a complete hard X-ray selected sample consisting of 27 Seyfert 1s, three Seyfert 2s, and one blazar (3C 273) (Piccinotti et al. 1982). These well-studied AGN are worthy of analysis with eROSITA, as this may be considered as the previous best, all-sky, $2-10$ keV selected sample of AGN.

Out of the 31 AGN in the Piccinotti sample, 20 are in the eROSITA_DE sky. All of these sources are detected by eROSITA (both in the main and hard sample) and can be found in Table 4.3. eROSITA-derived $2.3-5$ keV fluxes and $2-10$ keV *Swift* BAT follow-up fluxes (Ricci et al. 2017) are given for each source, as well as the fluxes derived in HEA0-1. It can immediately be seen that many of the eRASS1 and BAT fluxes are below the values measured in HEAO1, and indeed many are below the HEA01 limiting flux. This may be due to intrinsic variability, as HEAO1 detected only the brightest sources in the X-ray sky and likely detected only AGN in a bright state. The relative fluxes measured between missions are also model depedent, and may be in better agreement with improved spectral modelling.

The BASS (Ricci et al. 2017) and eRASS1 fluxes are plotted in Fig. 4.18 for Piccinotti AGN (black, open squares), as well as the other sources in the BAT 70 month sample (circles), with colours indicating the different measured absorption column densities. The 70 month catalog is used in order to obtain the soft X-ray fluxes released in Ricci et al. (2017). Open circles show the sources with non-detections in eROSITA which have measured column densities of $N_H<10^{22.5}$ cm^{-2}, and the eROSITA fluxes are plotted at the upper 90% flux limit of eRASS1. Sources which are more obscured but not detected in eRASS1 are not shown, since it is understood that these sources were likely not detected due to obscuration and the relatively lower sensitivity of eROSITA in the hard band. In order to compare the fluxes measured by different instruments and in different energy bands, we trace two solid lines; the black line shows the equivalent 1-to-1 line assuming a photon index of $\Gamma = 1.9$ and a column density of $10^{20.5}$ cm^{-2}, which are the approximately average

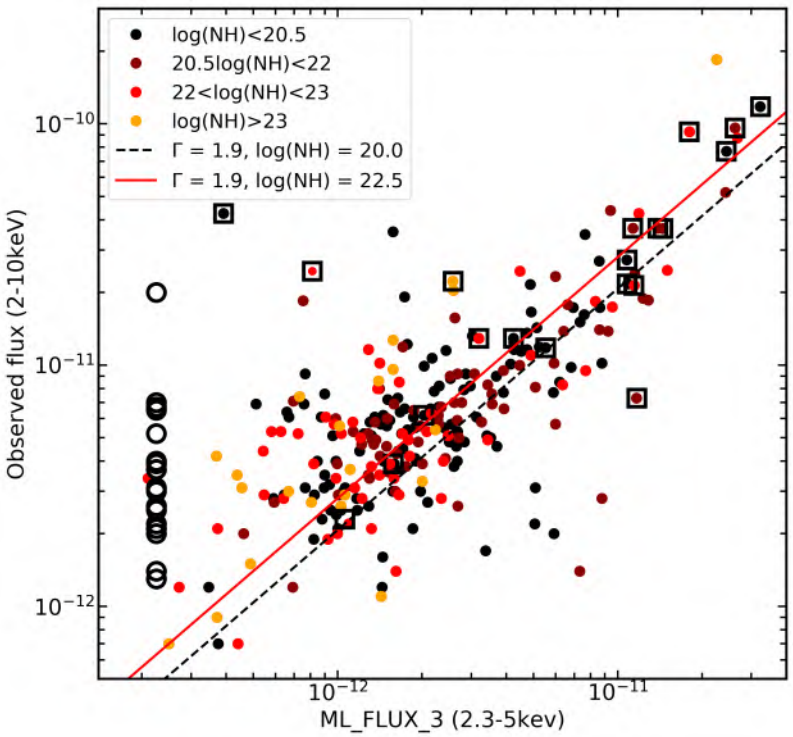

Abbildung 4.18: eROSITA $2.3-5\,\mathrm{keV}$ fluxes (ML_FLUX_3) and BASS fluxes ($2-10\,\mathrm{keV}$; Ricci et al. 2017) for Piccinotti AGN (black, open squares), as well as the other sources in the BAT 70 month sample (circles). Open circles show the sources with non-detections in eROSITA which have measured column densities of $N_H{<}10^{22.5}\,\mathrm{cm}^{-2}$. The black solid line shows the equivalent 1-to-1 line assuming a photon index of $\Gamma = 1.9$ and a column density of $10^{20.5}\,\mathrm{cm}^{-2}$. The red solid line shows the equivalent 1-to-1 line assuming a photon index of $\Gamma = 1.9$ and a column density of $10^{22.5}\,\mathrm{cm}^{-2}$. Sources with a measured column density of $< 10^{20.5}\,\mathrm{cm}^{-2}$, $10^{20.5-22.5}\,\mathrm{cm}^{-2}$, $10^{22.5-23}\,\mathrm{cm}^{-2}$, and $> 10^{23}\,\mathrm{cm}^{-2}$ are indicated with black, dark red, red and orange filled circles, respectively.

parameters of the matched eROSITA and BAT sample. While this tends to pass through the approximate parameter space occupied by most sources, a number lie significantly above or below this line.

Noticing that many of the sources which lie above the line show evidence for significant absorption, we add the second, red line which shows the equivalent 1-to-1 line assuming a photon index of $\Gamma = 1.9$ and a column density of $10^{22.5}\,\mathrm{cm}^{-2}$. Many sources with higher column densities lie closer to this line; while some still lie below it, the line indicates a very simple scenario with a fixed photon index of 1.9, and does not account for other effects including changing photon indices, a more complex spectral shape, or errors on the measured fluxes, and so should only be considered a rough indication of the parameter space. This analysis demonstrates that fluxes measured between *Swift* and eROSITA are moderately in agreement.

Tabelle 4.3: Sample of Piccinotti AGN in the eROSITA-DE sky. Column (1) gives the common name, column (2) gives the eROSITA name, and column (3) gives the *Swift*BAT name. Columns (4) and (5) give the optical RA and Dec, in degrees. Column (6) gives the eROSITA-derived observed $2.3-5$ keV flux. Column (7) gives the observed $2-10$ keV flux from the swift BAT 70 month catalog derived from soft X-ray follow-up, e.g. with *Swift* XRT or *XMM-Newton* (Ricci et al. 2017), and column (8) gives the $2-10$ keV HEAO-1 estimated flux quoted in Schartel et al. (1997).

(1) Name	(2) eROSITA Name	(3) BAT ID	(4) RA	(5) Dec	(6) eROSITA Flux $2.3-5$ keV ergs s^{-1} cm^{-2}	(7) BASS Flux $2-10$ keV ergs s^{-1} cm^{-2}	(8) HEAO1 Flux $2-10$ keV ergs s^{-1} cm^{-2}
Fairall 9	1eRASS J012345.7-584819	SWIFT J0123.9-5846	20.9408	-58.8057	1.08E-11	2.73E-11	2.8E-11
NGC 526a	1eRASS J012354.3-350356	SWIFT J0123.8-3504	20.9766	-35.0654	1.14E-11	2.14E-11	5.5E-11
ESO 198-024	1eRASS J023819.7-521133	SWIFT J0238.2-5213	39.5821	-52.1923	4.24E-12	1.29E-11	4.6E-11
3C 120	1eRASS J043311.1+052114	SWIFT J0433.0+0521	68.2962	5.3543	1.39E-11	3.69E-11	4.4E-11
3A 0557-383	1eRASS J055802.0-382007	SWIFT J0557.9-3822	89.5083	-38.3346	3.91E-13	4.25E-11	3.0E-11
H0917-074	1eRASS J092046.2-080323	SWIFT J0920.8-0805	140.1927	-8.0561	3.19E-12	1.29E-11	4.0E-11
NGC 2992	1eRASS J094541.9-141935	SWIFT J0945.6-1420	146.4252	-14.3264	1.17E-11	7.30E-12	7.5E-11
NGC 3227	1eRASS J102330.7+195153	SWIFT J1023.5+1952	155.8774	19.8651	1.13E-11	3.69E-11	3.7E-11
NGC 3783	1eRASS J113901.8-374418	SWIFT J1139.0-3743	174.7572	-37.7386	2.44E-11	7.69E-11	4.2E-11
3C 273	1eRASS J122906.7+020308	SWIFT J1229.1+0202	187.2779	2.0524	3.23E-11	1.18E-10	7.1E-11
NGC 4593	1eRASS J123939.4-052038	SWIFT J1239.6-0519	189.9143	-5.3443	1.58E-12	3.90E-12	3.8E-11
MCG -06-30-15	1eRASS J133553.8-341744	SWIFT J1335.8-3416	203.9741	-34.2956	1.45E-11	3.68E-11	4.6E-11
IC 4329A	1eRASS J134919.2-301834	SWIFT J1349.3-3018	207.3303	-30.3094	2.62E-11	9.63E-11	8.0E-11
NGC 5506	1eRASS J141315.0-031227	SWIFT J1413.2-0312	213.3119	-3.2075	1.80E-11	9.27E-11	5.9E-11
IRAS 18325-5926	1eRASS J183658.1-592408	SWIFT J1836.9-5924	279.2429	-59.4024	8.12E-13	2.45E-11	3.4E-11
ESO 103-035	1eRASS J183819.7-652537	SWIFT J1838.4-6524	279.5848	-65.4276	2.58E-12	2.23E-11	3.0E-11
H1846-786	1eRASS J184703.0-783148	SWIFT J1848.0-7832	281.7618	-78.5304	2.03E-12	6.20E-12	2.9E-11
ESO 141-G055	1eRASS J192114.2-584013	SWIFT J1921.1-5842	290.3090	-58.6703	1.08E-11	2.17E-11	3.7E-11
NGC 7213	1eRASS J220916.5-470959	SWIFT J2209.4-4711	332.3177	-47.1667	5.53E-12	1.18E-11	2.3E-11
NGC 7582	1eRASS J231906.0-420648	SWIFT J2318.4-4223	349.5979	-42.3706	1.06E-12	2.30E-12	5.5E-11

4.7 Variable point sources

Boller et al. (2024) have analysed a sample of sources from the eRASS1 main catalog to check for evidence of variability within eRASS1. Due to the scanning pattern of eROSITA, this typically serves as a test of variability on 4-24 hour timescales, where 4 hours is the length of one eroday, which is the length of one revolution of *SRG*in all-sky survey mode, and each source is covered typically six times in the survey, outside the ecliptic poles. For each source, a light curve was produced (Merloni et al. 2024). Two variability tests are then performed by Boller et al. (2024): the maximum amplitude variability, defined as the span between the maximum and minimum values of the count rate, and the normalised excess variance (NEV), defined as the difference between the observed variance and the expected variance based on the errors (Boller et al. 2022, 2024, see). These variability tests are run in the $0.2-2.3\,\mathrm{keV}$ band, so while they do not overlap with our hard sample selection, they still provide an interesting look into the variability in the energy band where eROSITA is most sensitive.

Matching the variable source catalog with the hard-band catalog, we find that 326 sources show evidence for variability. Out of these, 159 variable sources are within the LS10 area, and have good counterparts. Of these 159, 40 sources have spectroscopic redshifts of $z > 0.002$ and are thus confirmed to be extragalactic, and 109 are stars. Outside of the LS10 area, 125 variable sources have matches in the AGN catalogs used in this chapter, five of which have a spectroscopic redshift of $z > 0.002$ from Quaia. This gives a total of 45 variable AGN with a spec-z. These AGN are of particular interest for future follow-ups within later eRASSs or with other missions, to study both the short- and long-term variability of spectroscopic AGN. Further details of the specific variability properties of these sources and a more detailed study of their variability is left for future work.

4.8 A hard X-ray spectroscopic sample of AGN

4.8.1 AGN sample designation

A key output of the classification work described in sections 4.3 and 4.4 is a high purity, hard X-ray selected sample of AGN detected with eROSITA. To ensure a very pure sample, we select only sources with reliable counterparts, and only those with spectroscopic redshifts, where $z > 0.002$. More concretely, this sample is assembled from:

1. All LS10 sources with a good counterpart and a spectroscopic redshift $z > 0.002$, and removing sources with class_beamed > 0 (1243 sources, 23 hard-only)

2. Three AGN which were in the LS10 area but were missing counterparts, as listed in Table 4.2 (three sources, zero hard-only)

3. Sources from the *Swift* BAT 105 month catalog crossmatch which satisfy the following criteria: eROSITA position outside of the LS10 area, no flags in the X-ray catalog,

X-ray point source, BAT 105 month catalog Class of 40, 50, or 60 (corresponding to Seyfert 1, Seyfert 2, or LINER AGN, excluding beamed AGN and unknown type AGN) (82 sources, six hard-only)

In total, this selects 1328 sources, of which 29 are hard-only, as can be seen in Table 4.1. The luminosity-redshift distribution of these sources is shown in Fig. 4.19. They have median redshift of $z \sim 0.19$, which is slightly lower than the median from the eFEDS sample of 0.34 (see Chapter 2), albeit the eRASS1 hard sample has lower spectroscopic redshift coverage. Sources detected in the BAT 70 month sample are also shown in pink, and sources detected in the eFEDS hard sample are indicated with green triangles. The hard spectroscopic redshift sample is ideal for spectral analysis, since, as shown in Chapter 3, the presence of the photons above 2 keV will allow modelling of the contribution from the hot X-ray corona, as well as an investigation into soft X-ray emission and absorption. Furthermore, the eROSITA sample is able to probe dimmer sources than previous all-sky hard X-ray selected samples performed by HEAO-1, and Swift-BAT, and hence is also capable of finding many more sources at higher redshifts. The eRASS:1 hard sample covers a broad redshift range of $0.003 - 3$, and an even broader hard X-ray luminosity range (estimated from the $2.3 - 5$ keV band counts) of $10^{41} - 10^{47}$ ergs s^{-1}. We caution that while only few eRASS1 sources appear above the HEAO1 flux limit, all 20 Piccinotti AGN in eROSITA_DE are still detected, albeit at lower flux states.

Also shown here are the 339 sources with spectroscopic redshifts obtained from matching the sources outside the LS10 area to AGN catalogs, as described in chapter 4.4. These sources lie primarily within the Galactic plane and may be obscured by the local gas and dust, which may decrease the flux of these sources. A majority are also found at higher than median redshift, although examining the histogram in Fig. 4.11, we can see that the Quaia redshifts from the area in LS10 are similarly high, so this may be due to the sensitivity and bandpass of the spectrograph onboard Gaia. It should also be noted here that some of the higher redshift sources may be blazars.

4.8.2 Hard-only spectral fitting

In this section, I focus on the hard-only sources; 29 of which are not beamed and have spectroscopic redshifts $z > 0.002$, 23 obtained from the `NWAY` LS10 match and six from the *Swift* BAT match. This interesting subsample covers a large range in redshift $0.003 < z < 1.5$ and hard X-ray luminosity $10^{41} - 10^{46}$ ergs s^{-1}, and is likely to comprise heavily obscured AGN. We fit each eROSITA X-ray spectrum in the $0.3 - 8$ keV range with a torus model (`UXClumpy`; Buchner et al. 2019), using BXA (Bayesian X-ray Analysis; Buchner et al. 2014). BXA uses nested sampling combined with XSPEC (v. 12.12.1; Arnaud 1996) in order to explore the full model parameter space to best estimate parameters and enable model comparison. The model `UXClumpy` includes transmitted and reflected components with fluorescent lines obtained by assuming a clumpy torus with a specified opening angle and inclination. Since many of the eRASS1 spectra have very few counts, some model parameters are frozen to reasonable values obtained from Buchner et al. (2019), including

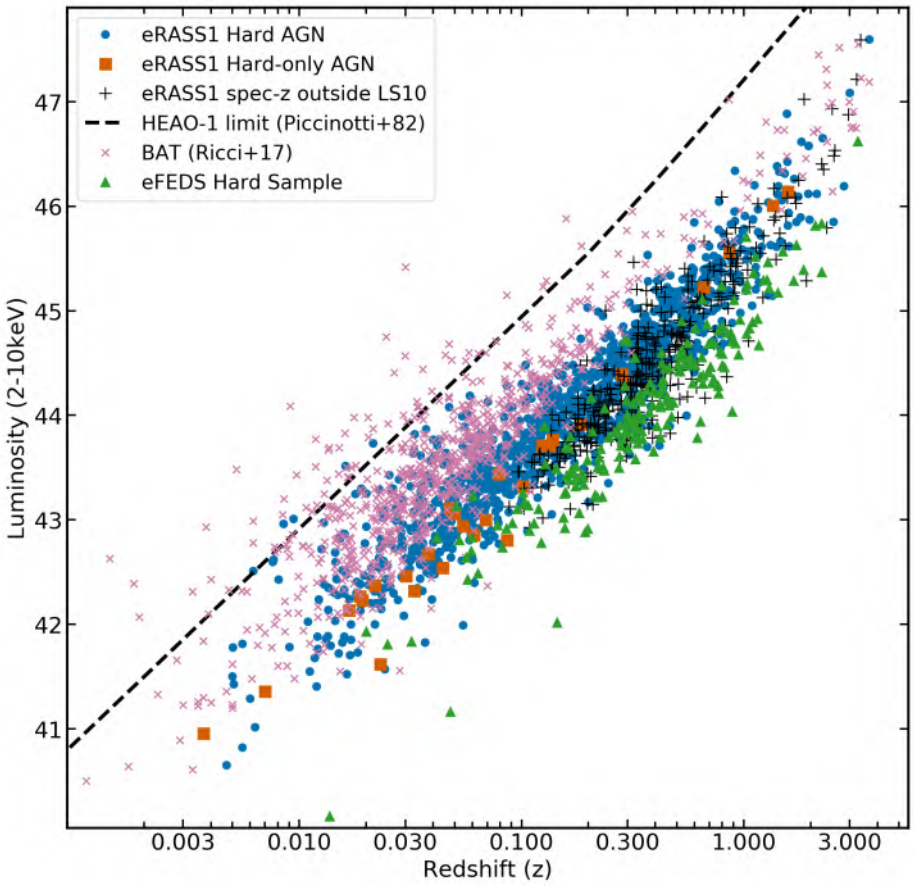

Abbildung 4.19: Hard X-ray luminosity versus redshift for various hard X-ray selected samples. eRASS1 AGN are shown as blue circles, eRASS1 hard-only AGN are shown with orange squares, *Swift* BAT AGN from the 70 month catalog are shown as pink x, and the eFEDS hard sample are shown with green triangles. The flux limit of HEAO-1 is shown with a black dashed line. Sources outside LS10 with spec-z are shown as black crosses.

the width of the Gaussian distribution modelling the vertical extent of the torus cloud population (`TORsigma`, 20), and the viewing angle of the torus (`Theta_inc`, 45^o). The photon index is given a Gaussian prior centered on 1.95 with a width of 0.15, and the column density varies uniformly in $\log_{10}$ space between 20 and 26.

In order to model an additional soft population of X-ray photons, an additional power law is added, with photon index and normalisation linked to the `UXClumpy` continuum. This scattered power law is then renormalised by a constant factor which is left free to vary between 0.01 and 0.1. There are a number of interpretations for this soft X-ray emission; likely it represents a population of scattered photons which leak through the torus and pass into our line of sight (Brightman & Nandra 2011; Liu & Li 2014; Furui et al. 2016; Buchner et al. 2019), but soft X-ray emission can also be produced by other mechanisms such as supernova remnants, ultraluminous X-ray sources and X-ray binaries (e.g. see discussion in Buchner et al. 2019, on Circinus), within the host galaxy which is unresolved in X-rays. An example of the corner plot for source 1eRASS J151012.0-021454 (LEDA 54130) is shown in Fig. 4.20. While some sources have double peaked solutions for the column density, others, as this one, are well-constrained, with little degeneracy between parameters.

The resulting column density distribution is shown in Fig. 4.21. The eRASS1 hard-only sources are shown in orange. A comparison is also given to the eFEDS sample, using spectral modelling as described in Nandra et al. (2024). The eFEDS hard sample is shown with a green dash-dot line, and the eFEDS hard-only sources are shown in black. While the hard sample as a whole contains many very bright, unobscured AGN, the hard-only sources recover a population of moderately obscured AGN, with column densities of $\sim 10^{23-24}\,\mathrm{cm}^{-2}$. These column densities are higher than what is typically found in the discs of galaxies, so may originate in obscuring material closer to the central engine, such as the torus.

Also of interest here is the scattering fraction, which shows how many photons leak through the torus. We find that the 1σ upper limits for all 29 hard-only sources with spec-z of the scattering fractions are all $< 3\%$, and the median values of the resulting posteriors are all $< 1\%$, shown in the filled and unfilled histograms of Fig. 4.22 (top left), respectively. These correspond to $0.5-2\,\mathrm{keV}$ X-ray luminosities on the order of $10^{40}\,\mathrm{ergs\,s}^{-1}$. These values are lower than found in previous samples (Brightman et al. 2014). These are shown in Fig. 4.22 top right (upper limits), and Fig. 4.22 bottom (median values) with the upper limits on the scattering fraction derived from COSMOS, AEGIS-X and CDFS shown compared to those from this work. Since eROSITA is extremely sensitive to soft X-ray photons, any small scattered emission should readily be detected, so by placing tight constraints on the detection threshold below 2.3 keV, a more extreme source population is selected. There is some hypothesis that sources with low scattering fractions are associated with a buried AGN having a spherical torus which may arise in the late- or post-merger stages (e.g. Yamada et al. 2019). However, looking at the LS10 images, we find many edge-on sources, in agreement with the very low b/a values found in chapter 4.3.7 (see also Fig. 4.15). Together, this may suggest that, particularly for the edge-on sources, some of the obscuration is associated with a ∼kpc scale obscurer in the host galaxy, versus solely in a parsec scale torus in the inner region. For the face-on sources, a uniform and spherical

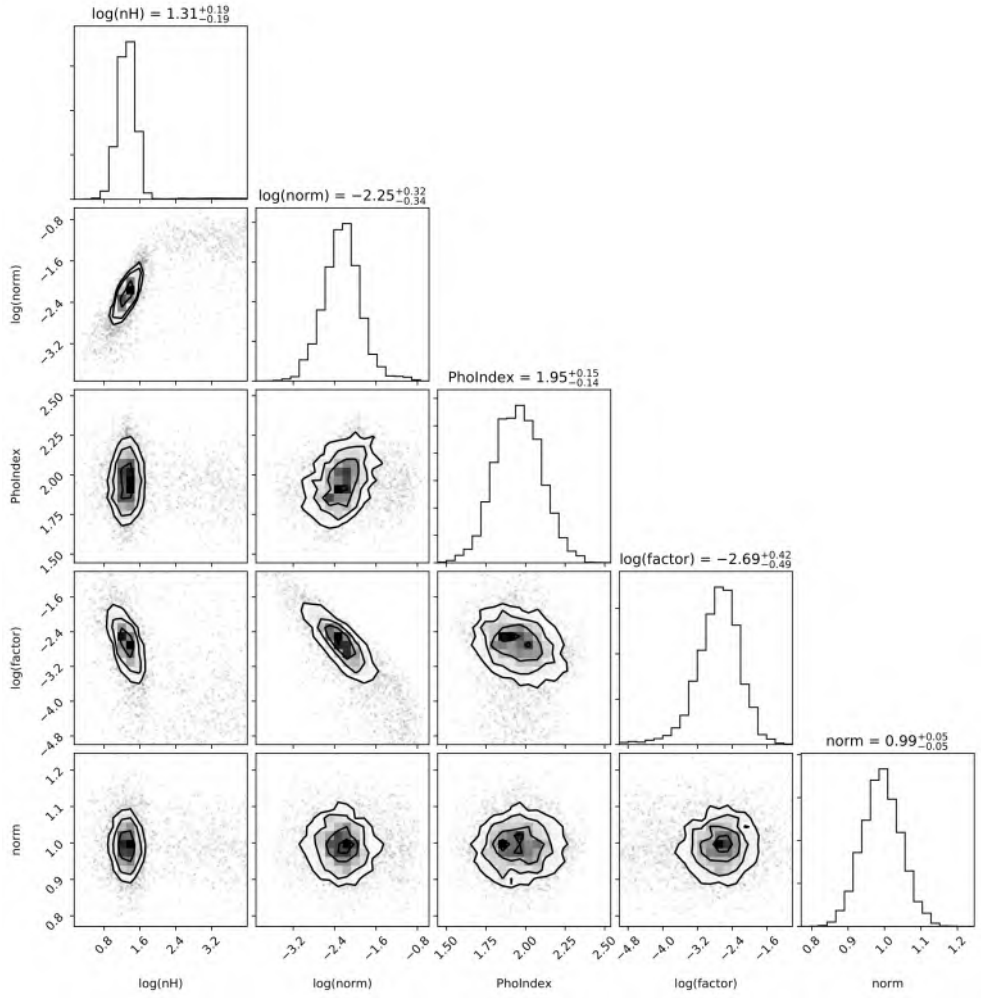

Abbildung 4.20: Corner plot resulting from applying the model described in section 4.8.2 to 1eRASS J151012.0-021454. The column log(nH) measures the torus column density in units of $\log(N_H) - 22$, such that a value of 1 corresponds to a column density of $10^{23}\,\mathrm{cm}^{-2}$. The column log(norm) is the normalisation of the `UXClumpy` model, which is also linked to the power law normalisation. PhoIndex is the photon index which is linked between the `UXClumpy` and power law components. The column log(factor) is the renormalisation constant between the reflected and transmitted components modelled by `UXClumpy` and the scattered power law component. Finally, norm is the relative normalisation of the background model within the source field, which should be ~ 1 as we have already accounted for the area differences between the source and background regions.

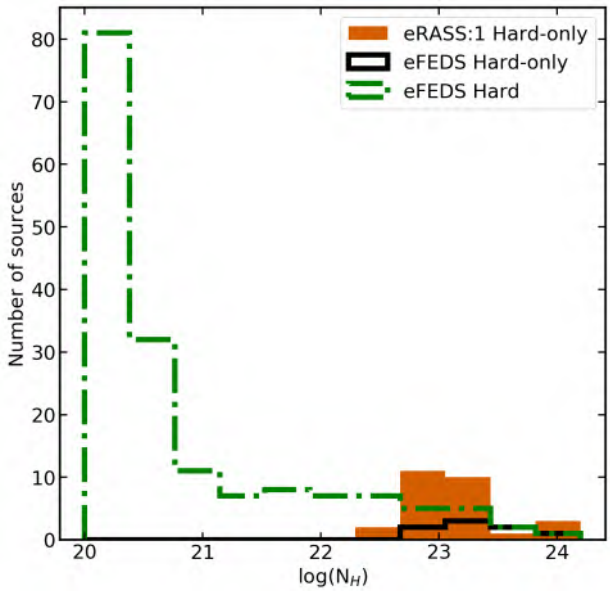

Abbildung 4.21: Distribution of column densities in the eFEDS hard (green dash-dot lines), eFEDS hard-only (black solid line), and eRASS1 Hard-only (orange filled histogram) samples. The column densities from eFEDS are shown for the AGN sample presented in Nandra et al. (2024), and the eRASS1 hard-only column densities are from the `UXClumpy` fits presented in this chapter.

torus is preferred to reproduce the high level of obscuration and low scattering. Obscuration in hard-only sources is discussed further in Sect. 4.9.2 and in Chapter 5.

4.9 Discussion

4.9.1 Counterparts and associations

From the first all-sky imaging survey in the $2-8$ keV band, in the western galactic sky, eROSITA has detected 5466 sources with an estimated spurious fraction of about 10%. In this chapter, I have identified and classified the 4895 clean, point-like sources present in the sample. For sources within the legacy survey DR10 area, we use `NWAY` to select counterparts for 2850 sources, and identify 2547 as having a good counterpart. We then flag stars and beamed AGN and match with public spectroscopic redshift catalogs to finally select 1243 sources inside the LS10 are with $z > 0.002$ and hence extragalactic.

An issue of particular importance is the relatively higher spurious fraction of the hard-only sources, shown in Fig. 4.2 and chapter 4.2.2. This fraction is expected to be around 65% at a `DET_LIKE_3` > 12, much higher than the $< 10\%$ for the sources with both oft and hard X-ray detections. Based on the positions of sources in Fig. 4.13, it is estimated that $\sim$ 20% of the hard-only sources with good `p_any` are expected to be spurious X-ray sources with chance associations to very dim galaxies. Most of these sources have a

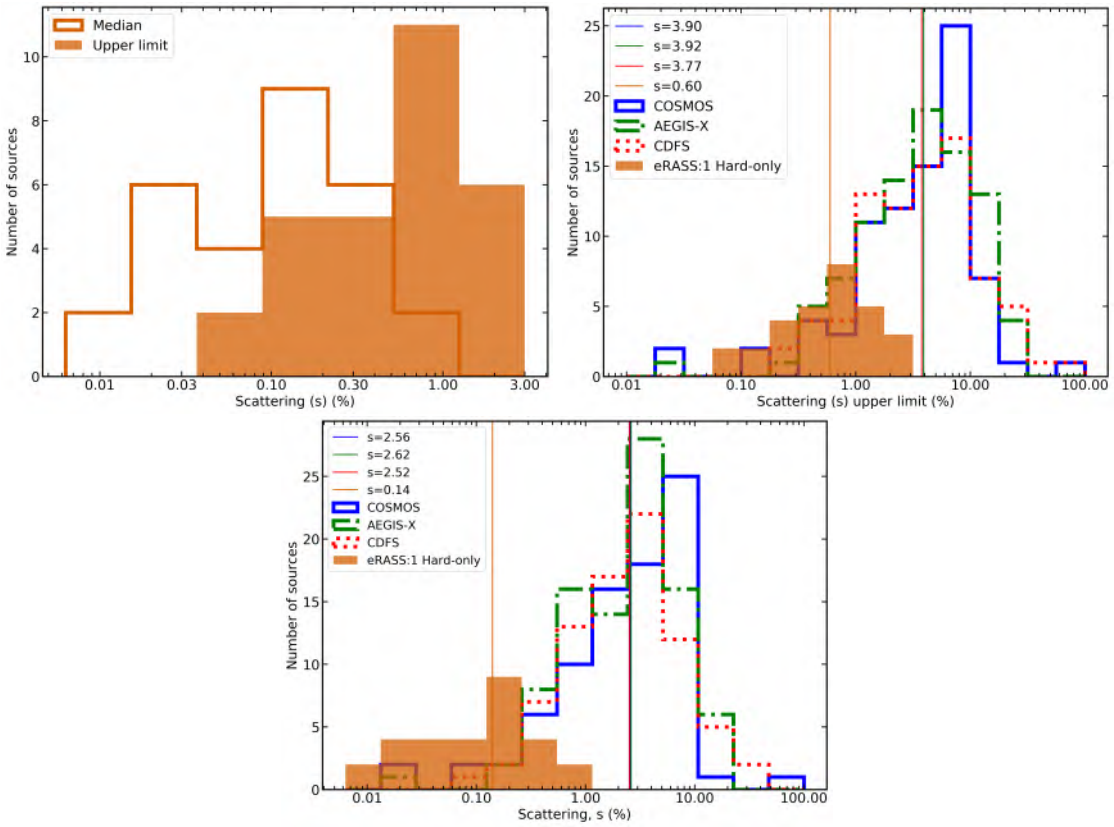

Abbildung 4.22: Top left: Distribution of the posterior median of the scattering fraction (orange unfilled histogram) and 1σ upper limit of the scattering fraction (orange filled histogram), shown as a percentage, for the eRASS1 spectroscopic hard-only sample. Top right: Distributions of the upper limits of the scattering fraction (shown as percentages) for this sample, in comparison to the best-fit *Chandra* results found by Brightman et al. (2014) The distributions measured from COSMOS, AEGIS-X and CDFS are shown as blue solid, green dash-dot and red dotted lines, respectively, all taken from Brightman et al. (2014) The eROSITA values are shown with an orange filled histogram. The median values are shown in vertical lines of corresponding colours, with almost exact overlap between the COSMOS, AEGIS-X and CDFS samples. Bottom: same as upper right, but shown for the median scattering fractions (%).

DET_LIKE_3 $\sim$ 12, while those with DET_LIKE_3 $>$ 15 tend to have brighter counterparts. Improvements could be made to the purity of the hard-only sample by applying further cuts on DET_LIKE_3 or p_any (e.g. at 0.1, corresponding to 95% association purity, or at 0.61, corresponding to 99% association purity); however, we stress that selecting only sources with spec-z also removes the optically dim sources.

One caveat to this analysis is that the priors in the NWAY run were constructed using a 4XMM and *Chandra* training sample which was created to match the eRASS1 main sample. The hard sample, however, is in general comprised of local AGN, similar to eFEDS finding a sample of bright and local AGN (see Chapter 2). Therefore, a majority of the AGN are brighter in optical and X-rays than the peak of the distributions from the training sample. The soft X-ray flux distributions are more comparable for the hard-only sources, but these may also have more extreme parameter values are they are a very distinct subclass of AGN. At the moment, there is not a large and well-defined enough training sample to select from, but using the counterparts from this sample to inform the training sample of future eROSITA all-sky surveys will help to improve counterparts and p_any values for future works. Further, despite this caveat, it should be noted that there is a large overlap between the *Swift* BAT counterparts and the NWAY counterparts from the eROSITA hard sample, suggesting that NWAY is performing well for this sample.

Examining the spectroscopic redshift distribution from Fig. 4.11, a majority of sources are very local, with a peak at $z \sim 0.3$ which is in agreement with eFEDS (Chapter 2). This is driven by the peaks of the sources with spectroscopic redshifts from NED, as well as from Quaia. Studying only the very secure spectroscopic redshifts from the compilation, the redshift peak is lower, at approximately $z \sim 0.07 - 0.08$. The hard-only sources also have a surprisingly large spread of redshift values, from $z \sim 0.03$ to $z > 1$. Studying the luminosity-redshift values from Fig. 4.19, we find that the all-sky survey selects many brighter sources than compared to eFEDS, but dimmer and higher redshift sources than BAT, providing a rich and unique sample well-suited for further exploration.

4.9.2 Absorption in hard-only AGN

A bona-fide hard-only sample can be produced by combining our hard-only selection with the information on the purity of the counterpart. We note that improvements in purity to the hard-only sample could be made by applying a higher threshold on DET_LIKE_3 (as can be seen in the Fig. 4.2) or p_any (e.g. at 0.1, corresponding to 95% association purity, or at 0.61, corresponding to 99% association purity), or requiring spectroscopic redshift. Sect. 4.8.2 has demonstrated the power of X-ray spectral fitting 29 hard-only AGN with spec-z, with interesting results. As can be seen from Fig. 4.19, the eROSITA hard-only sources span orders of magnitude in redshift and luminosity, but are more often found in low redshift, lower luminosity AGN. Given the nature of the eROSITA survey strategy, the sources being fit have very few photons and low exposure times, which would typically suggest that there is little to be gained from spectral analysis. In fact, all hard-only sources have less than 20 counts, which was identified by Liu et al. (2022b) as a cut-off below which little could be gained from parameter estimation and model comparison. However, spectral

fitting is aided by a good understanding of the typical photon indices of eROSITA spectra (e.g. from Liu et al. 2022b; Nandra et al. 2024) Crucially, for this work, fitting is enhanced not by what eROSITA detects, but what is not detected; simulations (Seppi et al. 2022) have given an accurate estimation of the significance and spurious fraction of different detection thresholds. By selecting a very low threshold combined with the knowledge of eROSITA's extreme sensitivity to soft X-ray photons, the hard-only sample comprises of sources with little to no soft X-ray emission (below 2.3 keV). Correspondingly, the hard-only sample appears to consist of sources with high level of obscuration.

Through spectral analysis, it is clear that not only are the column densities very high (with a median of $\log(N_H) = 23.1$), there are three sources for which the best-fit column density is $> 10^{24}\,cm^{-2}$, classifying them as Compton thick AGN. This represents 10% of the hard-only spec-z sample, which in turn is $\sim 1\%$ of the sample with counterparts. This may suggest that eROSITA can detect a small percentage of bright Compton thick AGN. We caution however that we were not able to detect sources which Swift-BAT also identified as Compton thick. Continued study of these sources in later eRASSs as well as searching for newly detected hard-only sources will help to confirm this result.

Spectral analysis reveals an extremely low percentage of photons scattered into the line of sight. The sample median of the relative normalisation of the unobscured second powerlaw is only 0.14% (see Fig. 4.22). If this component is instead interpreted as emission from X-ray binaries or supernovae in addition to scattered photons, this implies that any scattered emission fraction is even lower than measured. As shown in chapter 4.8.2, our values are much lower than what found from other samples (e.g. Brightman et al. 2014). In Brightman et al. (2014), data were obtained from *Chandra*, and the X-ray spectra were not analysed using `UXClumpy`, but rather with `BNtorus` (Brightman & Nandra 2011), using an older version of this code which allowed soft X-ray leakage, making the spectral shape extremely similar to `UXClumpy`. Therefore, comparison between measured values is justified. The differences do not suggest an intrinsic difference between AGN detected between the two missions, and spectral modelling of the full eRASS1 hard sample would likely return a more similar overall scattering fraction distribution to the *Chandra* samples from Brightman et al. (2014). Rather, this analysis demonstrates that the hard-only sample is a highly unique sub-sample whose properties may be of particular interest, and that this is distinct from other surveys and samples.

In chapter 4.8.2, it was also discussed that the small recovered scattering fractions may imply that the obscuration is associated with a ∼kpc scale obscurer on larger scales than the typical X-ray torus (e.g. Hönig et al. 2014). This result was further supported by the edge-on galaxies in the images, and also by the low b/a values measured in LS10. This may contrast to what is suggested in other works, where it is hypothesized that the low scattering fractions are more indicative of a spherical torus (e.g. Yamada et al. 2019). For sources that are face-on, a uniform and spherical torus is likely the preferred interpretation. This is also contrary to many works which suggest that the torus is patchy and not fully obscuring (e.g. Maiolino et al. 1999; Nenkova et al. 2008; Markowitz et al. 2014; Aird et al. 2015; Buchner et al. 2015, 2019). The hard-only sample therefore allows us to probe a unique and extreme parameter space of torus structure. This interesting result could be

studied in more detail using broad-band X-ray data, with a deeper observation of the soft X-ray to search for a scattered component or for time variability, and of the hard X-ray to model the shape and parameters of the obscurer and underlying power law. This topic is explored in Chapter 5, using *NuSTAR* data used to complement the eROSITA data and to estimate absorber parameters.

4.10 Lessons learned from the eRASS:1 Hard sample

SRG has performed the first all-sky imaging surveys above 2 keV. The hard X-ray selection decreases the bias against obscured sources compared to e.g. the eRASS1 main sample. The eROSITA hard sample presents an extraordinary improvement over the Piccinotti AGN sample, and allows us to probe deeper in flux and redshift than e.g. the *Swift* BAT all-sky survey. In addition to the counterparts and multiwavelength information provided, we identify a population of heavily obscured AGN which lack detection in eROSITAs most sensitive soft energy, but are detected in the hard band, similar to the population of hard-only sources identified in eFEDS In Chapters 2 and 3. Through X-ray spectral analysis of these sources, combined with the images and optical analysis, it was hypothesized that the obscuration of these sources could be associated with both a spherical torus and/or the galactic disc. Follow-up observations of these sources with other X-ray missions (e.g. *NuSTAR* will help to more accurately model the X-ray spectra (Chapter 5). Continued optical spectroscopic follow-up of these sources (e.g. with SDSS-V), and indeed of the entire hard sample, will help to identify more AGN, and hard-only AGN. Finally, combining these results with future X-ray all-sky surveys (including eRASS:2, eRASS:3 and eRASS:4) will help to further secure the source catalog and counterparts, shining a new light on the nature of X-ray sources above 2 keV.

Kapitel 5

eRASS1: Constraining absorber geometry in eROSITA hard sources

This chapter builds upon the findings on the hard-only X-ray selected sample of AGN presented in chapter 4, specifically with the goal of understanding the physical origin of the observed low scattering fractions. These hard-only sources are obscured AGN with column densities of $\sim 10^{23}\,\mathrm{cm}^{-2}$, and lack secondary soft X-ray emission components, as characterised by their low measured scattering fractions. This is a selection effect; given eROSITA's extreme sensitivity to soft X-ray emission, sources which do not show soft X-ray detections in eRASS1 must have both low nuclear scattered emission as well as no additional emission components on larger scales. This makes them ideal for investigating the nature of the obscurer. We aim to study the source of obscuration in these AGN. A comprehensive spectral analysis of the eROSITA hard-only sources is presented, as well as a complementary sample of high hardness ratio sources, which have also been observed using *NuSTAR*. This gives a total sample of 16 sources. Extending a broad energy range from $3-79\,\mathrm{keV}$, *NuSTAR* enables direct probing of the Fe Kα line and the Compton hump, which we model using a line of sight absorber, a smooth uniform obscurer, and a clumpy obscurer. The results are considered in the context of host galaxy properties, including the morphology type and orientation. Through X-ray spectral analysis, is is demonstrated that a clumpy torus is required for most sources, with the exception of AGN in edge-on spiral galaxies, where a smooth absorber is sufficient. A tentative link is identified between the host galaxy classification and orientation and the nature of the X-ray obscuration, where edge-on spiral galaxies have absorption in the galactic disc while other galaxies are obscured by a spherical obscurer seen nearly face-on.

This chapter is based on Waddell et al. (submitted to A&A). I performed the sample collection and classification, all X-ray spectral analysis, and relevant data analysis presented in this chapter.

5.1 Obscuration in AGN

The X-ray spectra of Seyfert 2 (or type-2) AGN typically show signatures of absorption (see e.g. Awaki et al. 1991; Turner et al. 1997; Risaliti et al. 1999; Brightman & Nandra 2011; Buchner et al. 2019), which is typically associated with a parsec-scale nuclear obscurer referred to as the torus (Antonucci 1993; Urry & Padovani 1995). Light from the inner region passes through this structure, resulting in the suppression of both the soft X-rays as well as broad lines in the optical spectrum. Compton scattering of X-ray photons in the line of sight obscurer also produces emission above the coronal power law in the $10-50\,\mathrm{keV}$ band, known as the Compton hump, as well as a fluorescent Fe Kα emission line at $6.4\,\mathrm{keV}$ and an associated absorption edge at $\sim 7.1\,\mathrm{keV}$. Various X-ray models employing different obscurer geometries and properties have been used to model these observed X-ray features (e.g. Magdziarz & Zdziarski 1995; Murphy & Yaqoob 2009; Brightman & Nandra 2011; Baloković et al. 2018, 2019; Buchner et al. 2019). Increasing sample size, signal-to-noise ration, and spectral resolution have also lead to continuous improvements of spectral models over time.

There is now strong evidence for a complex geometry for the torus, beyond a simple, smooth obscurer. Through the study of X-ray occultation events, we can determine that the torus is clumpy (e.g. Risaliti et al. 2002; Markowitz et al. 2014; Laha et al. 2020), such that many discrete Compton-thin clouds pass between the emitting region and the observer over time. Furthermore, the observed infrared spectra cannot be explained by smooth homogenous absorption, further implying that the torus is clumpy (see e.g. Nenkova et al. 2002; Mason et al. 2009). A recent review by Netzer (2015) further confirms the inhomogenous nature, stressing that the torus also has multi-phase (e.g. cold and hot) dust components (Netzer 2015). The nuclear absorber is likely closely linked to the accretion or ejection mechanisms, so understanding it is of key importance. Furthermore, it is a crucial element in the demographics of AGN, with a large portion of accretion power likely being obscured (e.g. Hasinger et al. 2007b; Ueda et al. 2014; Aird et al. 2015).

Hints on the geometry of the X-ray obscurer can also be found in the observed X-ray spectrum. As X-ray photons pass through the obscuring medium, they are Compton down-scattered, losing energy and producing the observed Fe Kα emission and absorption as well as the Compton hump. The precise shape of these components is dependent on the properties of the obscurer, specifically the overall density, the line of sight column density, the covering factor and the viewing angle. Constraints on the nuclear obscurer geometry can therefore be obtained using a sophisticated new model, `UXClumpy` (Buchner et al. 2019) which employs a clumpy absorber structure to reproduce the observed column density distribution, radial cloud distribution, and eclipse event rates. This model has also been applied successfully to individual sources and recovers the observed X-ray spectral shape of type-2 AGN (Buchner et al. 2019).

Beyond Compton scattering, there are also further potential soft X-ray emission components, including emission from hot gas or X-ray binaries in the host galaxy, or from emission which is not absorbed in the torus, and is instead transmitted directly. To measure the importance of such secondary soft emission, an additional power law component

to the model that undergoes no absorption is added. The normalisation of this power law component is then measured relative to that of the instrinsic power law component, with typical values of 1–10% (Turner et al. 1997; Cappi et al. 2006; Rivers et al. 2013; Brightman et al. 2014; Buchner et al. 2014; Ricci et al. 2017). In some sources, these scattering fractions reach lower values of 0.01-1%. A first exploration of such AGN in X-rays is presented in Ueda et al. (2007), where two sources with low scattering fractions of $\sim 0.2\%$ observed with *Suzaku* (Mitsuda et al. 2007) (ESO 005–G4 and ESO 297–G018) were analysed. The authors classify these sources as a new type of buried AGN, proposing that these sources are likely obscured by a geometrically thick, spherical obscurer with a small opening angle, estimated to be $< 20^{o}$. For one source it was also presumed that the system was being viewed at a small inclination (face-on), whereas the other source was proposed to be viewed more edge-on, but through a patchy obscurer with a dense component outside the line-of-sight. Interestingly, both sources are associated with spiral galaxies viewed extremely edge-on, so this would imply misalignment between the obscurer and the galactic disc. An alternate proposal was that both sources have low scattering fractions due to a lack of gas responsible for scattering. The optical spectra of these sources also revealed weak [OIII] λ5007A lines, meaning that these sources may be missed in optically selected surveys. These sources are of interest to be studied in X-rays, as they provide a pure look at the reflection spectrum, with minimal contribution from other emitting components.

Hönig et al. (2014) investigated absorption in a larger sample of low X-ray scattering fraction AGN.These sources reveal significant diversity in the host galaxy orientations and classifications (c.f. Fig. 1 of Hönig et al. (2014)). They show that a large fraction of their BAT selected sources (12/16; 75%) are in edge-on galaxies or merging/interacting systems, and conclude that low scattering fraction galaxies prefer more edge-on orientations, suggesting that absorption in the host galaxy can play an important role in low scattering AGN. Using a combination of X-ray, optical, and infrared properties, Yamada et al. (2019) also study low scattering fraction AGN, and find that these sources are more often found in merging or post-merger systems, suggesting that these are buried AGN, potentially as a result of the merger.

The first hard X-ray selected catalog from the eROSITA (Predehl et al. 2021a) all-sky survey (eRASS1) was presented (see also Merloni et al. 2024) and analysed in Chapter 4. The hard X-ray selection is less sensitive to obscuration effects than the main, soft (0.2−2.3 keV) selected sample, and can therefore more readily be used to search for obscured Sy2 AGN. In this chapter, a unique sub-population of source identified only in the hard ($2.3 - 5$ keV) band was identified, referred to as hard-only sources (see also Nandra et al. 2024). Some of these sources have spectroscopic redshifts of $z > 0.002$, and are therefore securely classified as AGN, likely obscured through a moderate column density medium ($\sim 10^{23}\,\mathrm{cm}^{-2}$) to absorb the emission below 2.3 keV. By studying their optical shapes derived from the half-light radii, it was found that hard-only sources are more likely to be found in edge-on galaxies than sources with soft X-ray detections (Chapter 4), and that these sources tended to show very low scattering fractions of $< 1\%$. This is due to eROSITA's extremely sensitive soft band; since we cut in significance of the soft X-ray detection, we by construction created a sample where not only is the nuclear light

completely absorbed, but any intrinsic soft emission within the host galaxy is also not present. Spectral modelling of eROSITA data alone did not enable a more detailed study of the properties of the absorber, as larger samples and broader energy ranges would be necessary to explore this in more detail. Nevertheless, the bright fluxes of these interesting sources makes them ideal for follow-up with hard X-ray telescopes to shed light on the competing scenarios: Either the obscuring medium is a uniform sphere that enshrouds the entire inner region (e.g. Yamada et al. 2019), or, the obscuration is associated with galaxy-scale absorbers (see e.g. Goulding et al. 2012; Hönig et al. 2014).

In this chapter, low scattering fraction AGN selected with (Merloni et al. 2012b; Predehl et al. 2021a; Sunyaev et al. 2021) are analysed in more detail, modelling eROSITA and *NuSTAR* (Harrison et al. 2016) data with state of the art absorption models to investigate the nature and structure of the absorption. In Sect. 5.2, the sample construction is discussed, using sources from the first eROSITA all-sky survey. Section 5.3 describes the data analysis technique, including X-ray data processing and host galaxy properties. Section 5.4 presents the spectral models and modelling approach used in this chapter. Results are presented in Sect. 5.5, and a discussion is given in Sect. 5.6. Finally, conclusions are given in Sect. 5.7.

5.2 Sample construction

The primary objective is to construct a sample of low scattering fraction sources which can be studied in order to gain more insight on the torus structure and geometry. To do so, we take advantage of eROSITA's sensitivty to soft X-ray photons, such that sources which show evidence for hard X-ray emission but no corresponding soft X-ray emission are very likely to be obscured AGN with low scattering and no other soft X-ray emission components, as suggested in Chapter 4 and published in (Waddell et al. 2024).

The starting point of the analysis is the eRASS1 hard sample (Chapter 4). The sample contains 5466 X-ray sources which have detection likelihoods of `DET_LIKE_3`>12 in the $2.3-5\,\mathrm{keV}$ band. After removing likely spurious or problematic sources and limiting to only X-ray point sources, Chapter 4 identified a sub-sample of 684 sources which were only detected in the hard band, defined as having detection likelihoods `DET_LIKE_1`<5 and `DET_LIKE_2`<5, where band 1 is $0.2-0.6\,\mathrm{keV}$ and band 2 is $0.6-2.3\,\mathrm{keV}$ (see Merloni et al. 2024, Chapter 4). These sources are likely to be obscured AGN with column densities $\gtrsim 10^{23}\,\mathrm{cm}^{-2}$. Counterparts were identified for all sources and a spectroscopic sample of AGN was presented, including spectroscopic hard-only AGN. Counterparts were found with NWAY (Salvato et al. 2018) based on the procedures laid out in Salvato et al. (2022). Matching was performed using the Legacy Survey Data Release 10 (LS10; Dey et al. 2019), which covers 15342 deg^2 of the sky, mostly excluding the Galactic plane. Sources are matched using a Bayesian algorithm, based on priors developed for the full eROSITA all-sky survey. This matching process yielded 23 hard-only sources with spectroscopic redshifts $z > 0.002$.

In order to also find counterparts for sources which did not lie within the LS10 sky area,

Chapter 4 also presented counterparts from matching to the the Swift-BAT 105 month catalog (Oh et al. 2018), an X-ray survey in the $15-195$ keV band which is relatively unbiased by absorption effects. This yielded six additional hard-only sources with spectroscopic redshifts $z > 0.002$. These sources are of particular interest, as many of these sources have been followed up with *NuSTAR* as part of a dedicated follow-up program, and many have other multiwavelength data which could be used for further comparison.

Combining these six sources with the 23 found with the LS10 match, and including one additional LS10 match for which a spec-z was not identified using the methodology outlined in Chapter 4, but where there is a spec-z in SIMBAD, this forms our golden sample of hard-only AGN, with a total of 30 sources. This is the same sample (with one additional source) presented in Chapter 4, where it was identified that these sources have extremely low scattering fractions and are prime for future study. These sources are then cross-matched with all *NuSTAR* observations using the heasarc[1] browse mission interface, and find seven matches, with a total of 11 individual observations. These are summarized in the top 11 rows of Table 5.1.

Recognizing the need to have *NuSTAR* data for as many sources as possible in order to better investigate the obscurer geometry, a secondary sample of hard, but not hard-only, sources, is defined. This is done by defining a hardness ratio, $HR = (H - S)/(H + S)$ where $\mathrm{H} = 2.3 - 5$ keV counts and $\mathrm{S} = 0.2 - 2.3$ keV counts. This distribution for sources with spec-z values is shown in Fig. 5.1. Examining this distribution, it can be seen that hard-only sources all have $HR > 0$, which is reasonable as the emission should be much stronger above 2.3 keV than below. Applying this cut to all X-ray sources with counterparts and spectroscopic redshifts adds 22 additional sources with good LS10 counterparts, spec-z and $HR > 0$. A further five sources are added with Swift-BAT counterparts and spec-z, which also have $HR > 0$, for a total of 27 additional sources. Cross-matching again with all *NuSTAR* observed sources returns nine additional matches with a total of 15 observations, summarized in the bottom rows of Table 5.1. In total, this gives 16 sources with *NuSTAR* matches. The source MCG-01-24-12 was observed on seven occasions, and the spectrum with the longest observation length is modelled; all other sources have a single observation. These sources all have relatively hard spectra by definition, have known spec-z, are believed to be non-jetted (Chapter 4), and are mostly Seyfert type 2. Given the extreme hardness of these spectra, all are expected to have low scattering fractions of $\lesssim 1\%$ (see Section 5).

To obtain a basic first look at the properties of these sources, Fig. 5.2 shows the luminosity-redshift space occupied by the sample. Sources from the golden sample are shown as red squares, and sources from the supplementary sample are shown as black circles. For reference, sources from the eRASS1 hard-only sample without *NuSTAR* detections are shown with grey crosses. Luminosities are obtained from eROSITA spectral modelling, and the redshifts are those presented in Table 5.1. Most sources are relatively nearby AGN, covering a range of hard X-ray luminosities. Most low-z sources come from the hard-only sample, perhaps because at higher redshifts, even more absorption is necessary to decrease the detection likelihood in the observed $0.2 - 2.3$ keV band. These

[1] https://heasarc.gsfc.nasa.gov/cgi-bin/W3Browse/w3browse.pl

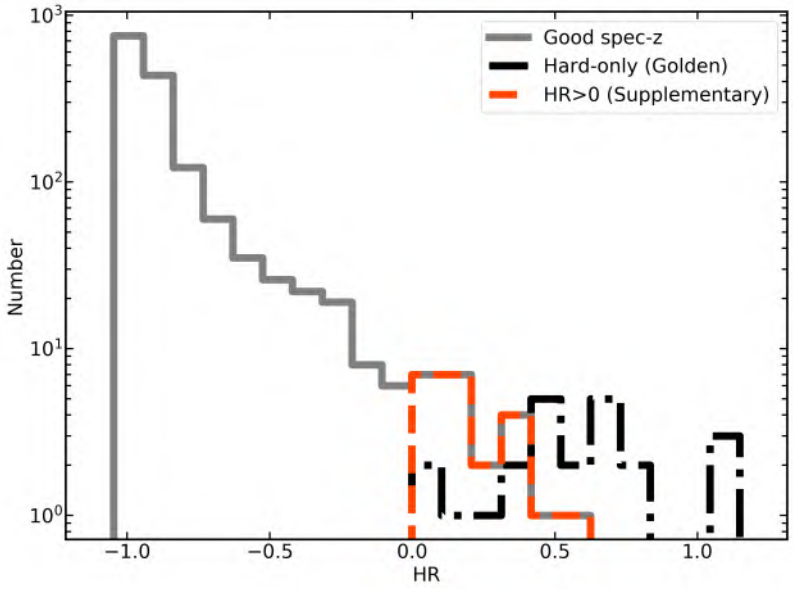

Abbildung 5.1: Hardness ratios shown for all sources with spectroscopic redshifts (gray), hard-only sources (red) and sources in the supplementary sample with $HR < 0$ (black).

sources also extend to very low luminosities, but most have been classified as Seyfert 2 AGN, and other works (e.g. Georgakakis & Nandra 2011) have revealed that AGN remain the dominant population.

5.3 Data processing

5.3.1 eROSITA

eROSITA data processing is described in detail in Section 4 and Section 7 of (Merloni et al. 2024). The task `srctool` is used from the standard eROSITA Science Analysis Software System (eSASS; Brunner et al. 2022), pipeline version 020. Source and background spectra and associated response files are produced using `srctool`, which performs source extraction using a circular region centred on the X-ray source, extracts a background region from a surround annulus, removing nearby sources, and calculates associated RMF and ARFs. Spectra are extracted centred on the positions from the eRASS1 hard catalog released in Merloni et al. (2024) and analysed in detail in Chapter 4. Data from all seven telescope modules (TM) are summed, and spectra are modelled in the energy range $0.3 - 8\,\mathrm{keV}$.

5.3.2 NuSTAR

Data from *NuSTAR* allows probing obscurer geometry and inclination, as discussed in section 2. This chapter uses `heasoft` v. 6.32.1 and follow standard *NuSTAR* data processing procedure and using the latest `caldb` v. 20220510. Raw event files are first pre-processed

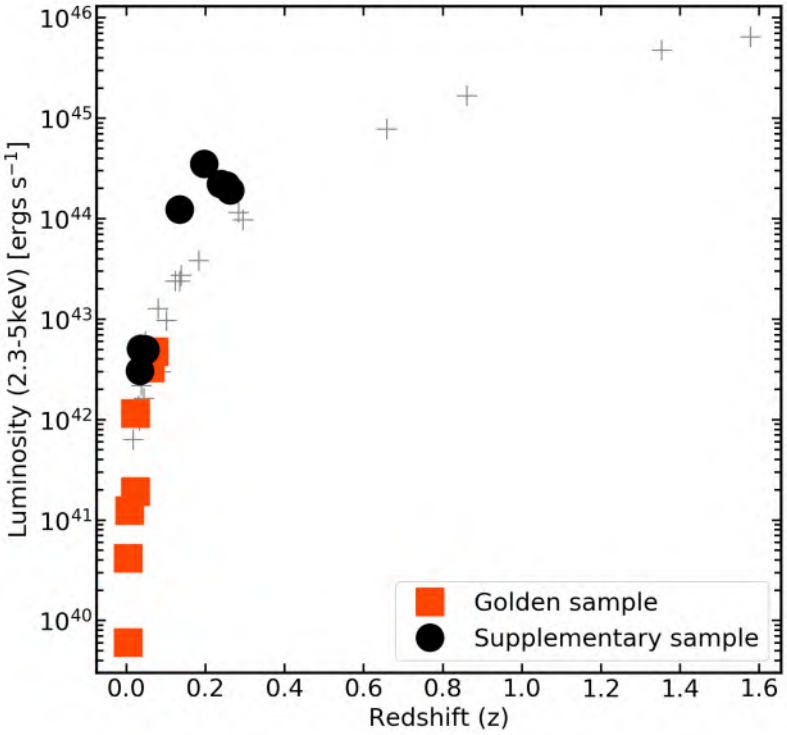

Abbildung 5.2: Luminosity redshift distribution for the golden sample (red squares) and supplementary sample (black circles), showing only sources with *NuSTAR* data, as studied in this chapter. Sources from the eRASS1 hard-only sample without *NuSTAR* detections are shown with grey crosses.

Tabelle 5.1: List of *NuSTAR* data sets used in this chapter. Column (1) gives the source name, column (2) gives the eROSITA name, and columns (3) and (4) give the SIMBAD RA and DEC(J2000). Column (5) gives the Galactic column density in units of $\times 10^{20}$ cm^{-2}, and column (6) gives the redshift used for spectral fitting. Column (7) gives the *NuSTAR* OBSID, and column (8) gives the exposure time of FPMA in units of kiloseconds.

(1) Name	(2) eROSITA Name (1eRASS)	(3) RA J2000 [°]	(4) Dec J2000 [°]	(5) N_H ($\times 10^{20}$ cm^{-2})	(6) Redshift (z)	(7) ObsID	(8) Exp [ks]
Golden sample							
ESO 119-8	J044856.6-573939	72.2364	-57.6592	1.29	0.0229	60061052002	26
ESO 121-G6	J060730.3-614819	91.8744	-61.8076	4.77	0.0040	60201063002	50
Fairall 272	J082300.9-045602	125.7546	-4.9349	5.39	0.0222	60061080002	24
IC 2461	J091958.1+371123	139.9918	37.1913	1.12	0.0075	60061353002	33
IRAS 00521-7054	J005356.1-703805	13.4840	-70.6345	5.78	0.0696	60301029002	73
						60301029004	122
						60301029006	74
LEDA 549777	J064038.3-432119	100.1583	-43.3558	7.06	0.0605	60061070002	22
NGC 6300	J171658.8-624911	259.2478	-62.8206	10.6	0.0037	60061277002	18
						60261001002	20
						60261001004	24
Supplementary sample							
2MASX J06411806 +3249313 (J06411806)	J064118.6+324928	100.3252	32.8254	23.6	0.047	60061071002	18
2MASX J14104482 -4228325 (J14104482)	J141044.5-422835	212.6867	-42.4758	7.26	0.0339	60160571002	22
CGCG 041-020	J120058.2+064830	180.2414	6.8064	1.23	0.0360	60061218002	19
ESO 103-35	J183820.4-652532	279.5848	-65.4275	7.22	0.1347	60301004002	44
ESO 197-27	J021053.0-494153	32.7189	-49.6985	1.79	0.0481	60061020002	18
IGR J09523-6231	J095218.9-623228	148.0841	-62.5430	34.1	0.252	60160376002	28
LEDA 93974	J104022.7-462521	160.0939	-46.4238	18.2	0.2392	60061202002	21
MCG-01-24-12	J092045.9-080323	140.1927	-8.0561	23.6	0.1961	60061091010	15
NVSS J122810-092702	J122809.7-092702	187.0420	-9.4508	2.91	0.2624	60464096002	21

using the tool `nupipeline` v. 0.4.9. Source photons are extracted from a 60 arcsecond circular region centered on the X-ray source position and background photons from a 120 arcsecond circular region off-center from the source. The task `nuproducts` was then run to extract source spectra, background spectra, and response files for the two focal plane modules (FPM), FPMA and FPMB. As is recommended, due to small calibration differences between the two FPMs, spectral data from FPMA and FPMB modelled jointly, with a cross-normalization left free to vary, typically recovered to be between $1 - 10\%$. Spectral modelling is performed in the $3 - 60$ keV band to avoid high background contamination above 60 keV in most spectra.

5.3.3 Host-galaxy properties

To connect the properties of the X-ray emission and absorption to the larger galaxy scale, sources are separated based on host galaxy classification and orientation. This is of particular interest for testing if some of the absorption component comes from the host-galaxy disc, rather than from a obscurer. Key host galaxy properties are tabulated in Table 5.2. Host galaxies are separated into spirals and ellipticals using classifications from NED[2] (NASA/IPAC Extragalactic Database) and SIMBAD[3] where available, and spiral galaxies are further classified into face-on and edge-on based on the ratio of semi-major and semi-minor axes (b/a) of the half light radius of the galaxy from NED and/or LS10 where available, as well as through visual inspection. A cut-off of $b/a < 0.4$ (corresponding to an inclination angle $i > 65^{\circ}$) is applied for edge-on galaxies to ensure that they sufficiently inclined that it is plausible that the galactic disc could obscure the X-ray emission from the central region, whereas sources with $b/a > 0.4$ ($i < 65^{\circ}$ are classified as face-on.

For four sources; LEDA 549777 from the golden sample, and 2MASX J14104482-4228325, NVSS J122810-092702 and IGR J09523-6231 from the secondary sample, no classification is immediately available. Studying the model residuals as well as the derived Sersic profile for LEDA 549777 obtained from LS10, a Sersic index of ~ 3 is found, and there is no visual evidence for spiral arms, only for an outer disc, so this source is likely to be an elliptical or lenticular galaxy. NVSS J122810-092702 is also in the LS10 area, but is a bright radio source at ($z = 0.2624$), so it is difficult to infer any host galaxy properties. The source is assumed to also be elliptical or lenticular as a bright radio source which looks green-red in the LS10 images.

For the remaining 12 sources, classification is more straightforward. From the golden sample, three sources are classified as face-on spiral galaxies, three sources as edge-on spiral galaxies, and one source as an elliptical galaxy. Adding in the supplementary sample, this yields five face-on spiral galaxies, three edge-on spiral galaxies, six elliptical or lenticular galaxies, and two sources with unknown class. S0 galaxies lack strong spiral arms and a clearly defined galactic plane. Therefore it is difficult to discuss an inclination angle based on b/a value, and these are discussed separately. Images for each galaxy are shown in

[2]The NASA/IPAC Extragalactic Database (NED) is funded by the National Aeronautics and Space Administration and operated by the California Institute of Technology: https://ned.ipac.caltech.edu/

[3]https://simbad.cds.unistra.fr/simbad/sim-fcoo

Tabelle 5.2: List of sources and classifications. Column (1) gives the source name, column (2) gives the ratio of semi-major to semi-minor axes (b/a), column (3) gives the Hubble class, and column (4) gives the Seyfert class. Column (5) gives the classification as will be used in this analysis, and column (6) gives any relevant notes on the spectral classification.

(1)	(2)	(3)	(4)	(5)	(6)
Name	b/a	Hubble class	Seyfert class	Classification	Notes
Golden sample					
ESO 119-8	0.98	SBa	Sy2	face-on spiral	
ESO 121-G6	0.20	Sc	Sy2	edge-on spiral	
Fairall 272	0.42	SA	Sy2	face-on spiral	borderline orientation
IC 2461	0.18	Sb/Sbc	Sy2	edge-on spiral	
IRAS 00521-7054	0.8	E-S0	Sy2	elliptical	alternatively lenticular
LEDA 549777	N/A	N/A	Sy2	unknown	likely elliptical
NGC 6300	0.66	SBb	Sy2	face-on spiral	
Supplementary sample					
J06411806	0.84	S0	Sy2	lenticular	
J14104482	0.6	N/A	Sy2	unknown	
CGCG 041-020	0.62	S0	Sy2	lenticular	
ESO 103-35	0.46	S0a	Sy2	face-on spiral	borderline lenticular
ESO 197-27	0.38	Sc	Sy2	edge-on spiral	
IGR J09523-6231	N/A	N/A	Sy1.9/Sy2	unknown	
LEDA 93974	0.60	S0	Sy2	lenticular	
MCG-01-24-12	0.96	SABc	Sy2	face-on spiral	
NVSS J122810-092702	N/A	N/A	AGN	unknown	likely elliptical, radio source

Fig. 5.3, with sources separated by classification to allow for visual comparison between classes.

5.4 Spectral models

There are a variety of physical models (e.g. see Buchner et al. 2019, for a review) which can be applied to X-ray data in order to model the observed cold absorption, all invoking different physical absorber structures and ranges of properties to explain the observed spectral shape. Here, results from three models will be compared; `ztbabs` (Wilms et al. 2000) which models absorption originating in the ISM, `borus02` (`borus`; Baloković et al. 2018, 2019) which models reflection and absorption in a uniform spherical absorber with polar cutouts, and `UXClumpy` (Buchner et al. 2019) which models absorption in clumpy absorber with a Compton-thick centre. Both `borus` and `UXClumpy` allow the user to input the viewing angle and the height of the obscurer, allowing us to test the effect of obscurer geometry on the observed spectral shape and determine if these effects will be observable. For the purpose of visual comparison in this section, models are always compared using a column density of $10^{23}\,\mathrm{cm}^{-2}$ and a photon index $\Gamma = 1.9$, in agreement with the typical values from eROSITA analysis in presented in Chapter 4. Photon indices are left free to vary when performing real spectral modelling later in this Chapter.

The model `ztbabs` (Wilms et al. 2000) combines X-ray absorption due to the gas phase, grain phase, and the molecules in the ISM of the host galaxy. The model also includes a redshift. The absorption is assumed to originate in the host galaxy, so the redshift is set to that of the galaxy. The column density is left free to vary, and is given a log-uniform prior between $10^{20}\,\mathrm{cm}^{-2}$ and $10^{25}\,\mathrm{cm}^{-2}$, which covers a broad range of possible obscuration column densities. The median of the prior on the photon indices is set to 1.95, in agreement with typical eROSITA photon indices (Nandra & Pounds 1994a; Liu et al. 2022b; Waddell et al. 2024; Nandra et al. 2024, e.g.).

It is of course also possible that the absorption originates not in the host galactic plane, but rather nearer to the central region, in the torus. In order to model this, this chapter makes use `UXClumpy` (Buchner et al. 2019), which models the transmitted and reflected components of X-rays passing through a clumpy structure. There are several interesting parameters which can be explored using this model. Specifically, the height or covering angle of the obscurer, TORsigma, which indicates how much of the central region is covered by the obscurer and the inclination angle θ_{inc}, in addition to the photon index of the incident X-ray spectrum, Γ, and the the column density of the obscurer, $\mathrm{N_H}$. There is also the possibility to include a high energy cut-off of the power law component (which is set to 300 keV), as well as to include a Compton thick ring of clouds surrounding the central region with some defined covering fraction, CTKcover. This inner ring consists of clouds which are spherical in shape and where the covering factor ranging from 0.0-0.6 dictates theradius and number of clouds (see Buchner et al. 2019).

X-ray spectra from different geometries of this model are shown in Fig. 5.4. Models are shown in the $3 - 60$ keV range, which can readily be studied with *NuSTAR*. The top

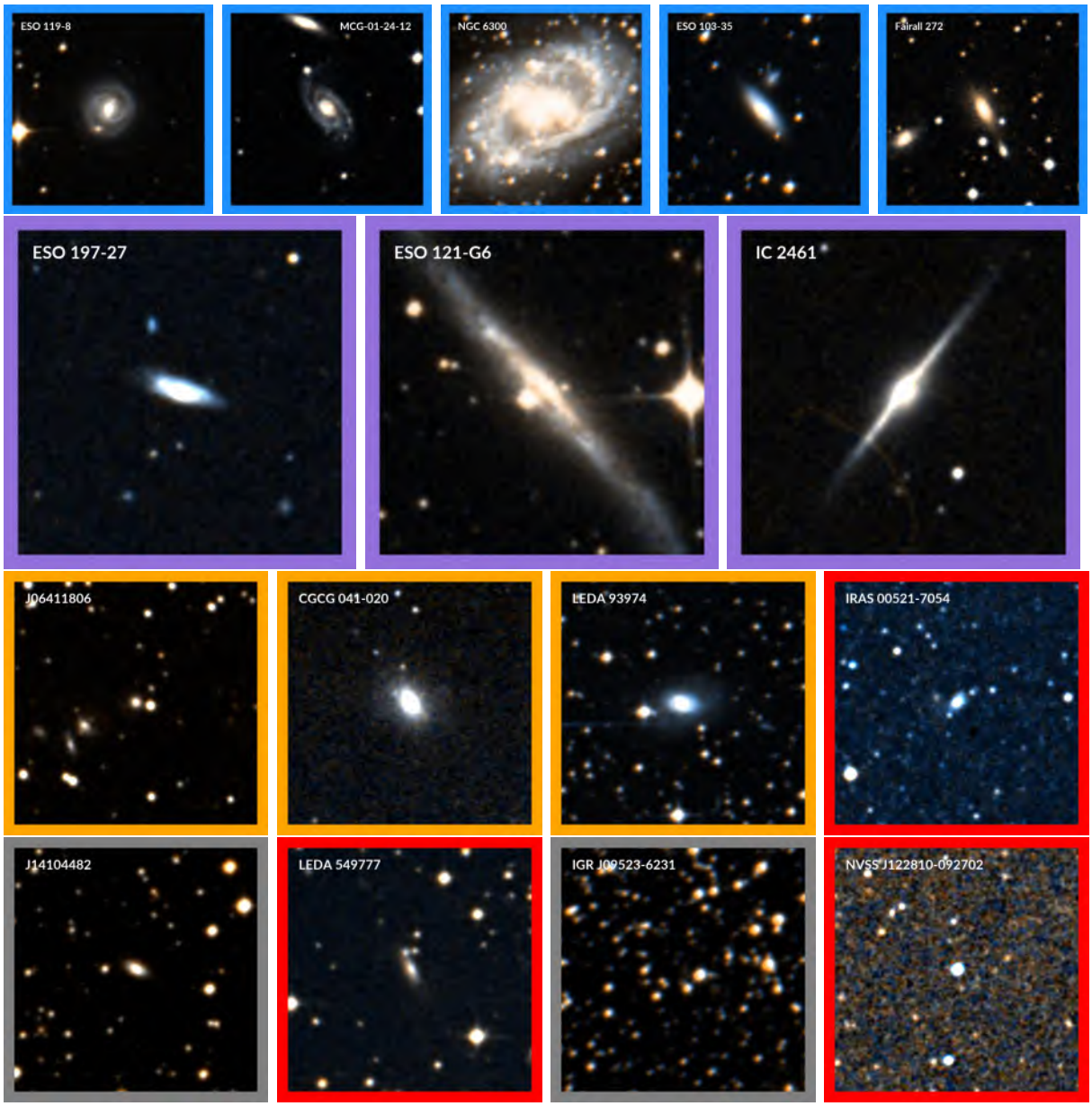

Abbildung 5.3: Optical image cutouts for the 16 sources in this chapter, with the host galaxy at the centre of each image. All cutouts are taken from coloured composites from DSS2, for ease of comparison, since all sources are not present in LS10. All have a 0.05 degree FOV. Source names are indicated at the top of each image. Border colours correspond to the source class; edge-on spiral galaxies are purple, face-on spiral galaxies are blue, lenticular galaxies are orange, elliptical galaxies are red, and unknown class galaxies are in grey. Sources are organised both by class and by decreasing b/a value. Note that IGR J09523-6231 is extremely dusty and cannot be readily seen in optical images; the source is visible in infrared bands.

Tabelle 5.3: List of parameters in each model as well as priors used in spectral fitting. Column (1) gives the relevant physical parameter, columns (2) and (3) give the parameter name and prior for `ztbabs`, columns (4) and (5) give the parameter name and prior for `borus02`, and columns (6) and (7) give the parameter name and prior for `UXClumpy`. Parameters which are only relevant for fitting with either eROSITA or *NuSTAR* are listed separately.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Model	`ztbabs`		`borus02`		`UXClumpy`	
Parameter	Name	Prior	Name	Prior	Name	Prior
All fitting						
Column density	nH	$20-25$	log(nH)	$22-22.5$	log(nH)	$20-25$
$\log(N_H/cm^{-2})$		lognorm		lognorm		lognorm
Photon Index	PhoIndex	1.95, $\sigma=0.15$	PhoIndex	1.95, $\sigma=0.15$	PhoIndex	1.95, $\sigma=0.15$
(Γ)		Gaussian		Gaussian		Gaussian
Norm (photons	log(norm)	$-10-10$	log(norm)	$-10-10$	log(norm)	$-10-10$
s^{-1} cm^{-2} keV^{-1})		lognorm		lognorm		lognorm
Torus height	-	-	thTor	$0-84$	TORsigma	$0-84$
(°)	-	-		uniform		uniform
Inclination	-	-	thInc	$19-87$	θ_{inc}	$18.2-87.1$
(°)	-	-		uniform		uniform
Compton thick covering	-	-	-	-	CTKcover	$0.0-0.6$
	-	-	-	-		uniform
High-energy cut	-	-	Ecut	300 keV	Ecut	300 keV
(keV)	-	-		fixed		fixed
Iron abundance	-	-	AFe	1.0	-	-
(solar)	-	-		fixed	-	-
Redshift	z	fixed	z	fixed	z	fixed
eROSITA only						
MW column density	nH	fixed	N/A	N/A	nH	fixed
Scattering	-	-	N/A	N/A	log(factor)	-5– -1
fraction	-	-	N/A	N/A		lognorm
eROSITA background	norm	1.0, $\sigma=0.2$	N/A	N/A	norm	1.0, $\sigma=0.2$
		Gaussian	N/A	N/A		Gaussian
***NuSTAR* only**						
Cross-norm	factor	1.0, $\sigma=0.1$	factor	1.0, $\sigma=0.1$	factor	1.0, $\sigma=0.1$
		Gaussian		Gaussian		Gaussian
FPMA background	norm	1.0, $\sigma=0.2$	norm	1.0, $\sigma=0.2$	norm	1.0, $\sigma=0.2$
		Gaussian		Gaussian		Gaussian
FPMB background	norm	1.0, $\sigma=0.2$	norm	1.0, $\sigma=0.2$	norm	1.0, $\sigma=0.2$
		Gaussian		Gaussian		Gaussian

panel, compares extreme obscurer geometries; a near-spherical (TORsigma = 84°) obscurer viewed face-on ($\theta_{\rm inc}$ = 0°), a near-spherical obscurer viewed edge-on ($\theta_{\rm inc}$ = 80°), and a wedge-shaped obscurer (TORsigma = 20°) viewed edge-on (see also Figure 3 of Buchner et al. 2019). The resulting spectrum for a wedge-shaped obscurer viewed face-on is not shown, as this is extremely unlikely to result in a significantly obscured X-ray spectrum, especially one with a low scattering fraction as those studied in this chapter. All three realisations differ significantly above ~ 5 keV, beyond the spectral range well-probed with eROSITA. The spherical absorber viewed face-on results in a strong Fe Kα line, a steep rise between $10 - 15$ keV, and a strong Compton hump. When a spherical absorber is viewed edge-on, more photons are obscured by the Compton thick ring (here with a CTKcover of 0.4), giving a slightly weaker Compton hump. In contrast, a wedge-shaped absorber viewed edge-on with the same CTKcover produces a much weaker Fe Kα line, a shallow rise between $10 - 15$ keV, and a less pronounced Compton hump. For the purpose of the analysis in this Chapter, the presence of a Compton thick ring is permitted for all sources, since this parameter is shown to primarily affect the spectral shape above 10 keV (see Fig. 5.5). Such differences should be readily detectable in the *NuSTAR* spectra studied in this chapter.

Fig. 5.5 shows the effect on the observed spectral shape when changing the covering factor of the Compton thick component, CTKcover, both for a near-spherical (TORsigma = 84°) obscurer viewed face-on ($\theta_{\rm inc}$ = 0°), and a thin obscurer (TORsigma = 20°) viewed edge-on ($\theta_{\rm inc}$ = 80°). The Fe Kα and Compton hump strengths are always stronger for the models showing a spherical obscurer, and unsurprisingly, increase in strength with increasing CTKcover due to the increasing number of visible Compton scattering events off of Compton thick material. In the thin wedge model versions, there is a less significant change in the spectral shape, with differences more visible in the Fe Kα band. Here the normalization decreases slightly with increasing CTKcover, since more reflected photons will have been blocked by the Compton thick region. No matter the value of CTKcover, it is still possible to distinguish the two obscurer geometries; however it should be cautioned that there is likely some degeneracy between viewing angle and CTKcover, which may be particularly relevant for exploring the inclination of the obscurer.

In this particular regime of studying low scattering fraction sources, it has also been suggested that the obscurer is not clumpy, but rather uniform and possibly spherical. In order to test for this, this chapter uses `borus` (Baloković et al. 2018, 2019). This model employs an absorber structure which is a sphere comprised of uniform gas with polar cutouts. The free parameters include the column density and photon index, the obscurer opening angle thTor (which is defined such that thTor = 0° is a spherical absorber and thTor = 90° would be a infinitely thin absorber), and the viewing angle, thInc. Since the table models only include the reprocessed component, the line-of-sight component must be added using a separate method, in this case by adding an additional power law component absorbed with `tbabs`. The resulting spectral shapes are shown in Fig. 5.6. Contrary to `UXClumpy` where there are significant differences in the spectral shape between different obscurer models, such differences are not observed with borus, due to the simple and smooth geometry employed. It is therefore unlikely that one would be able to distinguish

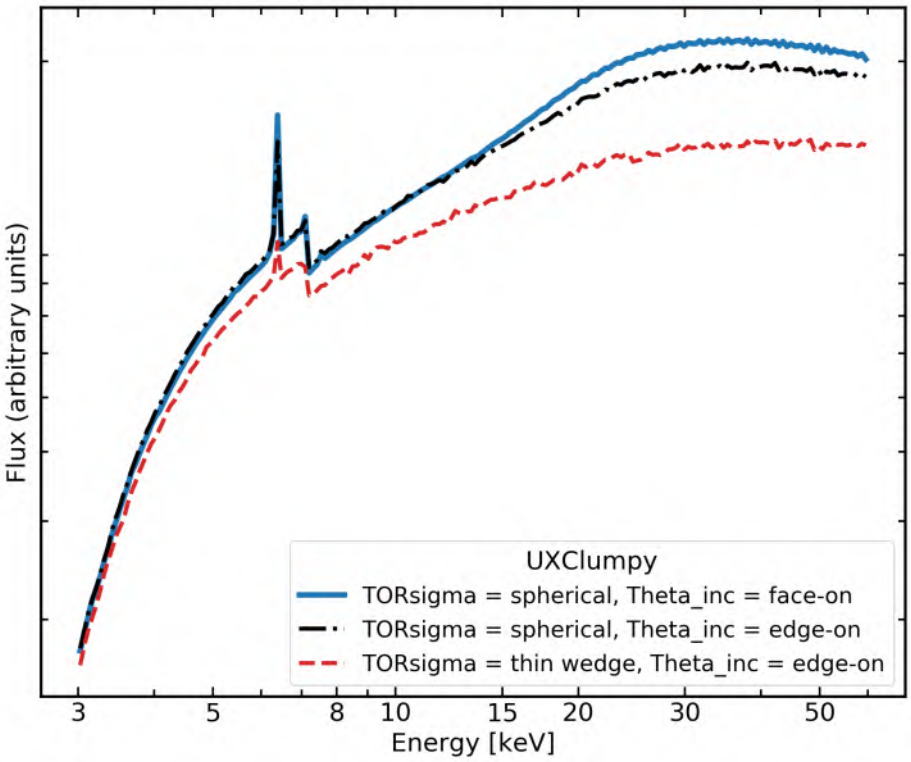

Abbildung 5.4: Comparison between X-ray spectra from different obscurer geometries in `UXClumpy`, showing a comparison between three extreme geometries and orientations; a near-spherical (TORsigma = 84°) obscurer viewed face-on ($\theta_{\rm inc} = 0°$) (blue solid line), a near-spherical obscurer viewed edge-on ($\theta_{\rm inc} = 80°$) (black dash-dot line), and a wedge-shaped obscurer (TORsigma = 20°) viewed edge-on (red dashed line). A column density of $10^{23}\,{\rm cm}^{-2}$ and a photon index $\Gamma = 1.9$ are used.

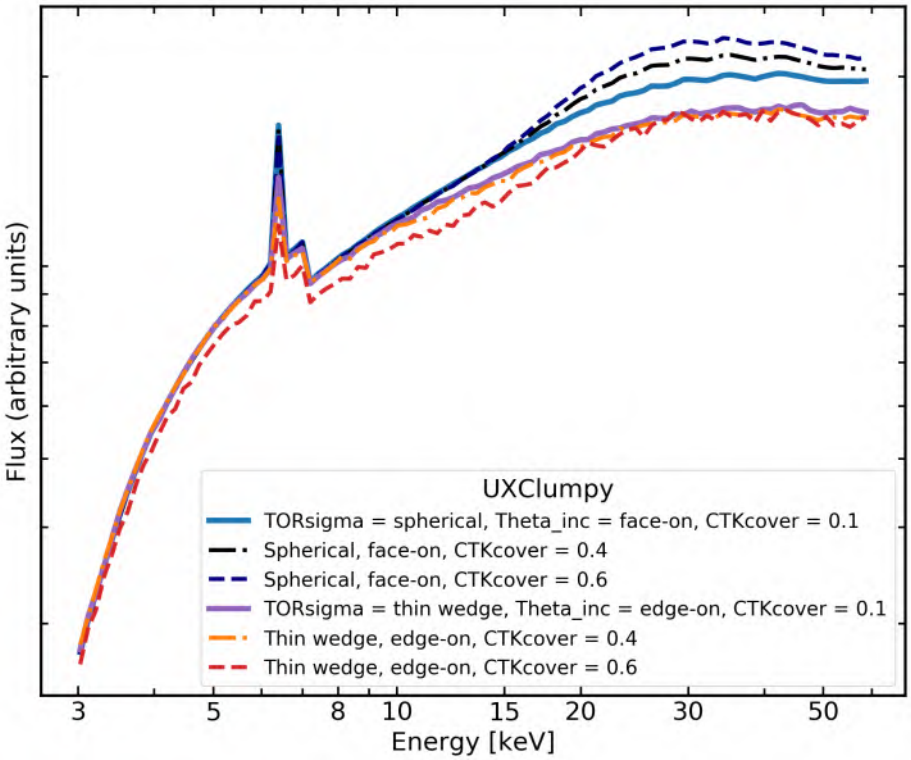

Abbildung 5.5: Comparison between X-ray spectra from different obscurer geometries in `UXClumpy`, showing two obscurer geometries, with varying column densities; a spherical (TORsigma = 84°) obscurer viewed face-on ($\theta_{\rm inc}$ = 0°) with Compton thick covering fractions of 0.1, 0.4 and 0.6 shown as a blue solid line, black dash-dot line, and navy dotted line respectively, as well as a wedge-shaped (TORsigma = 20°) obscurer viewed edge-on ($\theta_{\rm inc}$ = 80°) with Compton thick covering fractions of 0.1, 0.4 and 0.6 shown as a purple solid line, orange dash-dot line, and red dashed line respectively.

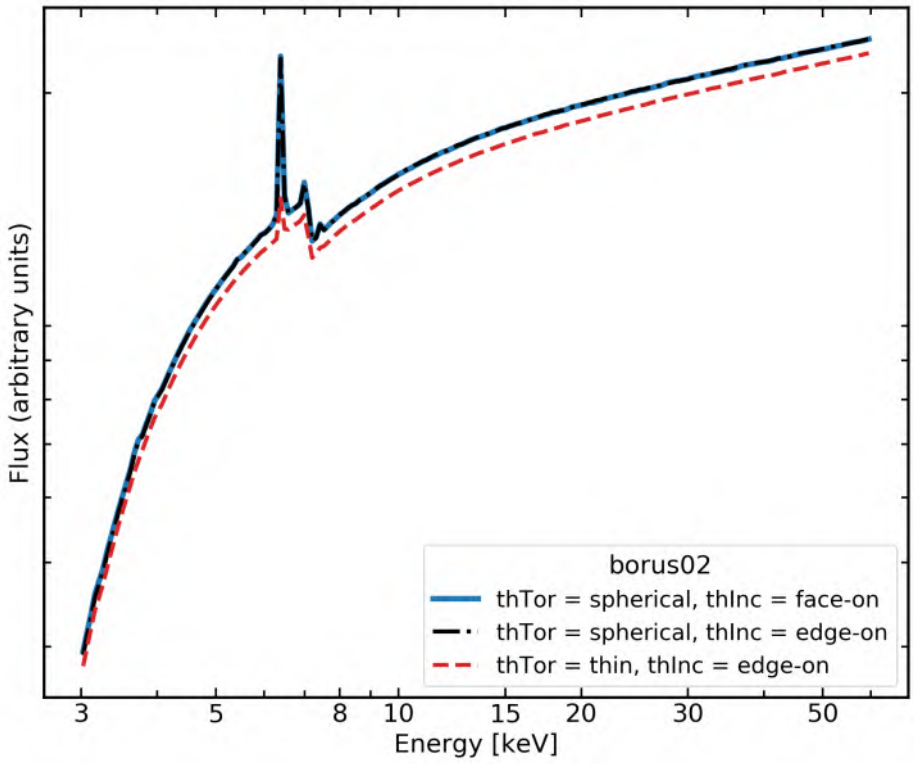

Abbildung 5.6: Comparison between X-ray spectra from different geometries in `borus`; a near-spherical obscurer (thTor = 0°) viewed face-on (thInc = 20°) (blue solid line), a near-spherical (thTor = 0°) obscurer viewed edge-on (thInc = 80°) (black dash-dot line), and a wedge-shaped obscurer (thTor = 80°) viewed edge-on (thInc = 80°) (red dashed line). A column density of $10^{23}\,\mathrm{cm}^{-2}$ and a photon index $\Gamma = 1.9$ are used.

different absorber geometries using this model.

To demonstrate the differences in spectral shapes produced by these models, Fig. 5.7 compares `ztbabs`, `borus` and `UXClumpy`, all compared to an unabsorbed power law with $\Gamma = 1.9$. The strong Compton hump feature in `UXClumpy` can be explained by the Compton thick ring found outside of the line of sight. The geometry assumed for both `UXClumpy` and `borus` is the same, using a obscurer with a moderate opening angle of $\sim 45°$ viewed at an intermediate viewing angle. The much stronger Compton hump feature is due to the addition of the Compton thick ring feature, which has a covering factor of 0.4.

In general, the analysis suggests that the geometry changes in `UXClumpy` result in much more dramatic spectral variation, likely due both to the inclusion of the Compton thick ring component and the clumpy geometry employed. Using `UXClumpy`, different obscurer geometries lead to different model realisations, with differences easily visible in the $5 - 70\,\mathrm{keV}$ region, which is observable with *NuSTAR*. This will be tested using data from low scattering fraction AGN, which should offer a relatively pure look at the reflection component.

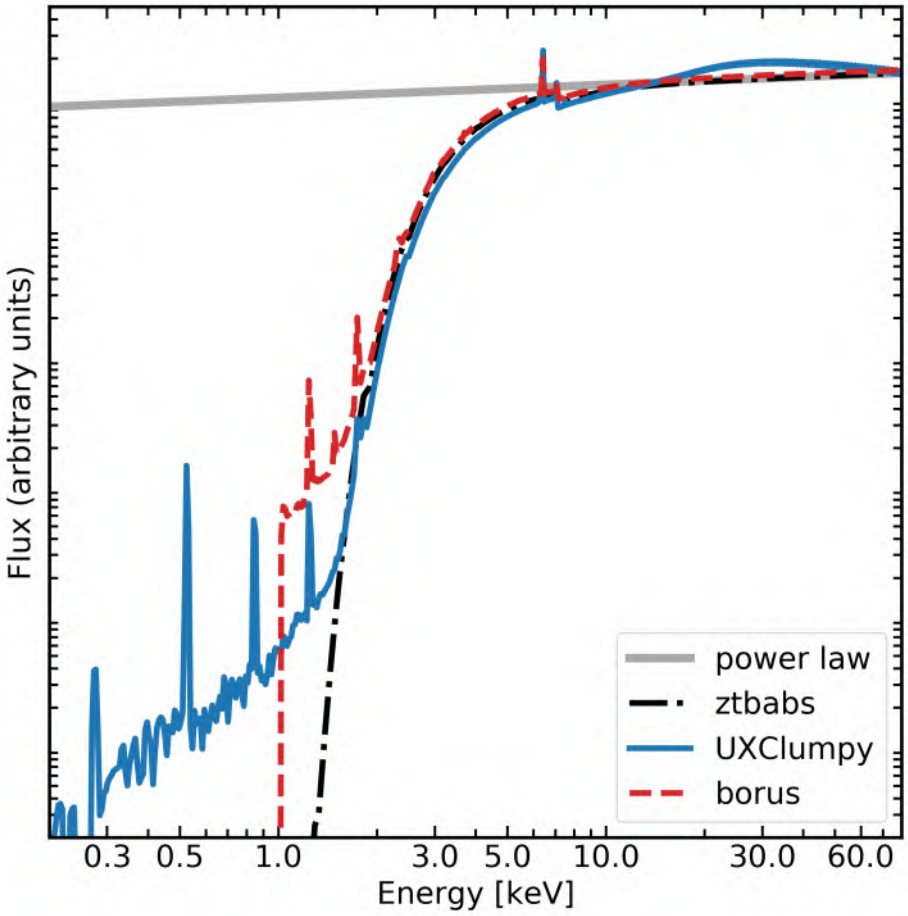

Abbildung 5.7: Comparison between X-ray spectra from `ztbabs` (black dash-dot line), `UXClumpy` (blue solid line) and `borus` (red dashed line), using the same column density of $10^{23}\,\mathrm{cm}^{-2}$, a typical photon index of $\Gamma = 1.9$, and the same obscurer geometry for both `UXClumpy` and `borus`. All are compared to an unabsorbed power law with $\Gamma = 1.9$ (grey solid line). A column density of $10^{23}\,\mathrm{cm}^{-2}$ and a photon index $\Gamma = 1.9$ are used.

Spectral fitting is performed using `XSPEC v. 12.13.1` (Arnaud 1996), implemented in `python` as `PyXSPEC`. In Table 5.3, we present an overview of all the parameters included in each of the three models discussed here, as well as of all of the parameter ranges and priors used in the spectral fitting. In order to effectively explore the full model parameter space, this chapter employs Bayesian X-ray Analysis (BXA; Buchner et al. 2014), which uses nested sampling algorithms (ultranest; Buchner 2021) in order to perform parameter estimation using Bayesian evidence. This is especially important for the eROSITA data, which have typically low counts and therefore require careful treatment. The Bayesian evidence obtained from each model can also be used to perform model comparison, in order to select the best model for each source. BXA also allows for priors to be placed on each parameter, and provides posterior distributions for each parameter, so that best-fit parameters can be more accurately examined. All models applied to the eROSITA spectra are modified by a `tbabs` component, using the column densities from Willingale et al. (2013). Spectroscopic redshifts are taken from Chapter 4 unless otherwise indicated.

The background spectra are modelled using the principal component analysis (PCA) approach described in Simmonds et al. (2018) and used for eROSITA data modelling in Chapters 3 and 4, among other works. This technique uses a PCA decomposition to determine the overall background shape of a typical source for each instrument, then fit to each background spectrum individually. Additional Gaussian components are then added for each background spectrum, and the fit improvement is tested using the Akaike Information Criterion (AIC). The additional Gaussians are added until the AIC no longer decreases, so no further information is gained. The background model is then applied to the source spectrum, using a normalization factor and allowing it to vary, using a Gaussian prior centered at 1 with a width of 0.2. The physical absorption models are then applied to the data. For other parameters, broad uniform or log-normal priors are applied as appropriate, covering the entire parameter space permitted by the model so as not to heavily bias the posteriors.

One important choice made for spectral analysis in this chapter is to model eROSITA and *NuSTAR* data separately, rather than performing joint modelling. This decision has two important consequences; first, no assumption is made regarding spectral or flux variability between the eROSITA and *NuSTAR* epochs, allowing for a more specific modelling of the spectral shape at each epoch. This also leads to more conservative error bar estimates, since the error bars would likely be reduced by including data from both instruments. To ensure that this choice did not significantly bias the results, we randomly selected a source and performed join eROSITA and *NuSTAR* modelling, and found no major changes in best fit parameter values, and only minor improvement in the constraining power. Future work should make use of simultaneous soft and hard X-ray coverage to more accurately and simultaneously model the scattering fractions and absorber parameters to verify the results reported in this chapter.

5.5 Results

5.5.1 Bayesian evidence comparison

A key advantage to using BXA is the ability to perform model comparison in a straightforward manner; by comparing the Bayesian evidence obtained from each model, the model with the lowest evidence can easily be selected and t can be determined that it is the preferred model (see also Waddell et al. 2024). Crucially, if the Bayesian evidence is lower for `UXClumpy`, one can infer that the data prefer a clumpy torus. We therefore first compare Bayesian evidence obtained for each model as applied to the eROSITA data, to see if there is any preference for one model. Since eROSITA is extremely sensitive to soft X-ray photons, any minor differences in spectral shapes below 2 keV should be revealed with eROSITA (see Fig. 5.7). Examining the golden sample, 5/7 sources prefer `UXClumpy`, albeit with small changes in the Bayesian evidence of $\Delta ln(Z)$ of $\sim 0.5 - 2.5$. Of the two that prefer `ztbabs`, LEDA 549777 and Fairall 272, both show small $\Delta ln(Z)$ of ~ 0.7, and both have b/a values which classify them as edge-on galaxies. This could suggest that the absorber in edge-on sources is more likely to be well described by a simple ISM absorption model than for face-on sources, where more complex absorber geometries and properties must be invoked.

We then move on to examine Bayesian evidence obtained from each model as applied to the *NuSTAR* data; these values are even more informative, and are given in Table 5.4. It is found that `UXClumpy` is the preferred model for 13/16 sources; one source (IC 2461) prefers `borus` and two sources (ESO 121-G6 and NVSS J122810-092702) show a preference for `ztbabs`. Typically, there is a large fit improvement when using `UXClumpy` over `borus`, often with $\Delta ln(Z) > 100$, suggesting a strong preference for a clumpy torus. In addition to looking at $\Delta ln(Z)$ for individual sources, the behaviour of the sample can also be studied by multiplying the Bayesian evidence Z (e.g. by summing $ln(Z)$; Buchner et al. 2014; Baronchelli et al. 2018). When summing the $ln(Z)$ evidence for each model and subtracting from the summed $ln(Z)$ the best fit model (last line of Table 5.4), it is even clearer that the clumpy obscurer geometry is preferred by most sources, followed by `ztbabs` and finally `borus`. Interestingly, both IC 2461 and ESO 121-G6 are edge-on spiral galaxies, the two most edge-on sources in our sample with b/a values of < 0.2, so the fact that they do not show strong evidence for absorption in a clumpy obscurer is immediately of interest.

5.5.2 Comparing photon indices between spectral models

Next, it is of interest to examine the individual parameters for each source. Key model parameters for `UXClumpy` fits are given in Table 5.5, since we have established that this is the preferred model for almost all sources. However, since the analysis obtains good constraints on the coronal photon indices for each model, so it is of interest to compare the values measured by each model to check for agreement. Fig. 5.8 shows the distributions of the mean Γ for each model. Histograms of photon indices obtained with `UXClumpy` are shown in black, photon indices obtained with `borus` are shown in dark red, and photon indices

Tabelle 5.4: Bayesian evidence $-ln(Z)$ returned from BXA from the spectral fitting for each model applied to the *NuSTAR* data. Column (1) gives the source name, columns (2), (3) and (4) give the evidence for `UXClumpy`, `zTBabs` and `borus02` respectively. Columns (5) and (6) give $\Delta ln(Z)$ ($-ln(Z_{\rm M1})$-$-ln(Z_{\rm M2})$) values, where a negative value indicates a fit improvement of model 2 over model 1. Column (7) gives the best-fit model for each source. The bottom row shows the normalized full sample evidence.

(1) Name	(2) -ln(Z) `UXClumpy`	(3) -ln(Z) `zTBabs`	(4) -ln(Z) `borus02`	(5) $\Delta ln(Z)$ `UXClumpy` - `zTBabs`	(6) $\Delta ln(Z)$ `UXClumpy` -`borus02`	(7) Best model
Golden sample						
ESO 119-8	-2515.18	-2515.74	-2532.64	0.56	17.46	`UXClumpy`
ESO 121-G6	-2888.12	-2887.17	-2948.22	-0.95	60.10	`ztbabs`
Fairall 272	-2887.28	-2895.11	-3013.16	7.83	125.88	`UXClumpy`
IC 2461	-2883.15	-2889.93	-2882.95	6.78	-0.20	`borus`
IRAS 00521-7054	-3092.50	-3109.16	-3118.64	16.66	26.14	`UXClumpy`
LEDA 549777	-2682.22	-2683.68	-2695.49	1.46	13.27	`UXClumpy`
NGC 6300	-2800.84	-2851.22	-2997.54	50.38	196.7	`UXClumpy`
Supplementary sample						
J06411806	-2664.91	-2666.17	-2703.35	1.26	38.44	`UXClumpy`
J14104482	-2525.22	-2527.04	-2530.15	1.82	4.93	`UXClumpy`
CGCG 041-020	-2719.41	-2723.52	-2744.65	4.11	25.19	`UXClumpy`
ESO 103-35	-3069.53	-3218.02	-3692.04	148.49	622.51	`UXClumpy`
ESO 197-27	-2370.73	-2377.50	-2382.05	6.77	11.32	`UXClumpy`
IGR J09523-6231	-2695.45	-2699.14	-2707.42	3.69	11.97	`UXClumpy`
LEDA 93974	-2515.61	-2525.44	-2521.28	9.83	5.67	`UXClumpy`
MCG-01-24-12	-2456.97	-2866.58	-2496.71	409.61	39.74	`UXClumpy`
NVSS J122810-092702	-2498.61	-2495.11	-2512.78	-3.5	14.17	`ztbabs`
Normalized sum	0.0	664.81	1213.35	664.81	1213.35	`UXClumpy`

obtained with `zTBabs` are shown in orange. Errors on the photon indices are typically of the order of ± 0.06, and the spread of mean values is ~ 0.1.

It can also be see that the photon indices are generally steeper when using `UXClumpy`, slightly flatter when using `zTBabs`, and much flatter when using `borus`; these sample means differ by ~ 0.3 between `UXClumpy` and `borus`, suggesting a much more dramatic effect. When using `UXClumpy`, the preferred spectral model for most sources, a mean value of $\Gamma \sim 1.88$ is obtained, which is similar to the mean obtained for larger samples of type-1 AGN (e.g. Lanzuisi et al. 2013; Waddell et al. 2019), and slightly steeper than values typically found in obscured AGN. In contrast, using a simple absorber model `zTBabs`, a median $\Gamma \sim 1.74$ is recovered, and using `borus`, a median $\Gamma \sim 1.57$ is recovered, both much flatter than the inputted Gaussian prior peak of $\Gamma = 1.95$ and closer to the median value found using simple models applied to *NuSTAR* data (e.g. Del Moro et al. 2017) and for type-2 AGN (e.g. Lanzuisi et al. 2013). This result will be discussed further in Sect. 5.6.

5.5.3 eROSITA derived scattering fractions

A key result of the analysis of the hard-only sources is that they have low scattering fractions $< 1\%$. This result is presented in more detail in Chapter 4, where Fig. 4.20 clearly shows a well measured scattering fraction, and Fig. 4.22 demonstrates that even the 1σ upper limits are well constrained and lower than reported in Brightman et al. (e.g. 2014). Examining the posterior distributions and their 1σ errors including the additional sources presented here, it is find that indeed all scattering fractions are lower than 1%, most (12/16) lower than 0.5%, and all are listed in Table 5.5. This is much lower than what found from other samples (e.g. Brightman et al. 2014), who analysed Chandra data using a similar analysis technique, finding median values of $\sim$2.6%. This is sensible, as eROSITA is extremely sensitive to soft X-rays, and therefore any small scattering component should be readily detected, resulting in the source not being included in the sample. This eROSITA selected sample is therefore simply exploring a more extreme parameter space than covered by other samples.

In order to better visualize differences between average obscurer properties of different groups, one can examine the parameter values derived from `UXClumpy` as a function of galaxy class. In order to derive these values, rather than simply averaging the medians, the probability distributions (posteriors) for each source in each class are combined, following the Hierarchical Bayesian Model (HBM) method described in detail in the Appendix of Baronchelli et al. (2018). In this way, the analysis is able to derive the sample properties using the summed (or stacked) posterior, so that the values are reflective of the tight parameter constraints obtained on some parameters for some sources. Once the stacked posterior is obtained, a Gaussian model is applied; it is the peak and associated error on this Gaussian that is taken as the value and error for each sub-sample.

The resulting values are shown in Fig. 5.9. For all classes except the face-on galaxies, modelling shows scattering fractions of $<0.5\%$, suggesting no detectable emission in soft X-rays beyond emission lines in the torus. Interestingly, variations in the median scattering fractions measured with eROSITA can be seen, where the edge-on spiral galaxies tend

Tabelle 5.5: Spectral fit parameters obtained from `UXClumpy` modelling. Column (1) gives the source name. Column (2) gives the scattering fraction derived from eROSITA spectral modelling. Columns (3), (4), (5), (6), and (7) give the best fit photon index Γ, the obscuring column density $\log(N_H)$, the height of the obscurer TORsigma, the Compton thick covering fraction of the obscuring region CTKcover, and the obscurer inclination Theta, respectively, and associated errors; all are measured with `UXClumpy` using *NuSTAR* data. (*) this parameter shows an extreme peak in the posterior at a obscurer height of 0 degrees; see text and Fig. 5.10.

(1) Name	(2) f_{scat} (%)	(3) PhoIndex Γ	(4) $\log(N_H)$ (cm^{-2})	(5) TORsigma [°]	(6) CTKcover	(7) θ_{inc} [°]
Golden sample						
ESO 119-8	$0.71^{+0.35}_{-0.42}$	$1.92^{+0.08}_{-0.09}$	$23.21^{+0.06}_{-0.07}$	39^{+30}_{-30}	$0.25^{+0.23}_{-0.16}$	47^{+23}_{-35}
ESO 121-G6	$0.04^{+0.03}_{-0.04}$	1.96±0.06	$23.19^{+0.04}_{-0.05}$	21^{+31}_{-15}	$0.32^{+0.17}_{-0.20}$	70^{+13}_{-29}
Fairall 272	$0.04^{+0.11}_{-0.04}$	$1.60^{+0.05}_{-0.03}$	$23.21^{+0.02}_{-0.03}$	37^{+25}_{-22}	0.24±0.16	56^{+25}_{-31}
IC 2461	$0.28^{+0.12}_{-0.17}$	$1.80^{+0.05}_{-0.04}$	$22.74^{+0.07}_{-0.06}$	52^{+21}_{-24}	$0.17^{+0.17}_{-0.13}$	36^{+29}_{-24}
IRAS 00521-7054	$0.022^{+0.023}_{-0.021}$	$2.08^{+0.07}_{-0.06}$	23.18±0.05	68^{+12}_{-22}	$0.37^{+0.17}_{-0.18}$	22^{+28}_{-18}
LEDA 549777	$0.03^{+0.02}_{-0.03}$	$1.83^{+0.07}_{-0.08}$	23.13±0.06	43^{+28}_{-27}	$0.22^{+0.21}_{-0.15}$	47±30
NGC 6300	$0.15^{+0.06}_{-0.14}$	$1.76^{+0.02}_{-0.03}$	$23.15^{+0.03}_{-0.02}$	63^{+15}_{-33}	$0.49^{+0.09}_{-0.17}$	17^{+26}_{-12}
Supplementary sample						
J06411806	0.89±0.21	$1.79^{+0.07}_{-0.06}$	23.07±0.04	24^{+34}_{-19}	$0.25^{+0.25}_{-0.20}$	48^{+31}_{-37}
J14104482	$0.19^{+0.07}_{-0.11}$	$1.75^{+0.07}_{-0.06}$	$22.74^{+0.10}_{-0.13}$	48^{+24}_{-27}	$0.31^{+0.20}_{-0.19}$	44^{+32}_{-28}
CGCG 041-020	$0.29^{+0.09}_{-0.08}$	1.96±0.05	$22.96^{+0.04}_{-0.06}$	50^{+24}_{-25}	$0.20^{+0.22}_{-0.15}$	46^{+29}_{-32}
ESO 103-35	0.11±0.03	$1.79^{+0.02}_{-0.01}$	$23.21^{+0.01}_{-0.02}$	79^{+4}_{-8}	$0.32^{+0.11}_{-0.05}$	24^{+30}_{-20}
ESO 197-27	$0.008^{+0.006}_{-0.005}$	$1.98^{+0.06}_{-0.07}$	$22.87^{+0.08}_{-0.09}$	0*	0.34±0.17	35^{+29}_{-24}
IGR J09523-6231	$0.55^{+0.11}_{-0.13}$	$1.99^{+0.06}_{-0.04}$	$23.07^{+0.05}_{-0.07}$	38^{+30}_{-24}	$0.23^{+0.20}_{-0.17}$	46^{+29}_{-31}
LEDA 93974	0.85±0.20	$1.94^{+0.04}_{-0.06}$	22.80±0.07	59^{+18}_{-25}	$0.29^{+0.20}_{-0.18}$	35^{+27}_{-25}
MCG-01-24-12	$0.43^{+0.09}_{-0.10}$	$1.94^{+0.03}_{-0.04}$	22.81±0.04	41^{+28}_{-24}	$0.15^{+0.19}_{-0.11}$	47^{+30}_{-27}
NVSS J122810-092702	$0.18^{+0.07}_{-0.10}$	$1.91^{+0.05}_{-0.07}$	23.18±0.06	10^{+20}_{-5}	$0.31^{+0.13}_{-0.24}$	68^{+17}_{-55}

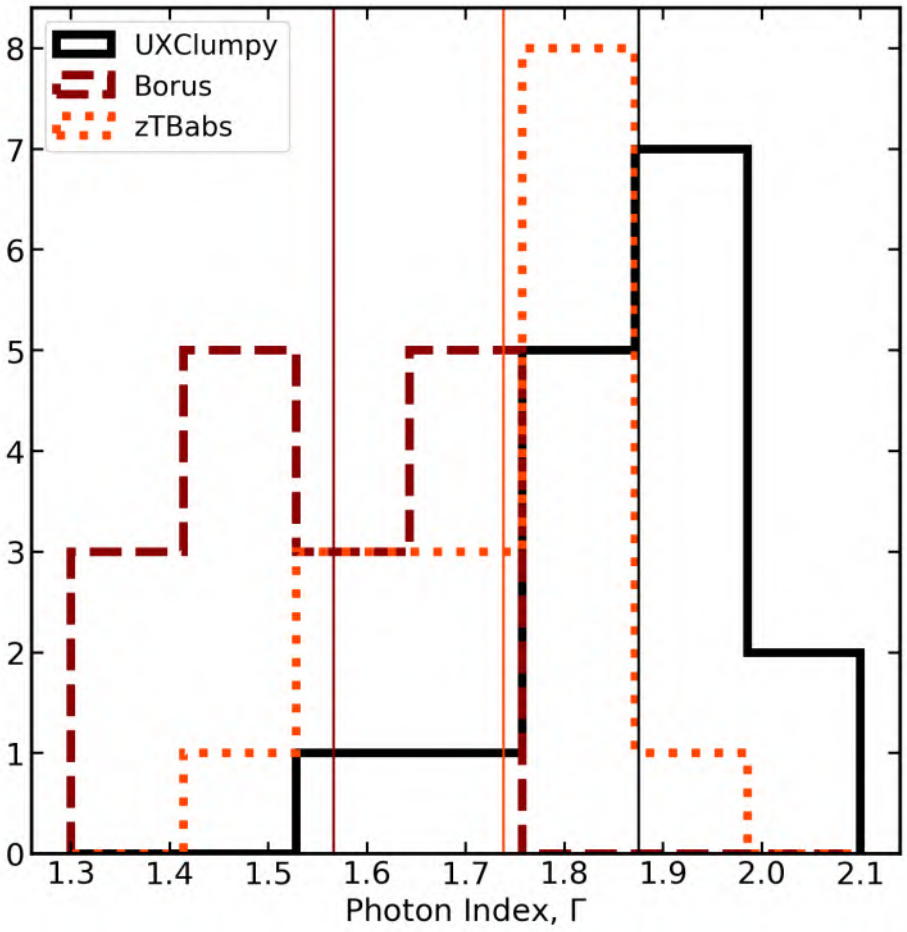

Abbildung 5.8: Best fit photon indices for each class, for each of the spectral models applied to the *NuSTAR* data. Photon indices measured from `UXClumpy`, `borus` and `zTBabs` are shown in black, red and orange, respectively. Mean values each spectral fit are shown with vertical lines in corresponding colours.

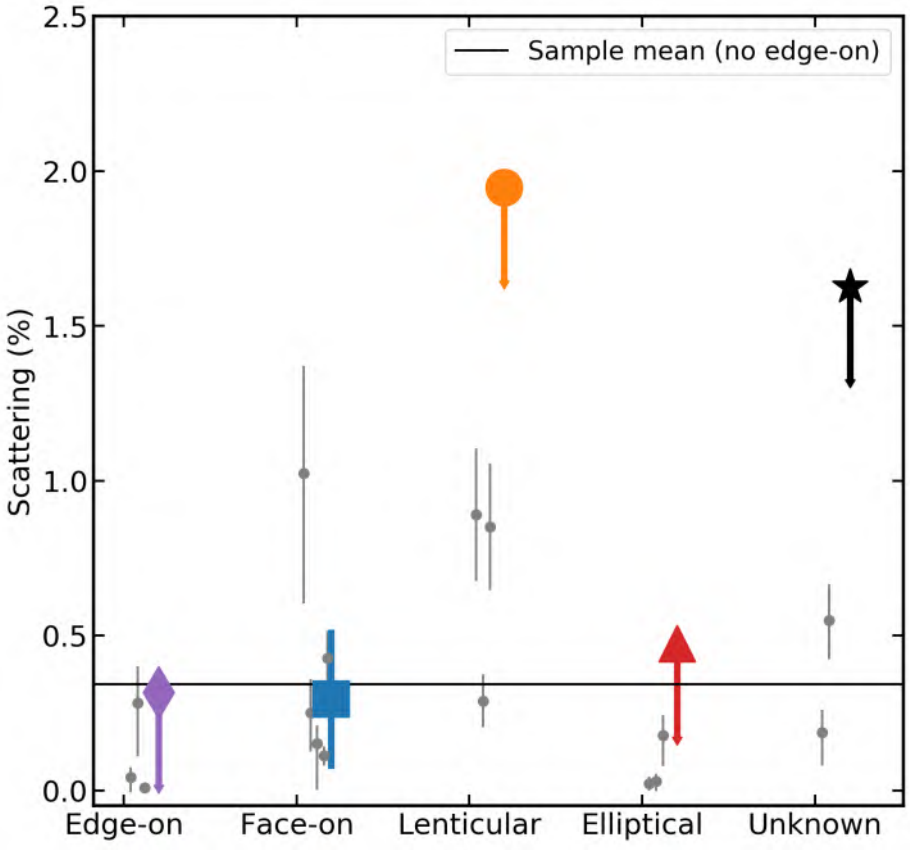

Abbildung 5.9: Distribution of scattering fraction (%) measured from the eROSITA fits. Edge-on spiral galaxies are plotted with purple diamonds, face-on spiral galaxies are plotted with blue squares, lenticular galaxies are shown with orange circles, elliptical galaxies are shown as red triangles, and unknown class galaxies are shown as black stars. Sample values and errors are derived by fitting a Gaussian to stacked posteriors. We stress that all sample values are upper limits to the derived scattering fractions, and only the face-on galaxies have non-zero scattering. Measurements and associated errors for individual sources are shown with small grey circles.

to have lower scattering fractions, as do the elliptical galaxies, whereas these values are non-zero for the face-on spiral galaxies (see Fig. 5.9), and are higher for lenticular and unclassified sources. This will be discussed further in Sect. 5.6.

5.5.4 Edge-on spiral galaxies

For all three edge-on galaxies; ESO 121-G6, IC 2461 and ESO 197-27, reasonable constraints on fit parameters from `UXClumpy` are obtained. However, as discussed in Section 5.1, ESO 121-G6 is actually better fit with `ztbabs` and IC 2461 by `borus`. For ESO 121-G6, the obscurer angle is small, implying a wedge shaped obscurer viewed at a large inclination angle (e.g. edge-on), and supported by the fact that the `ztbabs` model is preferred, it appears that the source is obscured at least in part by the ISM in the disc of the host galaxy. For IC 2461, the parameters derived from borus suggest both an intermediate viewing angle and inclination, but these change dramatically based on the assumptions made; when the normalisations of the reflected and transmitted components are not linked, for example, a thin obscurer is preferred. For ESO 197-27 Fig. 5.10 (top) shows the plot of the spectrum and residuals (top). The spectrum is well fit by the model. Examining the peak of the posterior, there is a large peak at a minimum obscurer height of $\sim 0-10$ and an inclination of $\sim 35^{o}$, consistent with a partially edge-on galaxy and a wedge-shaped obscurer.

Fig. 5.10 (bottom) shows the corner plot, showing the parameter posteriors derived from modelling ESO 197-27 with `UXClumpy`. Most parameters (PhoIndex (Γ), log(N_H), TORsigma, CTKcover, θ_{inc}) are described and presented in Table 5.5. There are also a number of additional parameter; row (1) shows log(norm), the logarithm of the normalisation parameter indicating the flux, row (7) is the factor which indicates the relative normalisation between the FPMA and FPMB instruments (which should be $0.9-1.1$), and rows (8) and (9) show the relative normalisation of the background for FPMA and FPMB respectively, where a value of ~ 1 indicates that the background is well fit.

Examining the key results in Table 5.5, it appears that the edge-on galaxies have lower values of torus heights, Fig. 5.11 shows the median distributions of each sample of the torus height, using the same posterior stacking approach as described in Section 5.3. As expected, the median TORsigma values are found to be lowest for edge-on spiral galaxies, suggesting that the absorber in these sources is thin, whereas all other source classes have larger TORsigma values, suggesting a more spherical obscurer shape (see Fig. 5.11). An exercise is also performed where all posteriors from the face-on spiral, lenticular, elliptical and unknown classes (all but the edge-on spirals) are summed, and a median TORsigma $= 60 \pm 13$ is derived. Values for different sub-classes do not differ within the derived 1σ errors, although this is not surprising given the small number of sources in the edge-on class, and given that 2/3 are not best-fit using `UXClumpy`.

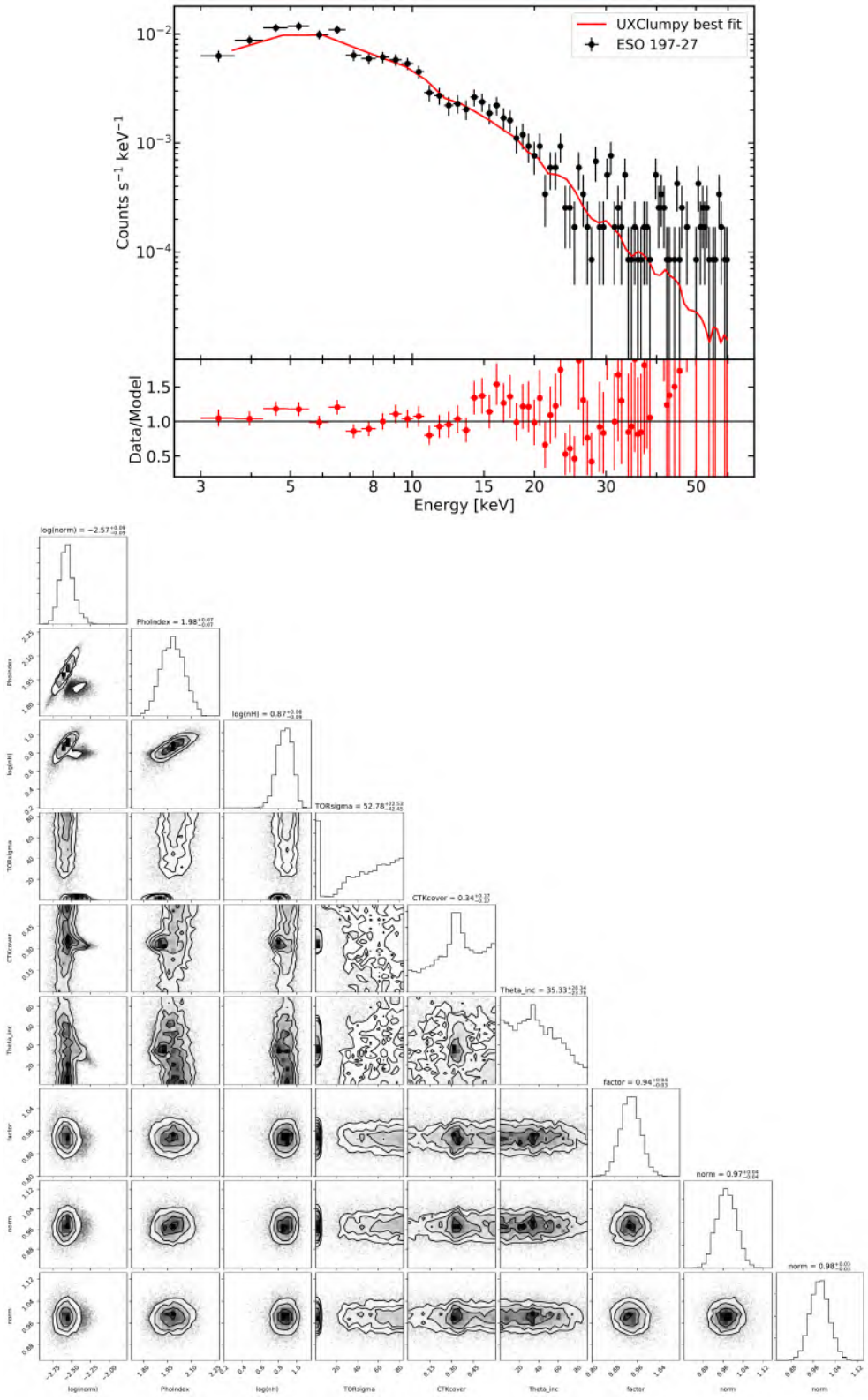

Abbildung 5.10: Top: Spectrum (black; shown only for FPMA), best-fit model (solid red line) and residuals (data/model; red points) best-fit `UXClumpy` model applied to edge-on spiral galaxy ESO 197-27, from the secondary sample. The data and model have been re-binned for display purposes, but not for spectral fitting. Bottom: corner plot including posteriors for best-fit `UXClumpy` model, with parameter names and meanings matching the column headers in Table 5.5 and in the text.

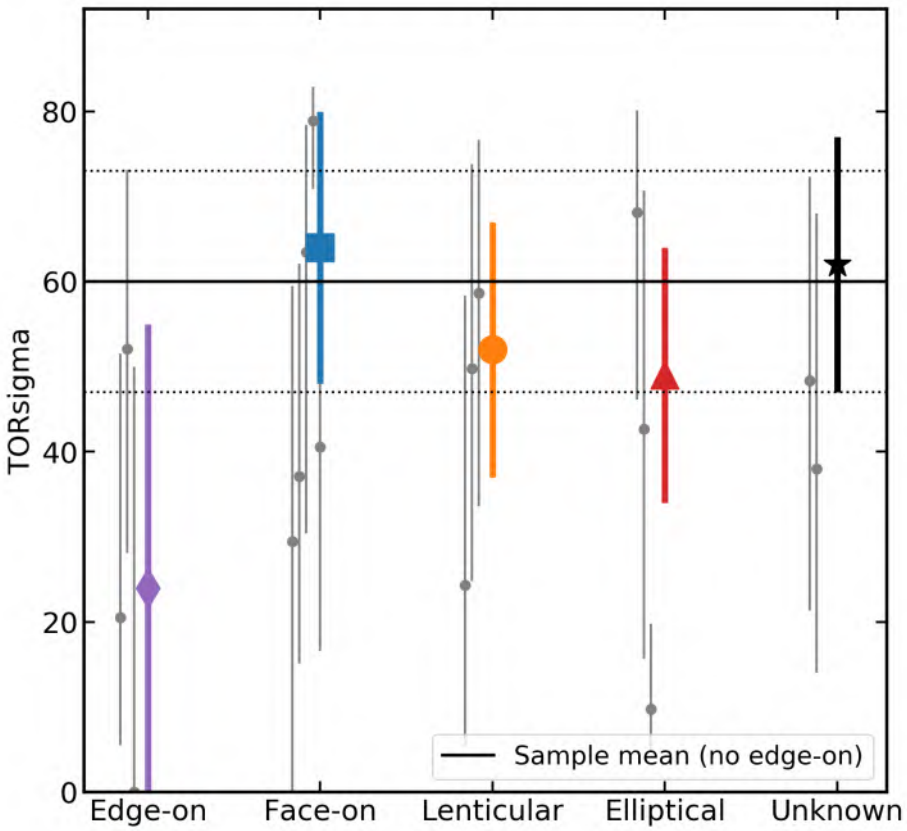

Abbildung 5.11: Distribution of torus obscurer heights, TORsigma, measured with `UXClumpy`, as a function of host-galaxy properties. Edge-on spiral galaxies are plotted with purple diamonds, face-on spiral galaxies are plotted with blue squares, lenticular galaxies are shown with orange circles, elliptical galaxies are shown as red triangles, and unknown class galaxies are shown as black stars. Measurements for individual sources are shown with small grey circles. Mean sample values derived from stacking the posterior of all sources excluding the edge-on galaxies are indicated with a horizontal black line, and the associated error is shown with grey dotted lines.

5.5.5 Face-on spirals and other galaxy classes

Of the five face-on galaxies, all of them are best fit with `UXClumpy`, suggesting that all sources are absorbed by a clumpy obscurer. Of the three sources in the golden sample, ESO 119-8 does not have sufficient data quality to constrain parameters, while Fairall 272 and NGC 6300 do have sufficient data. As an example, Fig. 5.12 shows the spectral plot (top) and corner plot (bottom) for NGC 6300. The spectrum is well-fit with `UXClumpy`, as can be seen from the residuals (data/model) in the bottom panel of Fig. 5.12 (top); only data from FPMA are shown for clarity and simplicity, but it is verified that data from FPMB are also well-modelled using `UXClumpy`.

Posteriors are shown Fig. 5.12 (bottom), with parameter names matching the column headers in Table 5.5 (see also Fig. 5.10 and Section 5.4). Examining these, the data strongly reject the lower values of the obscurer opening angle, suggesting that it is not a narrow wedge shape. For Fairall 272, a more intermediate angle of 37.1 is recovered, whereas for NGC 6300, it is near the maximum at 63.4. The viewing angle is also slightly higher for Fairall 272 at around 56 degrees, whereas NGC 6300 is viewed nearly face on, with an inclination of only 17 degrees, and more edge-on inclination angles are rejected. Since NGC 6300 was observed a total of three times, we also fit the other two data sets and can confirm that the inferred obscurer geometry is consistent across all epochs; we then jointly merge spectra across all epochs and re-fit the data, again yielding the same result. From the secondary sample, good parameter constraints for MCG-01-24-12 were not obtained, but parameters are well constrained for ESO 103-35. The parameters are similar to NGC 6300, with a large obscurer angle of 79 degrees and a low inclination of 24 degrees, suggesting a nearly spherical obscurer seen face on.

Moving on to look at the lenticular and elliptical systems, we again check the images and do not see obvious visual evidence for dust streams or discs in any sources. Of these, good parameter constraints are obtained on IRAS 00521-7054 from the golden sample, and LEDA93974 and NVSS J122810-092702 from the secondary sample, and some constraints can be placed on 2MASX J06411806+3249313 in the secondary sample. Both IRAS 00521-7054 and LEDA93974 show evidence for a spherical obscurer viewed face-on, while NVSS J122810-092702 prefers a wedge-shaped obscurer viewed more edge-on, but in fact the Bayesian evidence is lower for `ztbabs` than `UXClumpy`. Finally, there are two sources which remain entirely unclassified, and neither have particularly well constrained parameters, so it is difficult to make any further statements on these sources.

Examining the column densities shown in Fig. 5.13 (left), values around $1.1\times10^{23}\,\mathrm{cm}^{-2}$ are recovered for all classes, with relatively little variation between samples compared to the large scatter found within each sample. This is likely to be explained by selection effects stemming from the eROSITA selection criteria, where sources with significantly higher column densities but low scattering fractions should be obscured in the $2.3-5\,\mathrm{keV}$ band and therefore not appear in the eRASS1 hard sample, while sources with lower column densities would likely be brighter in the soft X-rays and therefore have $HR > 0$ and thus being excluded from this analysis. Little difference can be found between the median column densities for each sub-class. The same can be said for the Compton thick covering

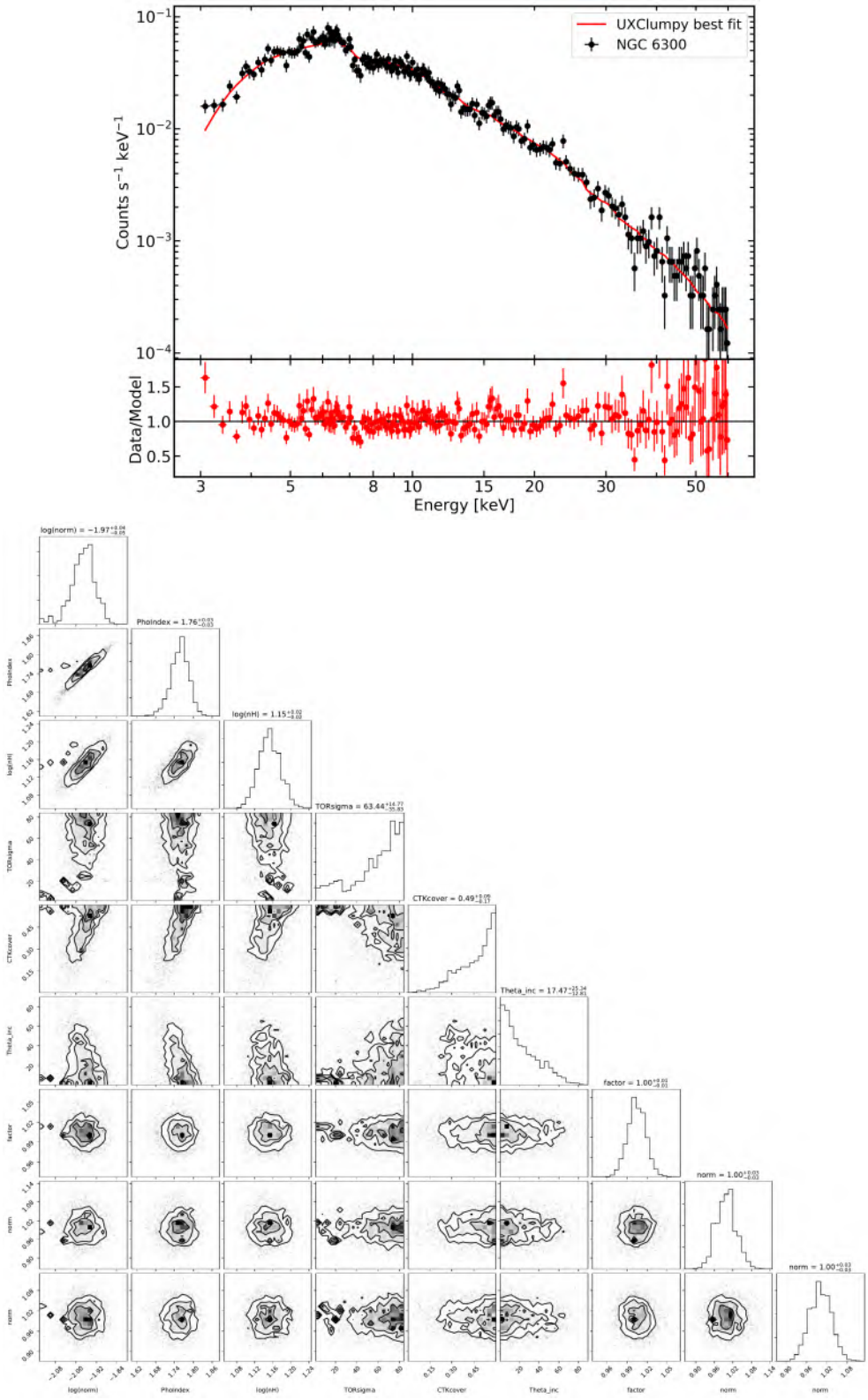

Abbildung 5.12: Top: Spectrum (black; shown only for FPMA), best-fit model (solid red line) and residuals (data/model; red points) best-fit `UXClumpy` model applied to face-on spiral galaxy NGC 6300, from our golden sample. The data and model have been re-binned for display purposes, but not for spectral fitting. Bottom: corner plot including posteriors for best-fit `UXClumpy` model, with parameter names matching the column headers in Table 5.5.

factor (Fig. 5.13, centre), where most sources recover a best-fit value near to the median of the prior at ~ 0.3, suggesting that this parameter is not well constrained for most sources; aside from NGC 6300, which shows strong evidence for a large -population of Compton thick clouds located outside the line of sight (e.g. has large CTKcover). Finally, Fig. 5.13 (right) shows the distributions of inclinations for each sub class. Differences here are less apparent, although there is some evidence that the inclinations are slightly lower for the face-on galaxies and larger for the edge-on galaxies. The viewing angle is degenerate with the Compton thick covering factor CTKcover, and that small changes in inclination as long as $\theta_{\rm inc} >$ TORsigma (for edge-on galaxies) and $\theta_{\rm inc} <$ TORsigma (face-on galaxies) may not be readily detectable in the spectra.

5.6 Comments on individual sources

5.6.1 Golden sample

ESO 119-8

ESO 119-8 (Fig. 5.14) is a face-on ($b/a = 0.98$), barred spiral (SB(r)a) galaxy at $z = 0.023260$. The source was observed once for about 26 ks with *NuSTAR* and has about 1800 counts (FPMA+FPMB) in the $3 - 60$ keV range, which is not sufficient to constrain the best-fit torus parameters. The source also has a relatively higher scattering fraction of 0.72%. The source was studied by Zhao et al. (2021) using *Chandra* and *NuSTAR* data and applying the model `borus`, and the geometry of the obscurer could not be constrained. Barua et al. (2021) apply a relativistic blurred reflection model to the *NuSTAR* data and find evidence for coronal temperature variations, suggesting that the spectrum may be complex and not well modeled using a torus model alone.

ESO 121-G6

ESO 121-G6 (Fig. 5.15) is an edge-on ($b/a = 0.2$) Sc spiral galaxy at a low redshift $z = 0.00403$. Interestingly, this source is actually best described using `ztbabs`, which has lower Bayesian evidence than `borus` or `UXClumpy`. With `UXClumpy`, the spectrum requires a narrow, wedge-shaped torus viewed edge-on. It is therefore highly plausible for this source that the emission originates in the larger scale galactic disc, and may not be nuclear. The best fit column density from `ztbabs` is $2.5 \times 10^{23}\,{\rm cm}^{-2}$, which is a rare but not impossible column density for host-galaxy scales; likely $\lesssim 15\%$ of sources would lie at this column density range (Buchner & Bauer 2017). The scattering fraction for this source is extremely low, at 0.04%, and is in agreement with 0%, which could also point to multiple absorption zones at different scales.

This source has been previously studied in X-rays, as well as at other wavelengths. These data have been studied by Annuar et al. (2020), who studied a sample of AGN to assess whether they are low-luminosity AGN or heavily obscured sources, and by applying physically motivated torus models, agree that it is a heavily obscured and possibly

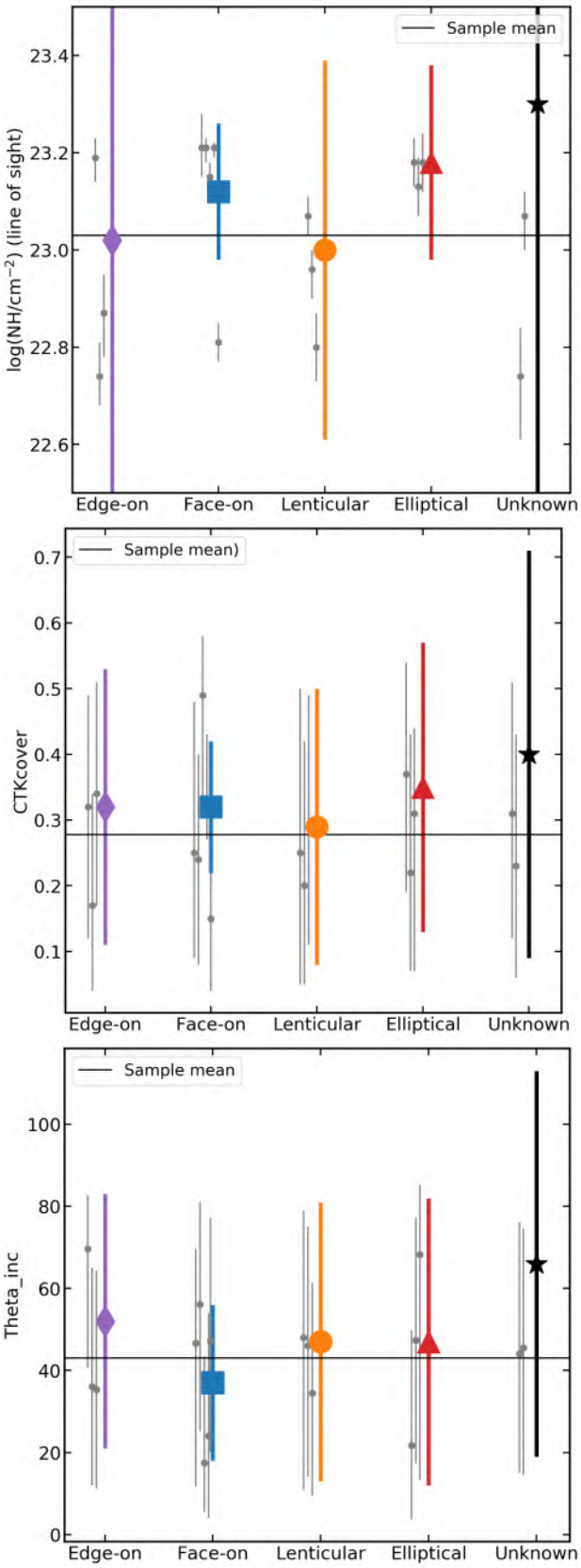

Abbildung 5.13: Top: Line of sight column density, $\log(N_H)$, measured with `UXClumpy`. Centre: obscurer heights, CTKcover, measured with `UXClumpy`. Bottom: obscurer inclination angles, θ_{inc}, measured with `UXClumpy`.

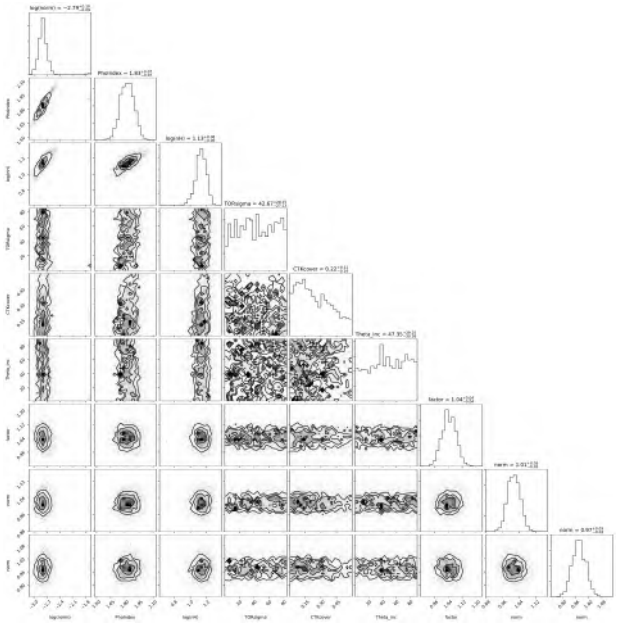

Abbildung 5.14: Corner plot including posteriors for best-fit `UXClumpy` model applied to ESO 119-8, with parameter names matching the column headers in Table 5.3. Additional parameters are also described in Fig. 5.10 and Section 5.4. The torus geometry parameters cannot be well constrained.

Compton-thick AGN, although with a best-fit column density of $2 \times 10^{23}\,\mathrm{cm}^{-2}$, in agreement with our measurement. There is also some evidence for a Compton thick ring with a covering factor of ~ 0.45, however there is not sufficient data to confirm this.

Fairall 272

We find Fairall 272 (Fig. 5.16) to be a borderline source, which for the purpose of this analysis, is classified as a face-on ($b/a = 0.42$; LS10 measured from available LS9/LS10 data) SA spiral galaxy at $z = 0.022$. Although the source is best fit with `UXClumpy`, the photon index is flatter than for other sources, with $\Gamma = 1.6$. Although the median of the TORsigma torus height is relatively low, the possibility of a thin wedge torus is strongly rejected, and the inclination is not well constrained with the available data. The scattering fraction is low and is in agreement with 0, suggesting a large covering fraction for the torus. The source was also studied using *NuSTAR* and *XMM-Newton* by Zhao et al. (2021), where modelling using `borus` suggested a high torus column density and moderate torus height and viewing angle.

IC 2461

IC 2461 (Fig. 5.17) is the most edge-on ($b/a = 0.18$) source in our sample, and is classified as a Sb/Sbc spiral galaxy at a low redshift of $z = 0.0075$. The source has one of the

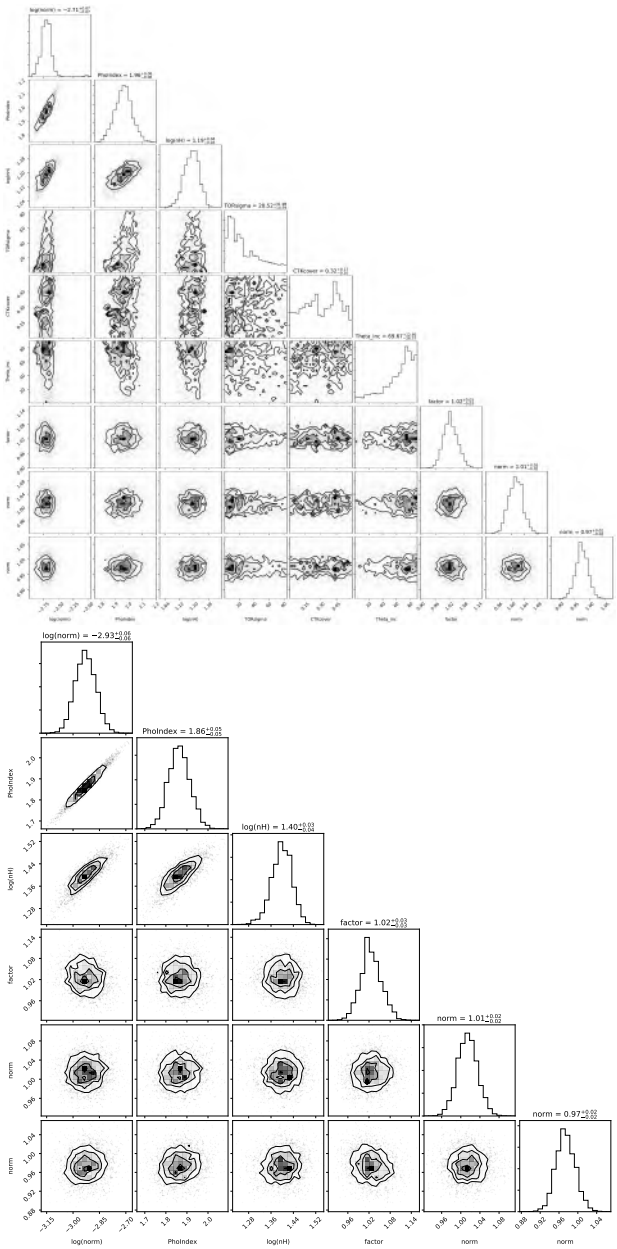

Abbildung 5.15: Top: As Fig. 5.14, but for ESO 121-G6. The torus height is near the minimum, suggesting a thin wedge torus, while the inclination is near the maximum, suggesting it is viewed edge-on. Bottom: corner plot for the `ztbabs` model applied to ESO 121-G6, where the column names have the same meanings. This model has the lowest Bayesian evidence and is therefore the preferred model, and provides good constraints on the column density and photon index.

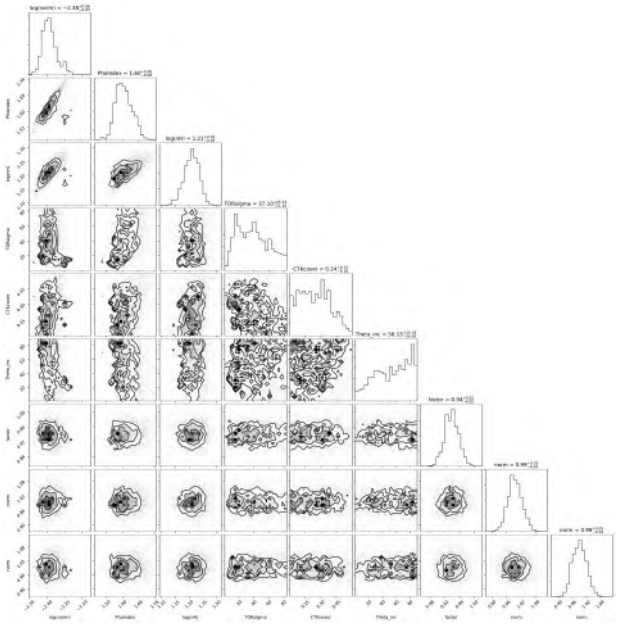

Abbildung 5.16: As Fig. 5.14, but for Fairall 272. Paramaters are not well constrained, but one can reject a very wedge-shaped torus with TORsigma$\sim 10^{o}$.

lowest column densities, with $\log(N_H) \sim 22.74$, and shows some evidence for a lack of any Compton thick ring component, unlike in some other sources; however, the source is actually best fit with `borus`. The best fit parameters for the torus geometry and viewing angle differ dramatically based on the interpretations made; when the column densities and normalisations of the transmitted and reflected components are linked, the torus appears wedge-shaped and is viewed edge-on; when these are left free to vary, the torus parameters are no longer well constrained.

IRAS 00521-7054

The source IRAS 00521-7054 (Fig. 5.18) has been classified as either elliptical or lenticular Sy2 galaxy, but that appears moderately circular on the sky with $b/a = 0.8$. It lies in behind the Small Magellanic Cloud, so there are many stars in the nearby area. It is the highest luminosity and highest redshift source in the golden sample, but with three separate exposures, there is significant data available for modelling. The parameters are well constrained and strongly suggest a spherical torus viewed nearly edge-on. There is also strong evidence for a Compton thick ring with a high CTKcover.

Several previous studies on this source have also revealed interesting and distinct properties; Frogel & Elias (1987) study the optical and infrared properties of the source, showing it to be particularly red/hot, with narrow optical Hα and some evidence for a broad underlying component, and strong forbidden lines. The authors propose that the source is in transition from a starburst galaxy to an active galaxy. The X-ray properties of this source are also distinct; the source shows evidence for a broad Fe Kα line, suggesting a

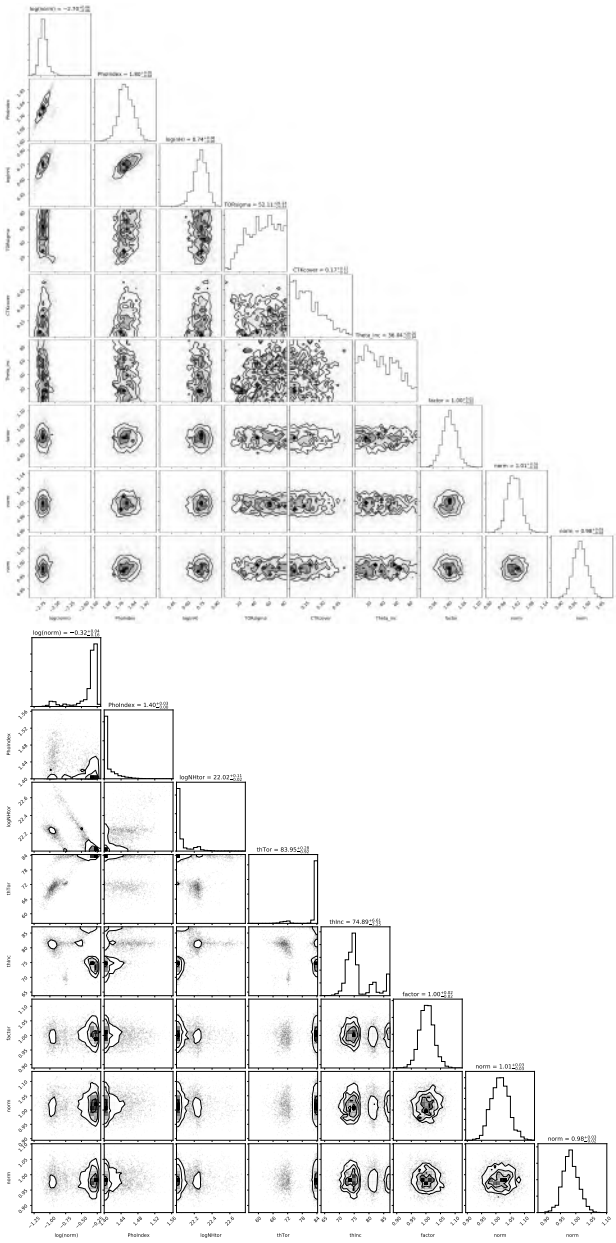

Abbildung 5.17: Top: As Fig. 5.14, but for IC 2461. One can only place weak constraints on the torus geometry, although this is not the preferred model. Bottom: corner plot for `borus`, the preferred model for IC 2461. Here, parameter values with the same names have the same meaning as above. Row (3) shows logNHtor, the logarithm of the torus column density, row (4) shows thTor, the covering angle of the torus (see Section 2), and row (5) shows thInc, the viewing angle of the torus (see Section 2). There is evidence for a wedge shaped torus viewed edge-on, cautioning that here thTor refers to the opening angle, not the height, such that large thTor is a very thin torus.

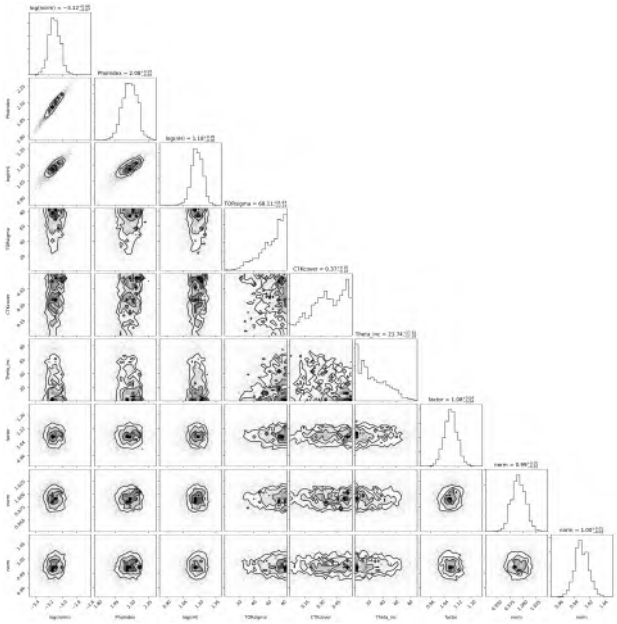

Abbildung 5.18: As Fig. 5.14, but for IRAS 00521-7054. The corner plot is similar to that of NGC 6300 (Fig. 5.12), with a maximum (spherical) torus height, a low inclination, and some preference for a large CTK covering factor.

strong blurred relativistic reflection component (Tan et al. 2012; Ricci et al. 2014; Walton et al. 2019), the first reported in a Sy2 galaxy. The source also may show evidence for for a relativistic outflow (Walton et al. 2019), revealed using *XMM-Newton* and *NuSTAR*, revealed to be strong enough to provide feedback on host-galaxy scales. Ricci et al. (2014) find evidence for a steep power law component ($\Gamma \sim 2.2-2.3$) and suggest that this source may be an obscured narrow-line Seyfert 1 galaxy; a slightly flatter slope of $\Gamma \sim 2.08$ is found in this chapter, but in agreement with these measurements.

LEDA 549777

LEDA 549777 (Fig. 5.19 is the only source in the golden sample which does not have an optical class; however, based on Sersic index provided in LS10, the source is probably elliptical/lenticular. For the purpose of this chapter, it is classified as lenticular, although the LS10 image may show some evidence for a disc. The *NuSTAR* data are low quality and are not sufficient to constrain most spectral parameters, which was also concluded by Zhao et al. (2021). X-ray and optical (imaging) follow-up would be of interest to better classify and model this source.

NGC 6300

NGC 6300 is a low redshift barred spiral galaxy ($b/a = 0.66$, SB(rs)b) which is also classified as a Sy2. The corner plot for this source is shown in Fig. 5.12. The source has

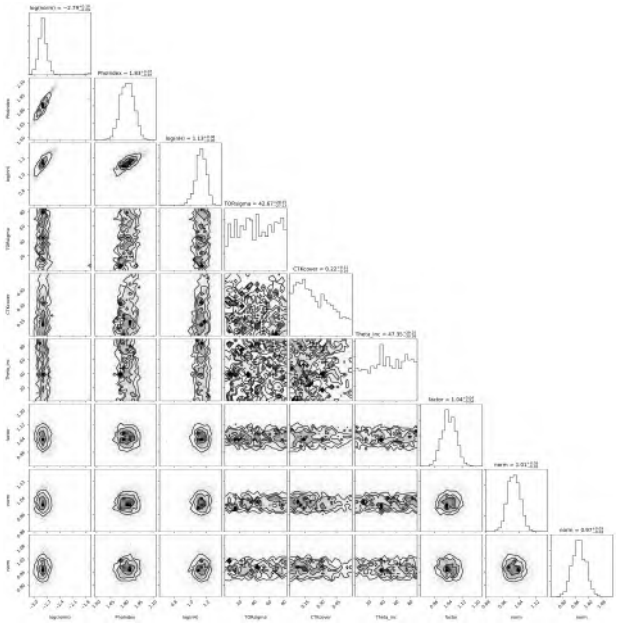

Abbildung 5.19: As Fig. 5.14, but for LEDA 549777. The torus geometry parameters cannot be well constrained.

been well studied in X-rays, e.g. by Jana et al. (2020). Due to its close proximity to the Milky Way, it is possible to separate the nuclear emission from the AGN emission, e.g. with *Chandra*; it is shown that there is a galactic-scale emission component, as well as a heavily obscured AGN, with a column density varying on timescales of days, suggesting a clumpy torus (Jana et al. 2020). The analysis in this chapter agrees with these results, recovering a similar column density of $\sim 1.5 \times 10^{23}\,\mathrm{cm}^{-2}$; however, Jana et al. (2020) find that the torus is viewed nearly edge-on (high inclination), whereas this chapter finds a more face-on view. This may be because of the model choice, as Jana et al. (2020) employ MYTorus (Murphy & Yaqoob 2009), which fixes the torus opening angle to 60°, whereas this chapter recovers an opening angle of $\sim 27°$ (TORsigma = 63°). The source is also studied by Zhao et al. (2021), who also recover a face-on inclination and high torus covering fraction (height).

Another interesting result is that a high CTKcover is recovered, of 0.49. This is the only source where this value is constrained to be > 0.3, which is the median of the prior; this suggests that there is a true Compton thick component in this source, which provides further evidence that the source is obscured by a torus with complex geometry and structure. However, this parameter varies slightly between the three observations of NGC 6300, and that this result is only recovered for 1/3 observations, as well as when jointly modelling all epochs. This should be investigated further in a more detailed individual study.

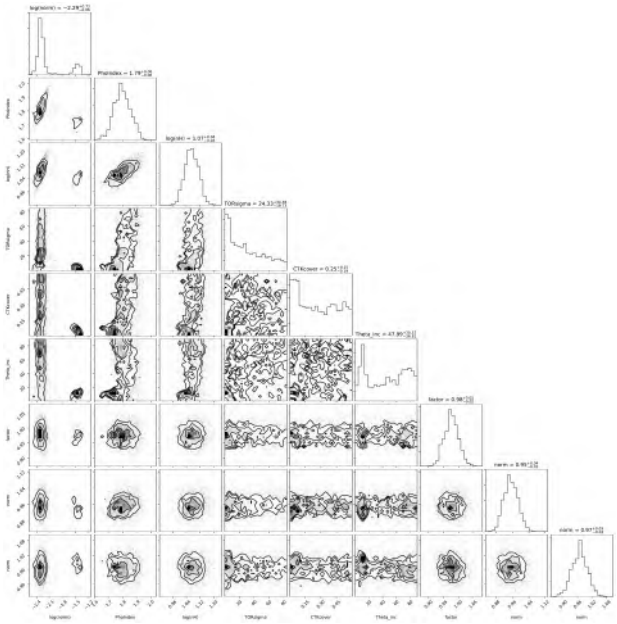

Abbildung 5.20: As Fig. 5.14, but for J06411806. A wedge-shaped torus is found, with a double peaked profile for the viewing angle, but likely the source has a relatively spherical torus viewed edge-on.

5.6.2 Secondary sample

2MASX J06411806+3249313

2MASX J06411806+3249313 (Fig. 5.20; J06411806) is a lenticular (S0) Sy2 galaxy at low redshift of $z = 0.047$. The b/a value is 0.84, however this value is not particularly meaningful as it is unclear if this galaxy has a disc or not. Using our defined fitting parameters recovers an the odd result of a wedge-shaped torus, and a double peaked profile for the viewing angle. However, Zhao et al. (2021) combine *NuSTAR* data with *XMM-Newton* data and find a spherical torus viewed very face-on. This source has the highest individual scattering fractions measured from eROSITA data, with a scattering fraction of $\sim 0.9\%$.

2MASX J14104482-4228325

2MASX J14104482-4228325 (Fig. 5.21) is an apparently face-on/circular source at low redshift ($z = 0.03394$), but the classification is unknown. The source is located outside of the LS10 footprint, but from examining the DSS image it appears more red than other spiral galaxies. The source has a slightly lower than typical column density of $5.5 \times 10^{22}\,\mathrm{cm}^{-2}$, and other obscurer parameters were unconstrained. The photon index is also flatter than the sample median, not as extreme as e.g. Fairall 272. The source appears in the Swift-BAT catalog, but is otherwise not well studied. X-ray and optical follow-up would be of interest to better classify and model this source.

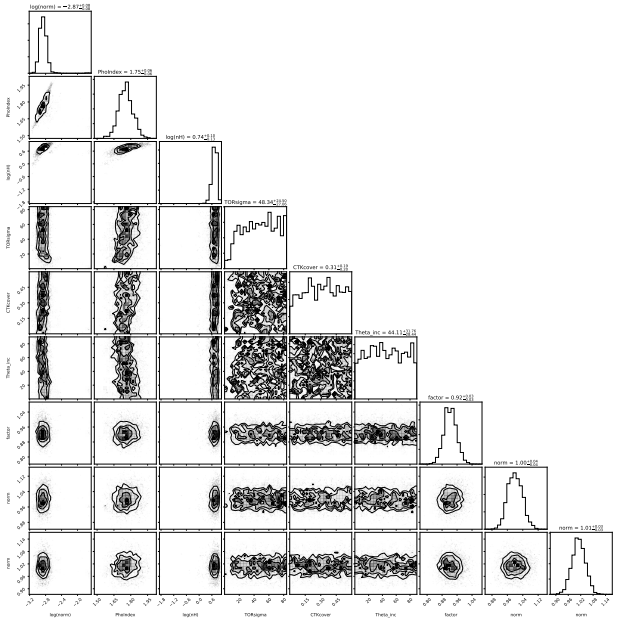

Abbildung 5.21: As Fig. 5.14, but for J14104482. The torus geometry parameters cannot be well constrained.

CGCG 041-020

CGCG 041-020 (Fig. 5.22) is a lenticular Seyfert 2 galaxy at low redshift $z = 0.03599$. The source appears in the INTEGRAL/IBIS 17-yr hard X-ray all-sky survey, suggesting that it has a gamma ray detection, and could potentially be a jetted AGN or blazar. However, a literature search did not identify any papers dedicated to the detailed study of this source. Examining the corner plot reveals that the torus geometry parameters cannot be well constrained, but there is some preference for a more spherical torus.

ESO 197-27

ESO 197-27 is an edge-on ($b/a = 0.38$) Sc spiral galaxy which has also been classified as a Seyfert 2. The corner plot for this source is shown in Fig. 5.10. The source has the lowest scattering fraction of any measured in our sample, at only 0.008%, which is essentially consistent with no scattering. The CTKcover is relatively well constrained around 0.34, suggesting a Compton thick ring covering factor. The posterior distribution of TORsigma parameters is very irregular, with an extremely strong peak at 0, which would imply a highly thin and wedge-shaped torus, but a broad secondary peak at larger values, which would suggest a more face-on torus. An intermediate inclination is also recovered. Further study with additional data would be required to confirm the shape and structure of the torus.

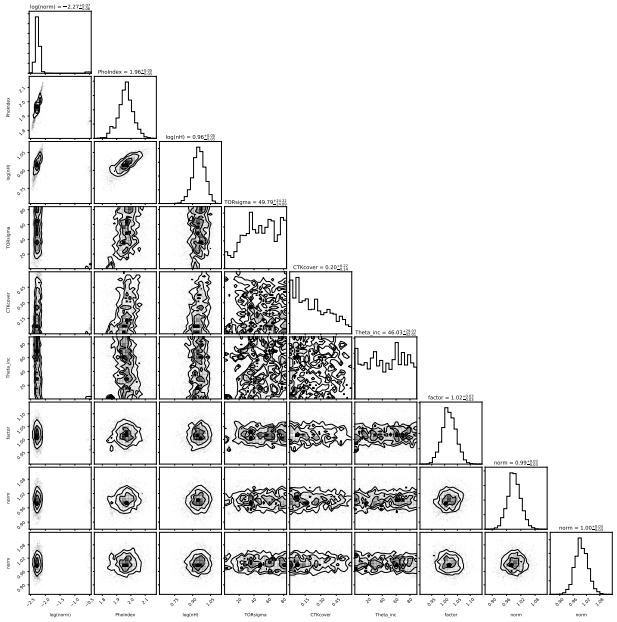

Abbildung 5.22: As Fig. 5.14, but for CGCG 041-020. The torus geometry parameters cannot be well constrained, but there is some preference for a more spherical torus.

ESO 103-35

ESO 103-35 (Fig. 5.23) is a borderline source; the morphological classification is an S0a, between a spiral and elliptical galaxy. Given the b/a value of 0.46, the source is classified the source as a face-on spiral. The source is typically classified as a Seyfert 2, but has also been classified as a Seyfert 1.9. The source was studied in detail with BeppoSAX Wilkes et al. (2001), who recover a column density of $1.8 \times 10^{23}\,\mathrm{cm}^{-2}$, in agreement with out measurement of $1.6 \times 10^{23}\,\mathrm{cm}^{-2}$ within errors. They also report a high energy cut-off of $29 \pm 10\,\mathrm{keV}$, while in this chapter, the value is frozen to 300 keV; this could be explored in more detail in future work. Evidence is found for a highly spherical obscurer, and can strongly reject a thin absorber. Some evidence is also revealed for a Compton thick ring with moderate covering factor CTKcover of 0.32. Our results appear in contrast with Zhao et al. (2021), who report a very Compton thick column density and moderate torus height and inclination, although this is found using `borus`.

IGR J09523-6231

IGR J09523-6231 (Fig. 5.24) is an extremely dusty source with almost no visible optical emission and unknown spiral or elliptical classification, found at a high redshift with respect to this sample, of $z = 0.252$. The source is classified as a Seyfert type 1.9/2, which suggests it has weak or no broad Hα and no other broad optical lines. The source is best fit with `UXClumpy`, but while the torus height and inclination cannot be constrained, there is a peak in the posterior distribution of the CTK covering factor around 0.4, which may give some

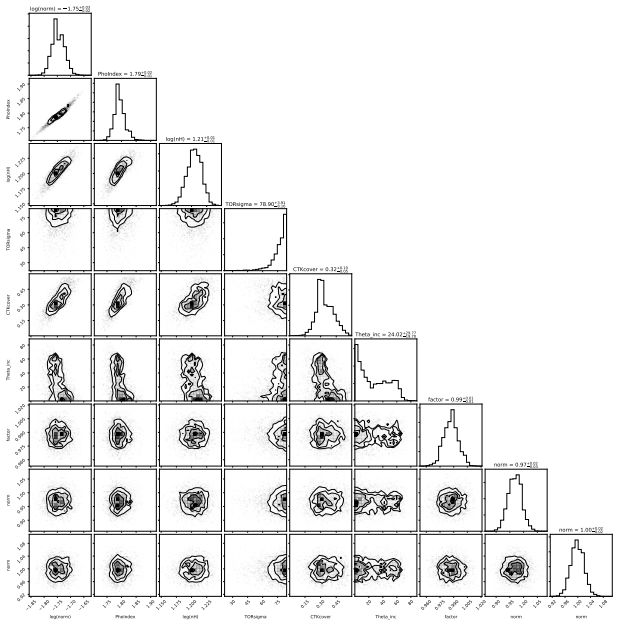

Abbildung 5.23: As Fig. 5.14, but for ESO 103-35. There is evidence for a very large TORsigma (spherical torus) viewed face-on, and evidence for a Compton thick ring with moderate covering fraction.

indication of a Compton thick ring of obscuring material, which is interesting given the high levels of observed dust throughout the galaxy.

LEDA 93974

LEDA 93974 (Fig. 5.25) is an S0 lenticular galaxy which is also classified as a Seyfert 2. The source was studied in some detail by Masetti et al. (2006), who confirm it to be a Compton-thin AGN which is also detected by Integral. This analysis also finds the source to be Compton thin, with a column density of $\sim 6 \times 10^{22}\,\mathrm{cm}^{-2}$, and find a high scattering fraction of 0.85%. Good constraints on the torus parameters cannot be constrained, but a thin torus and an edge-on orientation can both be rejected.

MCG-01-24-12

MCG-01-24-12 (Fig. 5.26) is an extremely face-on ($b/a = 0.96$) spiral (SAB) galaxy at an intermediate redshift of $z = 0.1961$. The source has been observed several times with *NuSTAR*, but for the purpose of the analysis, only the longest observation, ID 60061091010, i considered. The available data have been studied in detail by Middei et al. (2021), who suggest that the source is absorbed by a Compton thin medium, with a Compton thick reflector outside the line of sight. They also find evidence for variability in the Fe Kα line as well as evidence for an absorption feature, which is posited to arise in a variable disc wind (Middei et al. 2021). Unfortunately, we were not able to place good constraints on the

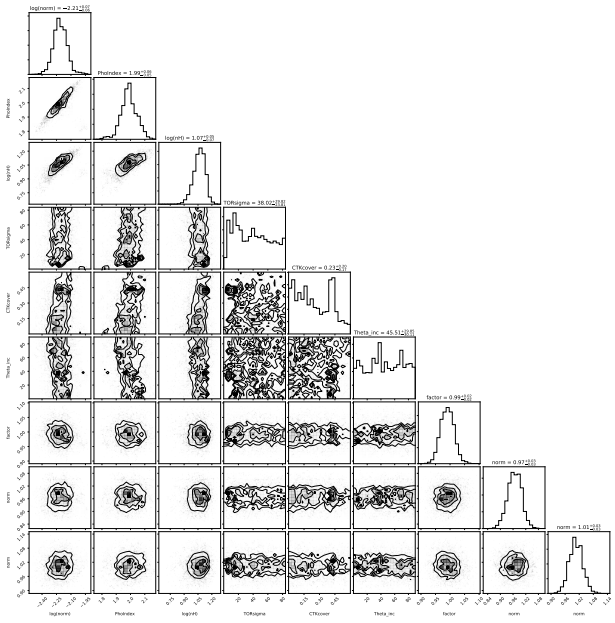

Abbildung 5.24: As Fig. 5.14, but for IGR J09523-6231. The torus geometry parameters cannot be well constrained.

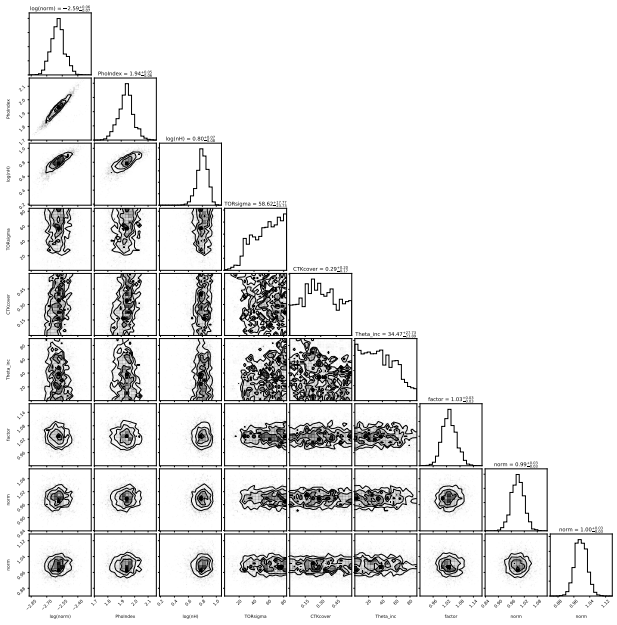

Abbildung 5.25: As Fig. 5.14, but for LEDA 93974. A thin torus and an edge-on orientation are rejected.

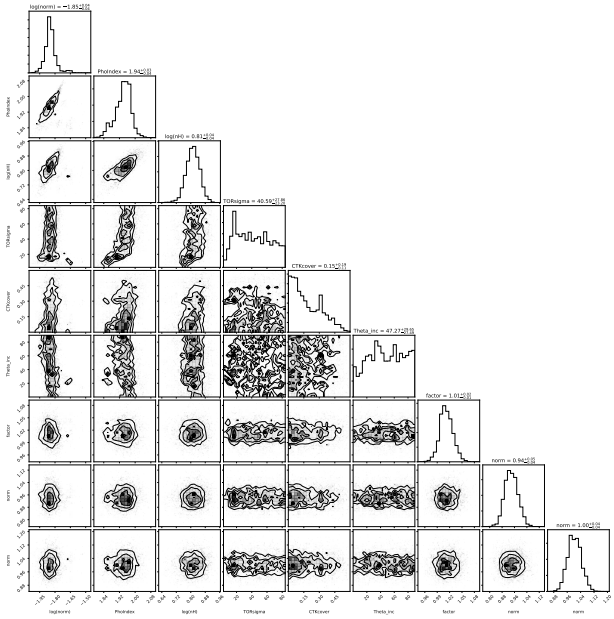

Abbildung 5.26: As Fig. 5.14, but for MCG-01-24-12. A very thin absorber is rejected and we find evidence for a low or absent Compton thick ring.

torus parameters by fitting individual epochs; aside from the Compton thick ring covering factor, which is likely to be low or 0. A highly wedge-shaped torus is strongly rejected.

NVSS J122810-092702

NVSS J122810-092702 (Fig. 5.27) is apparently a radio source with little known about the classification. At a redshift of $z = 0.2624$, it is the most distant source in our sample. Interestingly, it is the only source not immediately classified as edge-on which shows strong evidence for a wedge-shaped obscurer viewed edge-on, with very good constraints placed on both parameters. It is unclear on if the apparent presence of the jet could affect the measurement of these parameters, or if the source is indeed found in an edge-on spiral galaxy, as the source class is not apparent from the image.

5.7 Discussion

A sample of hard or hard-only, low scattering fraction AGN is identified with eROSITA, having $f_{\rm scat} < 1\%$, lower than median values from other samples. These low scattering AGN may help our understanding the most extreme realisations of obscurer geometry and structure, as well as potentially for understanding co-evolution of the central engine and large-scale galactic structure. This chapter sheds light on the origin of the obscuration in these low-scattering sources with *NuSTAR* reflection spectroscopy. The key findings are:

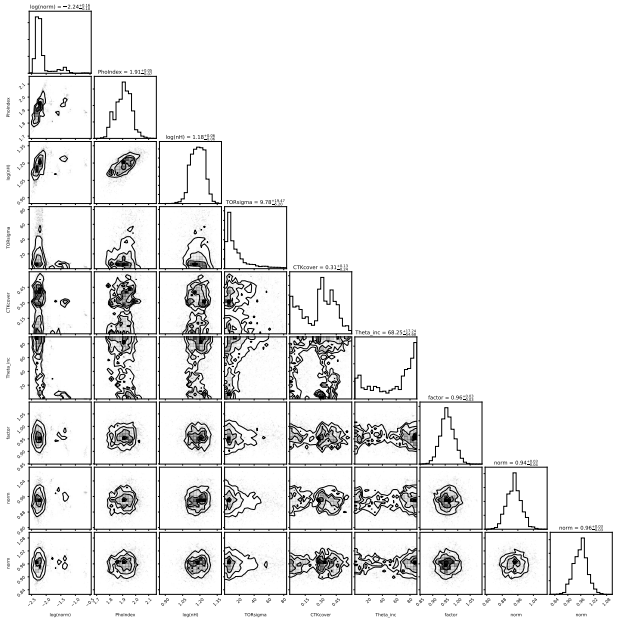

Abbildung 5.27: As Fig. 5.14, but for NVSS J122810-092702. The parameters are similar to those found in edge-on spiral galaxies, with a very thin torus viewed edge-on.

- The analysis reveals a statistical preference for `UXClumpy` for most sources, suggesting that the obscurer is predominantly clumpy/structurally complex (see Table 5.4)
- The photon indices obtained when modelling our sample of Sy2 AGN with `UXClumpy` are steeper than those obtained from `zTBabs` and `borus`, and are not in tension with values measured in Sy1 AGN (see Fig. 5.8)
- Edge-on galaxies show evidence for lower scattering fractions and TORsigma values (see Fig. 5.11 and Fig. 5.9), suggesting that the absorber is consistent with the galactic disc, whereas other classes (in particular face-on spiral galaxies) have higher TORsigma and scattering fractions, suggesting they are obscured by a spherical torus

The analysis reveals that almost all sources prefer `UXClumpy` over `borus`, showing strong preference for a clumpy obscurer. Furthermore, when `borus` is applied assuming a truly uniform obscurer where the column densities and normalisations between the reflected and transmitted components are linked, the Bayesian evidence values increase significantly, suggesting a much worse fit than achieved when allowing the line of sight and reflected column densities to vary separately, and with respect to the `UXClumpy` models. Many previous studies have also found evidence for a clumpy obscurer (Nenkova et al. 2008; Markowitz et al. 2014; Netzer 2015; Buchner et al. 2019; Laha et al. 2020), and this analysis adds further evidence to these results. The edge-on spiral galaxy class has only one source which prefers `UXClumpy`, suggesting that these sources may either not require a clumpy torus, or more likely, do have clumpy tori, but have additional absorber components which affect the spectral modelling.

Examining the photon indices obtained from spectral fitting, there are no major differences between different classes, however there exist significant differences between models, where `borus` and `ztbabs` recover much lower photon indices than `UXClumpy`. This was also investigated in detail by Saha et al. (2022), who find that photon indices measured by different absorption models differ significantly for simulated spectra with known $\Gamma = 1.9$, but each model is able to independently recover the correct Γ. In our analysis, `UXClumpy` best fits the spectral shape, suggesting that the low photon indices recovered by the other models may be a result of poorly fitting the clumpy structure of the torus. The recovered median of $\Gamma \sim 1.9$ is higher to the median value found using simple models applied to *NuSTAR* data (e.g. Del Moro et al. 2017) and for type-2 AGN (Lanzuisi et al. 2013; Liu et al. 2016), but is consistent with the expected values for type-1 AGN (Ishibashi & Courvoisier 2010; Liu et al. 2016; Waddell & Gallo 2020). This could be evidence that in fact, when careful spectral analysis with the correct model is employed, both Seyfert 1 and 2 AGN have the same intrinsic X-ray spectral shape, with observed differences driven only by differences in the absorption, in agreement with the unification model (Antonucci 1993).

By applying different absorber models to data sets using BXA, this chapter demonstrates that we are able to distinguish absorber geometries and reject certain geometries for individual sources. This is clear when examining the posterior distributions (see Fig. 5.10 and Fig. 5.12), many of which show a clear preference for the extremes of the parameter space. These are not necessarily likely for all sources showing evidence of absorption, but this is reasonable for this extreme sample with hard spectra and low scattering fractions. In particular for face-on spiral galaxies, preference is found for a near-spherical obscurer viewed roughly face-on, suggesting an alignment between the galactic disc and the accretion disc/absorber system, and implying that the low scattering fractions are a result of a fully covering torus (see also Ueda et al. 2007; Yamada et al. 2019). This result is particularly clear for ESO 103-35 and NGC 6300, which show clear preference for maximum obscurer height and minimum inclination angle, and a wedge shaped absorber viewed more edge on is strongly ruled out. This also seems to be the case for some AGN in elliptical or lenticular galaxies, where IRAS 00521-7054 and LEDA 93974 both show evidence for a spherical obscurer. This could suggest that overall, a majority of low scattering fraction sources can be interpreted as buried AGN hidden behind nearly spherical uniform absorbers.

The physical interpretation for edge-on sources, however, seems to differ from this. Out of ESO 121-G6, IC 2461 and ESO 197-27, only ESO 197-27 from the secondary sample prefers `UXClumpy`, but the source shows strong evidence for a wedge-shaped absorber. IC 2461 is slightly better fit with `borus`, which shows a relatively small obscurer opening angle (e.g. near spherical) viewed at an intermediate angle of about 43 degrees, which is similar to the best fit parameters obtained with `borus`. Finally, ESO 121-G6 is best fit with `UXClumpy`, requiring a thin absorber viewed edge-on. Overall, this suggests that a clumpy absorber is not necessarily required in edge-on galaxies, and also that these sources may be more likely to be absorbed by the host galaxy disc, or by a combination of a spherical obscurer and the disc(e.g. see also Hönig et al. 2014). It is likely that the underlying structure of the torus remains clumpy, but that the effects in the spectra are diluted due

to additional absorption in the disc. The analysis presented in this chapter also report much lower scattering fractions in these sources, which may be a result of the much larger amount of material available (the entire disc) to absorb emission from the central region rather than merely the obscurer, as seen in the face-on spiral galaxies.

It must be cautioned that, given the small sample size explored here, differences between different samples are rather qualitative and are typically not statistically significant. Clear results can be derived for some individual sources (e.g. see detailed notes in Section 6), but some sources lack sufficient data quality to place tight constraints on posteriors. This then can affect the stacked posteriors, increasing the size of the error bars. Still, the methodology here can be applied to other current and future data in order to expand this sample and clarify this result.

5.8 Lessons learned on low scattering fraction AGN

In this chapter, we build upon the analysis of hard-only AGN presented in Chapter 4 and demonstrate a clear preference for clumpy torus models in low scattering fraction type-2 AGN, recover a higher photon index in line with values derived for type-1 AGN, and find a tentative link between the nature of the X-ray absorber and the host galaxy orientation and class. Despite the small sample size used here, clear results are presented for posterior distributions using BXA, demonstrating an ability to measure absorber geometry and orientation using X-ray spectroscopy. The analysis presents a new pathway to study the central engine, as the orientation can be studied or even inferred from X-ray reflection spectroscopy in highly obscured AGN, and yields a new way to study the connection between the central region and the large scale environment. Further investigation of larger samples of such sources will be crucial to determine if these results continue to hold, and to further explore the nature of low scattering fraction AGN.

Kapitel 6

Summary, conclusions and outlook

Active Galactic Nuclei powered by supermassive black holes which are actively accreting material are among the most luminous and energetic sources of electromagnetic radiation in the Universe. Through a variety of physical mechanisms, they produce emission across all wavelengths of light, from radio to γ-rays. Through study of these sources in the high energy, X-ray regime, the innermost emission regions of the AGN can be probed, and the extreme general relativistic effects can be studied. As seen in Chapter 1, X-ray emission in AGN is primarily produced in the hot X-ray corona, a population of relativistic electrons which Compton upscatter UV photons from the accretion disc surrounding the SMBH. These are then radiated isotropically from the corona, and can then be detected using specifically designed X-ray telescopes. These X-rays are absorbed by the layers of the Earth's atmosphere, and can therefore only be collected using space-based observatories.

With the launch of the eROSITA instrument onboard SRG (Predehl et al. 2021a; Merloni et al. 2012b), an all-sky X-ray census of AGN has helped yield new insights into their nature. Considering the half-sky area belonging to the eROSITA-DE consortium, over one million X-ray sources were detected in the first six months of all-sky survey alone, a vast majority of which are believed to be AGN. This thesis presents an analysis of AGN selected in the hard, or higher energy, X-ray regime; specifically in the $2.3 - 5\,\mathrm{keV}$ energy range. This energy range is selected to complement eROSITA's sharp sensitivity drop at $\sim 2.3\,\mathrm{keV}$ due to the gold coating placed on the mirrors, and consequently produces a sample of the brightest objects in the hard X-ray sky, a majority ($\sim 80\%$) of which are again believed to be AGN. This thesis, describes the process of creating AGN samples starting with the eROSITA catalogs, through careful processing of the X-ray data and matching with multiwavelength catalogs to perform source classification and characterisation, starting first with the eFEDS performance verification field, and then moving to the first all-sky survey.

In Chapter 2, a first look was provided of the properties of eROSITA hard X-ray selected AGN, through analysis of the eFEDS field (Brunner et al. 2022). The presented methodology for counterpart identification and source classification provided the framework for analysis of the larger, half-sky eRASS:1 data set in Chapter 4. The eFEDS hard sample was found to consist of mostly extragalactic sources, primarily a bright and local ($< z >= 0.34$) sample of 200 AGN, but spanning redshifts from 0.003 to 3.3. Visual inspection of

available optical spectra reveals that most show strong evidence of broad lines making them Seyfert 1s, while a number show strong host galaxy contributions and show broad lines only in Hα, and are likely Sy1.9 AGN. Preliminary analysis of the spectra using hardness ratios showed that the broad line sources showed softer spectra, while the harder spectra of the Sy 1.9 and photo-z AGN suggested absorption. Spectral analysis of these 200 AGN was then performed in Chapter 3, revealing a wealth of spectral complexity, including warm absorbers, neutral partial covering absorbers, secondary warm coronae and relativistic blurred reflection.

In Chapters 4 and 5, data from the first eROSITA all-sky survey, the first all-sky imaging survey in the $2-10$ keV band, were analysed. Chapter 4 presents the source identifications, counterparts, and classifications starting from the sample of 5466 hard X-ray selected sources presented in Merloni et al. (2024). A sample of 1328 AGN with spectroscopic redshifts is identified, mostly comprised of local AGN with $< z >= 0.19$ but extending to $z > 3$, similar to eFEDS. Comparison with Swift BAT Ricci et al. (2017); Oh et al. (2018) shows that eROSITA reliably detects type-1 AGN and sources with column densities up to $\sim 10^{22}\,\mathrm{cm}^{-2}$. A small sample of hard-only spectroscopic AGN is also identified, similarly as in eFEDS (Chapter 2), and these sources are found to be obscured AGN typically hosted in more edge-on galaxies which are classified as type-2s. This subsample and similar sources are explored in more detail in Chapter 5, where spectral modelling revealed a strong preference for a clumpy torus model, and a potential link between differences between spectral host galaxies and different morphological types.

A central theme of this thesis is the use of hard X-ray selected samples from eROSITA to explore AGN properties. The main reasons for the use of hard X-ray samples in this thesis are discussed throughout Chapters 1-5, and the consequences of this choice are central to the results from each chapter. First, the hard X-ray, $2.3-5$ keV selected sample is less sensitive to selection bias related to absorption; in soft X-ray selected samples, sources with column densities of $> 10^{22}\,\mathrm{cm}^{-2}$ are less likely to be detected, since the absorption heavily affects the spectral shape below ~ 2 keV. This effect is clear when examining the hard-only AGN, sources which appear in the eROSITA hard sample but not the main sample, as discussed in more detail in chapters 2, 4 and 5. This is also clear in Chapter 4, when comparing the column density distribution between hard sample and hard-only sample sources matched to Swift-BAT, where hard-only sources are found to have high column densities. It is also crucial for population studies, since including both obscured and unobscured AGN is important for proper demographics. Finally, modelling both obscured and unobscured AGN can help to identify potential differences in properties between different classes (see Chapter 5), which may support or contend with the AGN unification model (see Chapter 1). Despite these important points, studying hard X-rays with eROSITA is challenging and comes with caveats; eROSITA does not have sufficient high energy coverage to study most Compton thick AGN, as this requires good coverage at energies above 10 keV. eROSITA also misses some moderately obscured AGN which are too dim for detection. The relatively high background of eROSITA above 2 keV also impacts our ability to detect dimmer obscured AGN (Brunner et al. 2022). Other instruments, such as Swift-BAT and *NuSTAR* provide better hard X-ray energy coverage; however, *NuSTAR* has a small field of view

and cannot provide an all-sky census of AGN, and Swift-BAT lacks sufficient sensitivity to detect dim, high redshift sources. Future eROSITA all-sky surveys and next-generation missions will help to fill in the gaps left by existing missions/surveys, and simultaneous observations of sources with multiple X-ray instruments (e.g. XMM-Newton + *NuSTAR*) can also help to provide better constraints on parameter over a broadened energy range.

A second and perhaps the most important advantage of the hard X-ray selected sample is the ability to model the spectral shape above 2 keV. In this energy range, the hot X-ray corona is broadly expected to dominate the spectral shape. As has been described throughout these chapters, this component is crucial for understanding the inner structure and emission mechanisms in the AGN. By first measuring the photon index, a search can the be performed for additional spectral components, including absorption and emission, and models can more readily be compared. This presents an important advantage over spectral modelling with the eFEDS (Liu et al. 2022b) and later eRASS:1 main catalogs, where the lack of significant emission above 2 keV for most sources makes it more difficult to search for the presence of a soft excess. This point is particularly crucial for Chapter 3, where multiple different physical emission and absorption mechanisms are tested on eFEDS spectra (see also Waddell et al. 2024).

Another central theme, explored particularly in chapters 3 and 5, is on the importance of accurate spectral modelling in parameter estimation. In chapter 3, the analysis demonstrates that the estimation of the photon index and of the column density can vary dramatically depending on the model, and a careful study of the parameters presented in Nandra et al. (2024) and Waddell et al. (2024) demonstrates that significant changes are measured in photon indices for the same source are recovered when different models are applied. This is particularly true for sources which show strong evidence for complex (warm or partial covering) absorption, where very flat photon indices are typically recovered when fitting with a power law, and for sources with a strong soft excess component, which often show very steep photon indices when fitting with a power law. When performing population studies with AGN, it is therefore of importance to consider the spectral modelling, and to verify that the addition of a soft excess, absorber, or other model component doe not significantly improve the fit prior to using the measurements directly to infer population behaviour. The same result is found in Chapter 5, where the torus model used had a strong influence on the recovered photon index; crucially, through using smooth torus models, a flatter photon index more typical of type 2 AGN was recovered, but through using a clumpy torus model, a steeper index in line with type-1s was found. This demonstrates the importance of using the best fitting spectral model when comparing average properties between different groups of classes. This point becomes more and more important in the advent of new, state-of-the-art, high resolution X-ray missions including XRISM and Athena, where the high energy resolution spectra obtained should enable spectral modelling following the methodology presented in this thesis.

The lessons learned from this thesis could readily be applied to future population studies, particularly of AGN in X-rays. Throughout this thesis, BXA (Buchner et al. 2014) combined with background modelling (Simmonds et al. 2018) has proved invaluable in performing spectral fitting, parameter estimation, posterior determination, and model com-

parison. This is particularly relevant for scanning surveys or other data sets where each spectrum has a low number of counts due to shorter exposures, or for modelling dim sources with fewer X-ray counts. Model comparison is demonstrated to be a useful tool both for individual sources, as well as for better understanding the behaviour for larger samples (see Baronchelli et al. 2018, Chapter 3, Chapter 5). The posterior determination also readily enables posterior stacking (e.g. Baronchelli et al. 2018, Chapter 5), in order to better understand the distribution of parameters within the global population. AGN population studies should utilise the Bayesian framework to both spectroscopic and variability studies, combined with careful modelling of the coronal power law and subsequent testing for the presence of additional, carefully selected model components including emission and absorption, in order to truly understand accreting supermassive black holes in the Universe.

Literaturverzeichnis

Abdollahi, S., Acero, F., Ackermann, M., et al. 2020, ApJS, 247, 33

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

Aird, J., Alexander, D. M., Ballantyne, D. R., et al. 2015, ApJ, 815, 66

Aird, J., Coil, A. L., Moustakas, J., et al. 2012, ApJ, 746, 90

Akiyama, M., Ueda, Y., Ohta, K., Takahashi, T., & Yamada, T. 2003, ApJS, 148, 275

Akylas, A. & Georgantopoulos, I. 2021, A&A, 655, A60

Akylas, A., Georgantopoulos, I., Ranalli, P., et al. 2016, A&A, 594, A73

Almeida, A., Anderson, S. F., Argudo-Fernández, M., et al. 2023, arXiv e-prints, arXiv:2301.07688

Annuar, A., Alexander, D. M., Gandhi, P., et al. 2020, MNRAS, 497, 229

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

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

Arcodia, R., Campana, S., Salvaterra, R., & Ghisellini, G. 2018, A&A, 616, A170

Arcodia, R., Merloni, A., Nandra, K., & Ponti, G. 2019, A&A, 628, A135

Arévalo, P., Bauer, F. E., Puccetti, S., et al. 2014, ApJ, 791, 81

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

Awaki, H., Koyama, K., Inoue, H., & Halpern, J. P. 1991, PASJ, 43, 195

Baldry, I. K., Liske, J., Brown, M. J. I., et al. 2018, MNRAS, 474, 3875

Baldwin, J. 1997, in International Astronomical Union Colloquium, Vol. 159, Cambridge University Press, 80–95

Baldwin, J., Ferland, G. J., Korista, K., Hamann, F., & Dietrich, M. 2003, The Astrophysical Journal, 582, 590

Ballantyne, D. R. & Xiang, X. 2020, MNRAS, 496, 4255

Baloković, M., Brightman, M., Harrison, F. A., et al. 2018, ApJ, 854, 42

Baloković, M., García, J. A., & Cabral, S. E. 2019, Research Notes of the American Astronomical Society, 3, 173

Baronchelli, L., Nandra, K., & Buchner, J. 2018, MNRAS, 480, 2377

Barthelmy, S. D., Barbier, L. M., Cummings, J. R., et al. 2005, Space Science Reviews, 120, 143–164

Barua, S., Jithesh, V., Misra, R., et al. 2021, ApJ, 921, 46

Bauer, F. E., Arévalo, P., Walton, D. J., et al. 2015, ApJ, 812, 116

Bentz, M. C., Walsh, J. L., Barth, A. J., et al. 2010, ApJ, 716, 993

Blaes, O. 2014, Space Sci. Rev., 183, 21

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

Blustin, A. J., Page, M. J., & Branduardi-Raymont, G. 2004, A&A, 417, 61

Blustin, A. J., Page, M. J., Fuerst, S. V., Branduardi-Raymont, G., & Ashton, C. E. 2005, A&A, 431, 111

Boella, G., Butler, R., Perola, G., et al. 1997, Astronomy and Astrophysics Supplement Series, 122, 299

Böhringer, H., Chon, G., & Fukugita, M. 2017, Astronomy & Astrophysics, 608, A65

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

Boller, T., Freyberg, M., Buchner, J., et al. 2024, arXiv e-prints, arXiv:2401.17280

Boller, T., Liu, T., Weber, P., et al. 2021, A&A, 647, A6

Boller, T., Schmitt, J. H. M. M., Buchner, J., et al. 2022, A&A, 661, A8

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

Bongiorno, A., Merloni, A., Brusa, M., et al. 2012, MNRAS, 427, 3103

Bongiorno, A., Schulze, A., Merloni, A., et al. 2016, A&A, 588, A78

Bonson, K. & Gallo, L. C. 2016, MNRAS, 458, 1927

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

Bowen, I. S. & Vaughan, A. H. 1973, Appl. Opt., 12, 1430

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

Brightman, M. & Nandra, K. 2011, MNRAS, 413, 1206

Brightman, M., Nandra, K., Salvato, M., et al. 2014, MNRAS, 443, 1999

Brightman, M., Silverman, J. D., Mainieri, V., et al. 2013, MNRAS, 433, 2485

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

Brusa, M., Urrutia, T., Toba, Y., et al. 2022, A&A, 661, A9

Brusa, M., Zamorani, G., Comastri, A., et al. 2007, ApJS, 172, 353

Buchner, J. 2019, PASP, 131, 108005

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

Buchner, J. & Bauer, F. E. 2017, MNRAS, 465, 4348

Buchner, J., Brightman, M., Nandra, K., Nikutta, R., & Bauer, F. E. 2019, A&A, 629, A16

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

Bulbul, E., Liu, A., Kluge, M., et al. 2024, A&A, 685, A106

Cappi, M., Panessa, F., Bassani, L., et al. 2006, A&A, 446, 459

Chainakun, P. & Young, A. J. 2017, MNRAS, 465, 3965

Chalise, S., Lohfink, A. M., Chauhan, J., et al. 2022, arXiv e-prints, arXiv:2206.05237

Chang, Y. L., Arsioli, B., Giommi, P., Padovani, P., & Brandt, C. H. 2019, A&A, 632, A77

Ciliegi, P. & Maccacaro, T. 1996, Monthly Notices of the Royal Astronomical Society, 282, 477

Civano, F., Marchesi, S., Comastri, A., et al. 2016, ApJ, 819, 62

Cocchia, F., Fiore, F., Vignali, C., et al. 2007, A&A, 466, 31

Comastri, A. 2004, in Astrophysics and Space Science Library, Vol. 308, Supermassive Black Holes in the Distant Universe, ed. A. J. Barger, 245

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

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

Czerny, B. & Elvis, M. 1987, ApJ, 321, 305

Czerny, B., Wang, J.-M., Du, P., et al. 2019, The Astrophysical Journal, 870, 84

D'Abrusco, R., Álvarez Crespo, N., Massaro, F., et al. 2019, ApJS, 242, 4

Daly, R. A. & Sprinkle, T. B. 2014, MNRAS, 438, 3233

Dauser, T., Garcia, J., Parker, M. L., Fabian, A. C., & Wilms, J. 2014, MNRAS, 444, L100

Dauser, T., García, J., Walton, D. J., et al. 2016, A&A, 590, A76

Dauser, T., Svoboda, J., Schartel, N., et al. 2012, MNRAS, 422, 1914

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

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

Del Moro, A., Alexander, D. M., Aird, J. A., et al. 2017, ApJ, 849, 57

Della Ceca, R., Caccianiga, A., Severgnini, P., et al. 2008, A&A, 487, 119

Della Ceca, R., Maccacaro, T., Caccianiga, A., et al. 2004, A&A, 428, 383

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

Done, C., Davis, S. W., Jin, C., Blaes, O., & Ward, M. 2012, MNRAS, 420, 1848

Dwelly, T., Salvato, M., Merloni, A., et al. 2017, MNRAS, 469, 1065

Eckart, M. E., Stern, D., Helfand, D. J., et al. 2006, ApJS, 165, 19

Ehler, H. J. S., Gonzalez, A. G., & Gallo, L. C. 2018, MNRAS, 478, 4214

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

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

Fabian, A. C., Lohfink, A., Kara, E., et al. 2015, MNRAS, 451, 4375

Fabian, A. C., Miniutti, G., Gallo, L., et al. 2004, MNRAS, 353, 1071

Fabian, A. C., Rees, M. J., Stella, L., & White, N. E. 1989, MNRAS, 238, 729

Fabian, A. C., Zoghbi, A., Ross, R. R., et al. 2009, Nature, 459, 540

Fang, T., Buote, D., Bullock, J., & Ma, R. 2015, ApJS, 217, 21

Fiore, F., Brusa, M., Cocchia, F., et al. 2003, A&A, 409, 79

Flesch, E. W. 2023, The Open Journal of Astrophysics, 6, 49

Freyberg, M., Perinati, E., Pacaud, F., et al. 2020, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11444, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 114441O

Frogel, J. A. & Elias, J. H. 1987, ApJ, 313, L53

Furui, S., Fukazawa, Y., Odaka, H., et al. 2016, ApJ, 818, 164

Gallo, L. 2018, in Revisiting Narrow-Line Seyfert 1 Galaxies and their Place in the Universe, 34

Gallo, L. C., Gonzalez, A. G., & Miller, J. M. 2021, ApJ, 908, L33

Gallo, L. C., Gonzalez, A. G., Waddell, S. G. H., et al. 2019, MNRAS, 484, 4287

Gallo, L. C., Wilkins, D. R., Bonson, K., et al. 2015, MNRAS, 446, 633

García, J., Dauser, T., Reynolds, C. S., et al. 2013, ApJ, 768, 146

García, J. A., Kara, E., Walton, D., et al. 2019, ApJ, 871, 88

García-Burillo, S., Combes, F., Ramos Almeida, C., et al. 2016, ApJ, 823, L12

Gehrels, N., Chincarini, G., Giommi, P., et al. 2004, ApJ, 611, 1005

Genzel, R., Eckart, A., Ott, T., & Eisenhauer, F. 1997, MNRAS, 291, 219

Georgakakis, A., Aird, J., Schulze, A., et al. 2017, MNRAS, 471, 1976

Georgakakis, A. & Nandra, K. 2011, MNRAS, 414, 992

George, I. M., Turner, T. J., Netzer, H., et al. 1998, ApJS, 114, 73

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., Artis, E., et al. 2024, arXiv e-prints, arXiv:2402.08458

Giacconi, R., Gursky, H., Paolini, F. R., & Rossi, B. B. 1962, Physical Review Letters, 9, 439

Gierliński, M. & Done, C. 2004, MNRAS, 349, L7

Giommi, P., Brandt, C. H., Barres de Almeida, U., et al. 2019, A&A, 631, A116

Goodrich, R. W. 1989, ApJ, 342, 224

Goulding, A. D., Alexander, D. M., Bauer, F. E., et al. 2012, ApJ, 755, 5

Goulding, A. D., Greene, J. E., Setton, D. J., et al. 2023, ApJ, 955, L24

GRAVITY Collaboration, Amorim, A., Bauböck, M., et al. 2020a, A&A, 643, A154

GRAVITY Collaboration, Dexter, J., Shangguan, J., et al. 2020b, A&A, 635, A92

Gravity Collaboration, Sturm, E., Dexter, J., et al. 2018, Nature, 563, 657

Grupe, D. 2004, AJ, 127, 1799

Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332

Guo, H., Shen, Y., & Wang, S. 2018, PyQSOFit: Python code to fit the spectrum of quasars, Astrophysics Source Code Library

Gursky, H., Giacconi, R., Gorenstein, P., et al. 1966, Astrophysical Journal, vol. 146, p. 310-316, 146, 310

Haardt, F. & Maraschi, L. 1991, ApJ, 380, L51

Haardt, F. & Maraschi, L. 1993, ApJ, 413, 507

Harrison, F., Boggs, S., Christensen, F., et al. 2010, in

Harrison, F. A., Aird, J., Civano, F., et al. 2016, ApJ, 831, 185

Harrison, F. A., Craig, W. W., Christensen, F. E., et al. 2013, ApJ, 770, 103

Hasinger, G., Cappelluti, N., Brunner, H., et al. 2007a, ApJS, 172, 29

Hasinger, G., Cappelluti, N., Brunner, H., et al. 2007b, ApJS, 172, 29

Hazard, C., Mackey, M. B., & Shimmins, A. J. 1963, Nature, 197, 1037

Healey, S. E., Romani, R. W., Taylor, G. B., et al. 2007, ApJS, 171, 61

Heger, A., Fryer, C. L., Woosley, S. E., Langer, N., & Hartmann, D. H. 2003, ApJ, 591, 288

Hönig, S. F., Gandhi, P., Asmus, D., et al. 2014, MNRAS, 438, 647

Huchra, J. P., Macri, L. M., Masters, K. L., et al. 2012, ApJS, 199, 26

Ider Chitham, J., Comparat, J., Finoguenov, A., et al. 2020, MNRAS, 499, 4768

Ikeda, S., Awaki, H., & Terashima, Y. 2009, ApJ, 692, 608

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841

Ishibashi, W. & Courvoisier, T. J. L. 2010, A&A, 512, A58

Israel, W. 1967, Physical Review, 164, 1776

Itoh, R., Utsumi, Y., Inoue, Y., et al. 2020, ApJ, 901, 3

Jana, A., Chatterjee, A., Kumari, N., et al. 2020, MNRAS, 499, 5396

Jansen, F. et al. 2001, Astron. Astrophys., 365, L1

Jiang, J., Fabian, A. C., Dauser, T., et al. 2019, MNRAS, 489, 3436

Jones, D. H., Read, M. A., Saunders, W., et al. 2009, MNRAS, 399, 683

Jones, M. L., Hickox, R. C., Black, C. S., et al. 2016, ApJ, 826, 12

Kaastra, J. S., Mewe, R., Liedahl, D. A., Komossa, S., & Brinkman, A. C. 2000, A&A, 354, L83

Kallman, T. & Dorodnitsyn, A. 2019, ApJ, 884, 111

Kang, J.-L. & Wang, J.-X. 2023, On joint analysing XMM-NuSTAR spectra of active galactic nuclei

Kara, E., Alston, W. N., Fabian, A. C., et al. 2016, MNRAS, 462, 511

Kara, E., Fabian, A. C., Cackett, E. M., Miniutti, G., & Uttley, P. 2013, MNRAS, 430, 1408

Kaspi, S., Brandt, W. N., Netzer, H., et al. 2000a, ApJ, 535, L17

Kaspi, S., Smith, P. S., Netzer, H., et al. 2000b, ApJ, 533, 631

Kelly, B. C. & Shen, Y. 2013, ApJ, 764, 45

Khachikian, E. Y. & Weedman, D. W. 1974, ApJ, 192, 581

Kluge, M., Comparat, J., Liu, A., et al. 2024, A&A, 688, A210

Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv e-prints, arXiv:1711.03234

Koss, M., Trakhtenbrot, B., Ricci, C., et al. 2017, ApJ, 850, 74

Koss, M. J., Trakhtenbrot, B., Ricci, C., et al. 2022, ApJS, 261, 1

La Franca, F., Fiore, F., Vignali, C., et al. 2002, ApJ, 570, 100

Laha, S., Dewangan, G. C., Chakravorty, S., & Kembhavi, A. K. 2013, ApJ, 777, 2

Laha, S., Guainazzi, M., Dewangan, G. C., Chakravorty, S., & Kembhavi, A. K. 2014, MNRAS, 441, 2613

Laha, S., Markowitz, A. G., Krumpe, M., et al. 2020, ApJ, 897, 66

Lang, D. 2014, AJ, 147, 108

Lanzuisi, G., Civano, F., Elvis, M., et al. 2013, MNRAS, 431, 978

Laurenti, M., Piconcelli, E., Zappacosta, L., et al. 2022, A&A, 657, A57

Liu, A., Bulbul, E., Ghirardini, V., et al. 2022a, A&A, 661, A2

Liu, T., Buchner, J., Nandra, K., et al. 2022b, A&A, 661, A5

Liu, T., Merloni, A., Comparat, J., et al. 2022c, A&A, 661, A27

Liu, T., Merloni, A., Sanders, J., et al. 2024, ApJ, 967, L27

Liu, T., Tozzi, P., Wang, J.-X., et al. 2017, ApJS, 232, 8

Liu, Y. & Li, X. 2014, ApJ, 787, 52

Liu, Z., Merloni, A., Georgakakis, A., et al. 2016, MNRAS, 459, 1602

Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, ApJS, 228, 2

Madejski, G. G. & Sikora, M. 2016, ARA&A, 54, 725

Magdziarz, P. & Zdziarski, A. A. 1995, MNRAS, 273, 837

Maiolino, R., Risaliti, G., & Salvati, M. 1999, A&A, 341, L35

Malkan, M. A. 1983, ApJ, 268, 582

Marchesi, S., Lanzuisi, G., Civano, F., et al. 2016, ApJ, 830, 100

Marinucci, A., Muleri, F., Dovciak, M., et al. 2022, MNRAS, 516, 5907

Markowitz, A. G., Krumpe, M., & Nikutta, R. 2014, MNRAS, 439, 1403

Masetti, N., Pretorius, M. L., Palazzi, E., et al. 2006, A&A, 449, 1139

Mason, R. E., Levenson, N. A., Shi, Y., et al. 2009, ApJ, 693, L136

Massaro, E., Maselli, A., Leto, C., et al. 2015, Ap&SS, 357, 75

Massaro, F., Masetti, N., D'Abrusco, R., Paggi, A., & Funk, S. 2014, AJ, 148, 66

McClintock, J. E., Narayan, R., & Steiner, J. F. 2014, Space Sci. Rev., 183, 295

McKernan, B., Yaqoob, T., & Reynolds, C. S. 2007, MNRAS, 379, 1359

Merloni, A. & Fabian, A. C. 2003, MNRAS, 342, 951

Merloni, A., Fabian, A. C., & Ross, R. R. 2000, MNRAS, 313, 193

Merloni, A., Lamer, G., Liu, T., et al. 2024, A&A, 682, A34

Merloni, A., Predehl, P., Becker, W., et al. 2012a, arXiv e-prints, arXiv:1209.3114

Merloni, A., Predehl, P., Becker, W., et al. 2012b, arXiv e-prints, arXiv:1209.3114

Middei, R., Bianchi, S., Marinucci, A., et al. 2019, A&A, 630, A131

Middei, R., Matzeu, G. A., Bianchi, S., et al. 2021, A&A, 647, A102

Middei, R., Vagnetti, F., Bianchi, S., et al. 2017, Astronomy amp; Astrophysics, 599, A82

Mitsuda, K., Bautz, M., Inoue, H., et al. 2007, PASJ, 59, S1

Miyakawa, T., Ebisawa, K., & Inoue, H. 2012, PASJ, 64, 140

Mizumoto, M., Done, C., Tomaru, R., & Edwards, I. 2019, MNRAS, 489, 1152

Morgan, C. W., Kochanek, C. S., Dai, X., Morgan, N. D., & Falco, E. E. 2008, ApJ, 689, 755

Mosquera, A. M., Kochanek, C. S., Chen, B., et al. 2013, ApJ, 769, 53

Motch, C. 2001, in AIP Conference Proceedings, Vol. 599 (AIP), 244–253

Murphy, K. D. & Yaqoob, T. 2009, MNRAS, 397, 1549

Nandra, K., George, I. M., Mushotzky, R. F., Turner, T. J., & Yaqoob, T. 1997a, ApJ, 476, 70

Nandra, K., George, I. M., Mushotzky, R. F., Turner, T. J., & Yaqoob, T. 1997b, ApJ, 477, 602

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. 1994a, MNRAS, 268, 405

Nandra, K. & Pounds, K. A. 1994b, MNRAS, 268, 405

Nandra, K., Waddell, S. G. H., Liu, T., et al. 2024, arXiv e-prints, arXiv:2401.17300

Nenkova, M., Ivezić, Ž., & Elitzur, M. 2002, ApJ, 570, L9

Nenkova, M., Sirocky, M. M., Nikutta, R., Ivezić, Ž., & Elitzur, M. 2008, ApJ, 685, 160

Netzer, H. 2015, ARA&A, 53, 365

Oh, K., Koss, M., Markwardt, C. B., et al. 2018, ApJS, 235, 4

Oppenheimer, J. R. & Snyder, H. 1939, Physical Review, 56, 455

Osterbrock, D. E. 1981, ApJ, 249, 462

Osterbrock, D. E. & Pogge, R. W. 1985, ApJ, 297, 166

Özel, F., Psaltis, D., Narayan, R., & McClintock, J. E. 2010, ApJ, 725, 1918

Padovani, P., Alexander, D., Assef, R., et al. 2017, The Astronomy and Astrophysics Review, 25, 1

Pal, I., Stalin, C. S., Chatterjee, R., & Agrawal, V. K. 2023, Journal of Astrophysics and Astronomy, 44, 87

Panagiotou, C. & Walter, R. 2019, Astronomy amp; Astrophysics, 626, A40

Panagiotou, C., Walter, R., & Paltani, S. 2021, Astronomy amp; Astrophysics, 653, A162

Papaloizou, J. C. B. & Lin, D. N. C. 1995, ARA&A, 33, 505

Parker, M. L., Alston, W. N., Härer, L., et al. 2021, MNRAS, 508, 1798

Parker, M. L., Fabian, A. C., Matt, G., et al. 2015, MNRAS, 447, 72

Parker, M. L., Matzeu, G. A., Matthews, J. H., et al. 2022, MNRAS, 513, 551

Parker, M. L., Schartel, N., Komossa, S., et al. 2014, MNRAS, 445, 1039

Pavlinsky, M., Tkachenko, A., Levin, V., et al. 2021, A&A, 650, A42

Peterson, B. M. 1993, PASP, 105, 247

Peterson, B. M. 2006, in Physics of Active Galactic Nuclei at all Scales (Springer), 77–100

Peterson, B. M. 2014, Space Sci. Rev., 183, 253

Peterson, B. M. & Wandel, A. 1999, ApJ, 521, L95

Peterson, B. M. & Wandel, A. 2000, ApJ, 540, L13

Petrucci, P. O., Gronkiewicz, D., Rozanska, A., et al. 2020, A&A, 634, A85

Petrucci, P. O., Haardt, F., Maraschi, L., et al. 2001, ApJ, 556, 716

Petrucci, P. O., Ursini, F., De Rosa, A., et al. 2018, A&A, 611, A59

Piccinotti, G., Mushotzky, R. F., Boldt, E. A., et al. 1982, ApJ, 253, 485

Piconcelli, E., Jimenez-Bailón, E., Guainazzi, M., et al. 2005, A&A, 432, 15

Pierre, M., Pacaud, F., Adami, C., et al. 2016, A&A, 592, A1

Pounds, K. A., Done, C., & Osborne, J. P. 1995, MNRAS, 277, L5

Pravdo, S. H., Nugent, J. J., Nousek, J. A., et al. 1981, ApJ, 251, 501

Predehl, P., Andritschke, R., Arefiev, V., et al. 2021a, A&A, 647, A1

Predehl, P., Andritschke, R., Arefiev, V., et al. 2021b, A&A, 647, A1

Predehl, P., Sunyaev, R. A., Becker, W., et al. 2020, Nature, 588, 227

Pringle, J. E. & Rees, M. J. 1972, A&A, 21, 1

Reeves, J. N. & Turner, M. J. L. 2000, MNRAS, 316, 234

Reis, R. C. & Miller, J. M. 2013, ApJ, 769, L7

Remillard, R. A. & McClintock, J. E. 2006, ARA&A, 44, 49

Revnivtsev, M., Sazonov, S., Jahoda, K., & Gilfanov, M. 2004, A&A, 418, 927

Reynolds, C. S. 2019, Nature Astronomy, 3, 41

Reynolds, C. S. 2021, ARA&A, 59, 117

Ricci, C., Tazaki, F., Ueda, Y., et al. 2014, ApJ, 795, 147

Ricci, C., Trakhtenbrot, B., Koss, M. J., et al. 2017, ApJS, 233, 17

Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006, ApJS, 166, 470

Risaliti, G., Elvis, M., Fabbiano, G., et al. 2007, ApJ, 659, L111

Risaliti, G., Elvis, M., & Nicastro, F. 2002, ApJ, 571, 234

Risaliti, G., Maiolino, R., & Salvati, M. 1999, ApJ, 522, 157

Risaliti, G., Nardini, E., Salvati, M., et al. 2011, MNRAS, 410, 1027

Risaliti, G., Young, M., & Elvis, M. 2009, ApJ, 700, L6

Rivers, E., Markowitz, A., & Rothschild, R. 2013, ApJ, 772, 114

Robinson, D. C. 1975, Phys. Rev. Lett., 34, 905

Ross, R. R. & Fabian, A. C. 2005, MNRAS, 358, 211

Ross, R. R., Fabian, A. C., & Young, A. J. 1999, MNRAS, 306, 461

Rothschild, R., Boldt, E., Holt, S., et al. 1979, Space Science Instrumentation, 4, 269

Rybicki, G. B. & Lightman, A. P. 1979, Radiative processes in astrophysics

Saha, T., Markowitz, A. G., & Buchner, J. 2022, MNRAS, 509, 5485

Salpeter, E. E. 1964, ApJ, 140, 796

Salvato, M., Buchner, J., Budavári, T., et al. 2018, MNRAS, 473, 4937

Salvato, M., Hasinger, G., Ilbert, O., et al. 2009, ApJ, 690, 1250

Salvato, M., Ilbert, O., Hasinger, G., et al. 2011, ApJ, 742, 61

Salvato, M., Wolf, J., Dwelly, T., et al. 2022, A&A, 661, A3

Salviander, S., Shields, G. A., Gebhardt, K., & Bonning, E. W. 2007, The Astrophysical Journal, 662, 131

Sandage, A. 1965, ApJ, 141, 1560

Sanfrutos, M., Miniutti, G., Agís-González, B., et al. 2013, MNRAS, 436, 1588

Saxton, R. D., Read, A. M., Esquej, P., et al. 2008, A&A, 480, 611

Schartel, N., Schmidt, M., Fink, H. H., Hasinger, G., & Truemper, J. 1997, A&A, 320, 696

Schmidt, M. 1963, Nature, 197, 1040

Schmidt, M. 1978, , 17, 329

Schneider, P. C., Freund, S., Czesla, S., et al. 2022, A&A, 661, A6

Schulze, A., Bongiorno, A., Gavignaud, I., et al. 2015, MNRAS, 447, 2085

Scott, A. E., Stewart, G. C., & Mateos, S. 2012, MNRAS, 423, 2633

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

Shemmer, O., Brandt, W. N., Netzer, H., Maiolino, R., & Kaspi, S. 2008, ApJ, 682, 81

Shen, Y. 2013, Bulletin of the Astronomical Society of India, 41, 61

Shen, Y. & Ho, L. C. 2014, Nature, 513, 210

Shen, Y. & Liu, X. 2012, The Astrophysical Journal, 753, 125

Shen, Y., Liu, X., Greene, J. E., & Strauss, M. A. 2011, ApJ, 735, 48

Shimizu, T. T., Davies, R. I., Koss, M., et al. 2018, ApJ, 856, 154

Shu, Y., Koposov, S. E., Evans, N. W., et al. 2019, MNRAS, 489, 4741

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

Soltan, A. 1982, MNRAS, 200, 115

Storey-Fisher, K., Hogg, D. W., Rix, H.-W., et al. 2023, arXiv e-prints, arXiv:2306.17749

Suh, H., Civano, F., Trakhtenbrot, B., et al. 2020, ApJ, 889, 32

Sunyaev, R., Arefiev, V., Babyshkin, V., et al. 2021, A&A, 656, A132

Sunyaev, R. A. & Truemper, J. 1979, Nature, 279, 506

Sutherland, W. & Saunders, W. 1992, MNRAS, 259, 413

Tan, Y., Wang, J. X., Shu, X. W., & Zhou, Y. 2012, ApJ, 747, L11

Tanaka, Y., Boller, T., Gallo, L., Keil, R., & Ueda, Y. 2004, PASJ, 56, L9

Tanaka, Y., Inoue, H., & Holt, S. S. 1994, PASJ: Publications of the Astronomical Society of Japan (ISSN 0004-6264), vol. 46, no. 3, p. L37-L41, 46, L37

Tanaka, Y., Nandra, K., Fabian, A. C., et al. 1995, Nature, 375, 659

Tashiro, M., Maejima, H., Toda, K., et al. 2020, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11444, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1144422

Taylor, B., Andresen, R., Peacock, A., & Zobl, R. 1981, in X-Ray Astronomy: Proceedings of the XV ESLAB Symposium held in Amsterdam, The Netherlands, 22–26 June 1981, Springer, 479–494

Thorne, K. S. 1974, ApJ, 191, 507

Tombesi, F., Cappi, M., Reeves, J. N., et al. 2010, A&A, 521, A57

Tortosa, A., Bianchi, S., Marinucci, A., Matt, G., & Petrucci, P. O. 2018, A&A, 614, A37

Trakhtenbrot, B., Ricci, C., Koss, M. J., et al. 2017, MNRAS, 470, 800

Tripathi, S., Waddell, S. G. H., Gallo, L. C., Welsh, W. F., & Chiang, C. Y. 2019, MNRAS, 488, 4831

Truemper, J. 1982, Advances in Space Research, 2, 241

Trump, J. R., Sun, M., Zeimann, G. R., et al. 2015, ApJ, 811, 26

Tsunemi, H., Kitamoto, S., Manabe, M., et al. 1989, Astronomical Society of Japan, Publications (ISSN 0004-6264), vol. 41, no. 3, 1989, p. 391-403., 41, 391

Tsuzuki, Y., Kawara, K., Yoshii, Y., et al. 2006, The Astrophysical Journal, 650, 57

Turner, T. J., George, I. M., Nandra, K., & Mushotzky, R. F. 1997, ApJ, 488, 164

Turner, T. J. & Pounds, K. A. 1989, MNRAS, 240, 833

Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, ApJ, 786, 104

Ueda, Y., Eguchi, S., Terashima, Y., et al. 2007, ApJ, 664, L79

Ueda, Y., Takahashi, T., Ishisaki, Y., Ohashi, T., & Makishima, K. 1999, ApJ, 524, L11

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

Vaona, L., Ciroi, S., Di Mille, F., et al. 2012, MNRAS, 427, 1266

Vasudevan, R. V., Fabian, A. C., Reynolds, C. S., et al. 2016, MNRAS, 458, 2012

Vaughan, S., Edelson, R., Warwick, R. S., & Uttley, P. 2003, MNRAS, 345, 1271

Vestergaard, M. & Peterson, B. M. 2006, The Astrophysical Journal, 641, 689

Vestergaard, M. & Wilkes, B. J. 2001, The Astrophysical Journal Supplement Series, 134, 1

Waddell, S. G. H., Buchner, J., Nandra, K., et al. 2024, The SRG/eROSITA all-sky survey: Hard X-ray selected Active Galactic Nuclei

Waddell, S. G. H. & Gallo, L. C. 2020, MNRAS, 498, 5207

Waddell, S. G. H. & Gallo, L. C. 2022, MNRAS, 510, 4370

Waddell, S. G. H., Gallo, L. C., Gonzalez, A. G., Tripathi, S., & Zoghbi, A. 2019, MNRAS, 489, 5398

Waddell, S. G. H., Nandra, K., Buchner, J., et al. 2024, arXiv e-prints, arXiv:2306.00961

Walton, D. J., Nardini, E., Gallo, L. C., et al. 2019, MNRAS, 484, 2544

Webster, B. L. & Murdin, P. 1972, Nature, 235, 37

Weisskopf, M. C., Tananbaum, H. D., Van Speybroeck, L. P., & O'Dell, S. L. 2000, in X-Ray Optics, Instruments, and Missions III, ed. J. E. Truemper & B. Aschenbach (SPIE)

Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, A&AS, 143, 9

Wilkes, B. J., Mathur, S., Fiore, F., Antonelli, A., & Nicastro, F. 2001, ApJ, 549, 248

Wilkins, D. R. & Gallo, L. C. 2015, MNRAS, 449, 129

Wilkins, D. R., Gallo, L. C., Silva, C. V., et al. 2017, MNRAS, 471, 4436

Willingale, R., Starling, R. L. C., Beardmore, A. P., Tanvir, N. R., & O'Brien, P. T. 2013, MNRAS, 431, 394

Wilms, J., Allen, A., & McCray, R. 2000, ApJ, 542, 914

Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, PASP, 131, 055001

Wolf, J., Salvato, M., Coffey, D., et al. 2020, MNRAS, 492, 3580

Woo, J.-H. & Urry, C. M. 2002, ApJ, 579, 530

Woosley, S. E. & Weaver, T. A. 1995, ApJS, 101, 181

Wu, Q. & Shen, Y. 2022, ApJS, 263, 42

Xu, W., Ramos-Ceja, M. E., Pacaud, F., Reiprich, T. H., & Erben, T. 2022, Astronomy amp; Astrophysics, 658, A59

Xu, Y., García, J. A., Walton, D. J., et al. 2021, ApJ, 913, 13

Yamada, S., Ueda, Y., Tanimoto, A., et al. 2019, ApJ, 876, 96

Zdziarski, A. A., Johnson, W. N., & Magdziarz, P. 1996, MNRAS, 283, 193

Zel'dovich, Y. B. & Novikov, I. D. 1971, Theory of gravitation and the evolution of stars.

Zhang, S. N., Cui, W., & Chen, W. 1997, ApJ, 482, L155

Zhao, X., Marchesi, S., Ajello, M., et al. 2021, A&A, 650, A57

Zheng, X. C., Xue, Y. Q., Brandt, W. N., et al. 2017, The Astrophysical Journal, 849, 127

Zhou, Y. Y., Yu, K. N., Young, E. C. M., Wang, J. M., & Ma, E. 1997, ApJ, 475, L9

Zoghbi, A., Fabian, A. C., & Gallo, L. C. 2008, MNRAS, 391, 2003

Zoghbi, A., Reynolds, C., Cackett, E. M., et al. 2013, ApJ, 767, 121

Życki, P. T., Done, C., & Smith, D. A. 1999, MNRAS, 309, 561

Acknowledgements

The journey of my PhD has been a story of trying, failing, learning, forgetting, winning, losing and persevering through it all, with an incredible group of people at my side. I have come a long way from moving across continents in the middle of a global pandemic, to tears when leaving a city that four years later truly felt like home.

I would like to thank Paul Nandra for the quarantini celebrations, many useful scientific discussions, long and productive meetings, and hard work reading and editing this thesis. I have learned so much from you and am grateful for all the time spent on furthering my education and development.

Thanks to Johannes Buchner, for never once failing to answer my pleas for help, for his patience in repeatedly explaining numerous concepts, and for contributing enormously to my growth as a writer and as a scientist.

To Andrea Merloni, Thomas Boller, and Maria Diaz-Trigo, thank you for the many useful dicussions over the years that lead to the completion of this thesis.

To Mara Salvato, Teng Liu, and so many others within eROAGN, thank you for listening and giving feedback to the endless updates on soft excesses ad hard X-ray selected samples.

Thanks to all of my co-authors, referees, and others who gave comments on my papers, helping me work so hard to pull them over the finish line to submission.

To the HEG PhD students and postdocs, past and current; Adam, Alena, Ana, Ania, Bob, Catarina, David, Emre, Francesco, Isabelle, Iuliia, Julien, Michael, Pietro, Riccardo A., Riccardo S., Silas, Soumya, Toska, William, Xueying, Yi, Zsofi, and anyone else I've forgotten, thank you for the endless lunchtime conversations about food, too-long coffee breaks, the jokes, the laughter, the celebrations, and the lovely environment you all created in our research group. Thank you for all the lovely outings and adventures over the years. Special thanks to Emre, Francesco, Fulvio, Riccardo and Soumya, for forming such a strong friendship during the craxy COVID days that will last for decades to come.

Soumya and Zsofi, thank you for being the most beautiful people in the most beautiful office, full of plants and sunsets and much needed gossip sessions. I will miss it dearly!

Thanks to David for his incredible help with the German component of my abstract, because despite living in Munich for four years, my German astrophysics vocabulary does not extend beyond *Schwarze Löcker*.

Endless gratitude is also due again to Soumya and Fulvio for printing and binding my thesis for submission at LMU, I never would have submitted without you!

Thanks to two of my longtime Canadian best friends, Liz and Lucie, whose visits

brought so much joy to my four years abroad and reminding me of the joys of life outside the PhD.

To my fiancé (almost husband!), James. Thank you for never leaving my side over this journey, bother figuratively and literally as we made our home in a studio apartment in lockdown. I have the incredible joy and honour to love you more and more every day, and have been so proud watching you finish you Master's and now start a PhD journey of your own. I can't wait to share adventures together for the rest of our lives.

Last but certainly not least, thanks to my loving family; Tanya, Greg, Lucas and Charlotte. I have missed you terribly for the last four years, but your visits, messages, and video calls made the distance seem bearable, and my apartment felt more like home with you all there. Thank you for always believing in me, for supporting me, and for listening to my endless stories and commenting on my endless photos. Thank you for the 26 years of your love. I love you too!

She's swinging up to the stars
And she just passed Mars.
She better look around soon.

Because if she's not careful
To watch where she's going.
I might just bump into the moon.

- Swingin' Up to the Stars, Philip Parker