Active Galactic Nuclei: accretion and feedback across the mass scale with eROSITA and LOFAR

Zsofi Igo

SIGILL·UNIVERSI·TAT·LUDOVIC·MAXIM

München 2026

Active Galactic Nuclei: accretion and feedback across the mass scale with eROSITA and LOFAR

Zsofi Igo

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

vorgelegt von
Zsofi Igo
aus Eger, Ungarn

München, den 25.11.2025

Erstgutachter: Prof. Dr. Kirpal Nandra
Zweitgutachter: Dr. Céline Péroux
Tag der mündlichen Prüfung: 02.02.2026

To Mom and Dad and loved ones, near and far.
For readers looking to get the full experience, the soundtrack of
this Thesis includes: Ludovico Einaudi, Alpha Waves,
Two Steps From Hell and other inspirational movie theme songs.

Contents

List of Figures

List of Tables

Zusammenfassung

Supermassereiche Schwarze Löcher (SMBHs) befinden sich im Zentrum nahezu jeder massereichen Galaxie. Etwa $1 - 10\%$ der SMBHs akkretieren Materie aus ihrer Umgebung und werden dadurch zu aktiven Galaxienkernen (AGN). AGN zählen zu den leuchtkräftigsten und dauerhaftesten Quellen radiativer und kinetischer Energie im gesamten Universum. Ihre Strahlung, die sich über das gesamte elektromagnetische Spektrum von Radiowellen bis hin zu Röntgenstrahlen erstreckt und in Form mächtiger Ausströmungen freigesetzt wird, ist so stark, dass sie die Wirtsgalaxie erheblich beeinflussen kann. Dieser Prozess wird als 'AGN-Feedback' bezeichnet. AGN-Feedback tritt in verschiedenen Formen auf, etwa in Form von Winden oder gebündelten Jets. Diese enstehen in der Nähe des Schwarzen Lochs und können sich auf größere Skalen ausbreiten, Gas und Staub verdrängen oder erhitzen und dadurch die Sternentstehung unterdrücken. Obwohl AGN eine entscheidende Rolle für das Verständnis der Galaxienentwicklung spielen, ist bislang unklar, wie die einfallende Akkretionsenergie und die abgegebene Feedback-Energie von den Eigenschaften des AGN und seiner Wirtsgalaxie abhängen und wie sie sich im Laufe der kosmischen Zeit verändern.

Aus diesem Grund beginne ich diese Dissertation mit der Frage, wie die Feedbackemission von Jets von der stellaren Masse der Galaxie abhängt. AGN-Jets emittieren hauptsächlich Synchrotronstrahlung, welche bei Radiowellenlängen messbar ist. Daher können wir erdgebundene Radioobservatorien wie das Low Frequency Array (LOFAR) nutzen, um große Stichproben von AGN, welche jets aufweisen, zu untersuchen. Ich präsentiere eine neue statistische Stichprobe von 764 LOFAR-Radio-AGN bei niedriger Rotverschiebung ($z < 0{,}4$) und bestimme erstmals den Anteil der Galaxien, welche solche Radio-AGN beherbergen (d. h. die Radio-AGN-Inzidenz), als Funktion der spezifischen kinetischen Leistung des Schwarzen Lochs, $\lambda_{\rm Jet}$, einem Maß dafür, wie stark die Jets im Verhältnis zur Masse der Galaxie (und damit indirekt zur Masse des Schwarzen Lochs) sind. Meine Ergebnisse zeigen, dass die Radio-AGN-Inzidenz eine deutliche Massenabhängigkeit nachweist: Massereichere Galaxien beherbergen mit höherer Wahrscheinlichkeit einen Radio-AGN bei allen Werten von $\lambda_{\rm Jet}$. Dies ist bemerkenswert, da frühere Arbeiten zur Inzidenz von Röntgen-AGN gezeigt haben, dass die Mechanismen, welche AGN-Aktivität in massereichen Galaxien auslösen und antreiben, weitgehend massenunabhängig sind (zumindest in erster Näherung). Es ist daher auffällig, dass Jet-Feedback, welches vermutlich durch denselben Akkretionsprozess gespeist wird, ein anderes Verhalten zeigt. Um dieses Ergebnis zu erklären, untersuche ich die Kopplung zwischen Akkretionsscheibe und Jet in AGN, indem ich die statistischen Inzidenzverteilungen von Röntgen- und Radio-AGN vergleiche.

Neben der Massenabhängigkeit zeigt die Radio-AGN-Inzidenz auch deutliche Unterschiede zwischen verschiedenen Radiomorphologien. Kompakte Radio-AGN weisen eine steile, potenzgesetzartige Inzidenz auf. Dies bedeutet, dass sie bei niedrigen Jet-Leistungen dominieren und bei höheren Jet-Leistungen schnell abfallen. Im Gegensatz dazu zeigen komplexe Radio-AGN, d.h. solche mit ausgedehnter, möglicherweise doppellappiger Morphologie, eine flachere Inzidenzverteilung mit einer deutlich stärkeren Massenabhängigkeit. Dies deutet darauf hin, dass sie bei größeren stellaren Massen und höheren λ_{Jet}-Werten zunehmend dominieren. Dank der gut charakterisierten, vollständigen, spektroskopischen Radio-AGN-Stichprobe verfüge ich über detaillierte Kenntniss sowohl der Wirtsgalaxienmassen als auch der Jet-Leistungen einzelner Radio-AGN, sowie darüber, wie häufig ein Radio-AGN einer bestimmten Jet-Leistung in Galaxien einer gegebenen Masse vorkommt. Unter Ausnutzung dieser drei Größen bestimme ich die mittlere Jet-Leistung der Population massereicher Galaxien bei niedriger Rotverschiebung. Ich zeige, dass das kinetische Feedback durch Radio-AGN gegenüber allen plausiblen Formen von strahlungsgetriebenem Feedback, wie z.B. den zuvor erwähnten Winden, in massereichen Galaxien dominiert. Interessanterweise sind es die kompakten Radio-AGN, die dieses globale kinetische Energiebudget bestimmen, mit Ausnahme der massereichsten Galaxien. Darüber hinaus vergleiche ich die durchschnittlich injizierte Jet-Energie mit der Bindungsenergie der Galaxien und ihrer umgebenden Dunklen-Materie-Halos sowie mit der gesamten thermischen Energie des Gases innerhalb dieser Halos. Ich zeige, dass Radio-AGN, egal ob kompakt oder komplex, zwar nicht genügend Energie besitzen, um die globale Gasverteilung massereicher Galaxien vollständig zu stören, aber eine bedeutende Wärmequelle darstellen, welche das lokale thermodynamische Gleichgewicht selbst in den Kernen der massereichsten Galaxienhaufen im lokalen Universum beeinflussen kann.

Abschließend untersuche ich die Details des Akkretionsprozesses in einem bislang unerforschten Parameterraum von Galaxien mit niedriger Masse bei niedriger Rotverschiebung. Dies wurde durch die jüngsten Fortschritte in tiefen, multiwellenlängen, Himmelsdurchmusterungen möglich. Insbesondere nutze ich in dieser Dissertation die vierte, tieftse Durchmusterung des eROSITA All-Sky Survey (eRASS:4) in Kombination mit optischen Bilddaten aus dem zehnten Datenrelease des DESI Legacy Imaging Survey (LS10), um die Inzidenz von Röntgen-AGN in Abhängigkeit der spezifischen radiativen Leistung des Schwarzen Lochs, λ_{Edd}, zu untersuchen. Dies ist ein komplementärer, massenskalierender Leistungsindikator zu λ_{Jet}. Ausgehend von einer optisch ausgewählten Stichprobe aus LS10 von über fünf Millionen Galaxien über den gesamten Massenbereich hinweg identifiziere ich 874 Röntgen-AGN in Galaxien mit niedriger Masse ($\log M_*/M_\odot \leq 10$), von denen mehr als 600 neu entdeckt sind. Mithilfe einer bayesianischen Methode, welche die Röntgeninformationen aller Galaxien der Ausgangsstichprobe berücksichtigt, kann ich die bislang engsten Einschränkungen der Verteilungen spezifischer Akkretionsraten (zumindest am hochakkretierenden Ende) ableiten und dabei massenabhängige Eigenschaften zweiter Ordnung aufdecken. Anschließend leite ich daraus die kumulative AGN-Fraktion als Funktion der stellaren Masse ab. Interessanterweise finde ich ein Maximum dieser kumulativen AGN-Fraktion bei $\log M_*/M_\odot \sim 10 - 10.5$ für moderat und stark akkretiertende Röntgen-AGN. Dies weist möglicherweise auf eine abnehmende Effizienz der AGN-Entstehung bei sehr niedrigen und sehr hohen Massen hin. Außerdem fällt dieses Maximum genau in den Bereich, in dem die Sternentstehungseffizienz bei einer bestimmten Halomasse am größten ist. Das deutet darauf hin, dass eine gemeinsame physikalische Ursache sowohl die AGN-Akkretion als auch die Sternentstehung beeinflusst. Zukünftige detaillierte Röntgenspektral- und Timing-Analysen dieser neuen Stichprobe von Röntgen-AGN könnten neue Erkenntnisse über die Details von Akkretion und Feedback im Niedermassenregime liefern.

Abstract

Supermassive black holes (SMBHs) are found at the centre of almost every massive galaxy. Around $1-10\%$ of SMBHs are accreting matter from their surroundings, making them *active* galactic nuclei (AGN). AGN are some of the most luminous, persistent emitters of radiative and kinetic energy in the entire Universe. This energy, released across the electromagnetic spectrum from radio to X-ray wavelengths and through powerful outflows, can significantly impact the host galaxy, in a process called 'AGN feedback'. AGN feedback occurs in various forms, such as wide-angle winds or collimated jets launched from the vicinity of the black hole that propagate to larger scales, displacing or heating gas and dust, and thereby suppressing star formation. While AGN are a crucial ingredient to understanding galaxy evolution, it is still unclear how the accretion energy coming in and the feedback energy coming out depend on AGN and host galaxy properties, as well as how they vary across cosmic time.

As such, I begin this Thesis by tackling the question of how jetted feedback emission depends on the stellar mass of the galaxy. AGN jets primarily emit via synchrotron radiation, observable at radio wavelengths. Therefore, we can use radio observatories here on Earth, such as the Low Frequency Array (LOFAR), to probe large samples of jetted AGN. I compile a new statistical sample of 764 LOFAR radio AGN at low redshift ($z < 0.4$) and calculate for the first time the fraction of galaxies hosting such radio AGN (i.e. the radio AGN incidence) as a function of specific black hole kinetic power, λ_{Jet}, a measure of how powerful the jets are compared to the mass of the galaxy (a proxy for the mass of the black hole). I find that the incidence of radio AGN shows a mass-dependence, whereby higher mass galaxies are more likely to host a radio AGN at all values of λ_{Jet}. This is interesting as previous work investigating the incidence of X-ray AGN find that the mechanisms responsible for triggering and fuelling AGN across all massive galaxies are mass-invariant (to first order). Therefore, it is curious that jetted feedback, supposedly fuelled by the accretion process, behaves differently. I try to reconcile this by analysing the coupling between the accretion disk and jet in AGN using the statistical incidence distributions of X-ray and radio AGN.

In addition to the mass-dependence, the radio AGN incidence shows a striking variation for different radio AGN morphologies. Compact radio AGN show a steep power-law-like incidence, meaning that they dominate at lower jet powers and drop out rapidly at higher jet powers. In contrast, complex radio AGN, those with extended, possible double-lobed morphologies, show a flatter incidence distribution with a much stronger mass-dependence, indicating that they progressively dominate for higher stellar mass and $\lambda_{\rm Jet}$ values. Thanks to the well-characterised, complete, spectroscopic radio AGN sample, I have detailed knowledge of the host galaxy mass, the jet power of a given radio AGN and how often this radio AGN of a given jet power populates a galaxy of a given mass. Capitalising on these three factors, I derive the average jet power of the population of massive galaxies at low redshift and find that the kinetic feedback from radio AGN dominates over any plausible inventory of radiatively-driven feedback, such as the aforementioned winds. Interestingly, it is the compact radio AGN that dominate this global kinetic energy budget for all but the most massive galaxies. I also compare this average injected jet energy against the galaxy and larger-scale dark matter halo binding energy and against the total thermal energy of the host gas within the halos. I show that although radio AGN, be it compact or complex, do not have enough energy to fully disrupt the global gas distribution of massive galaxies, they provide a significant source of heat that can impact the local thermodynamical balance in even the cores of the most massive clusters (groups of galaxies) in our local Universe.

Finally, I probe the details of the accretion process in an as-of-yet unexplored parameter space of low-mass galaxies at low redshift. This has been made possible by the recent advancements in deep, multi-wavelength all-sky surveys. In particular, this Thesis makes use of the deepest four-pass eROSITA All Sky Survey (eRASS:4), in combination with optical imaging from the 10th Data Release of the DESI Legacy Imaging Survey (LS10), to investigate the incidence of X-ray AGN as a function of specific black hole radiative power, $\lambda_{\rm Edd}$. This is a complementary mass-scaled power indicator to $\lambda_{\rm Jet}$. Starting from an optically selected parent sample from LS10 of over 5 million galaxies across the mass scale, I find 874 X-ray AGN in low-mass galaxies ($\log M_*/M_\odot \leq 10$), more than 600 of them newly discovered. Thanks to a Bayesian framework that makes use of the X-ray information from all parent sample galaxies, I place the tightest constraints on the specific accretion rate distributions (at least at the high accretion rate end), to date, revealing second-order mass-dependent properties. Then, I derive the cumulative AGN fraction as a function of stellar mass by integrating this distribution. Interestingly, I find a peak in this cumulative AGN fraction around $\log M_*/M_\odot \sim 10 - 10.5$, for moderately- and highly-accreting X-ray AGN, potentially highlighting a decrease in the efficiency of AGN fuelling at both very low and very high masses. Additionally, the peak aligns relatively well with the location where the star-formation efficiency as a function of halo mass is maximal, potentially pointing to a common physical condition that is responsible for AGN accretion and star-formation. Future detailed X-ray spectral-timing analysis on this new sample of X-ray AGN could reveal unique insights about the details of accretion and feedback in the low-mass regime.

Publications

During the writing of this Thesis, the following first-author, second-author and co-author refereed publications were produced (in reverse chronological order):

First-author publications

1. **The incidence of eROSITA X-ray AGN in the local Universe: from dwarf to massive galaxies**
   To be submitted to *Astronomy & Astrophysics (A&A)* in 2026.
   Forms the basis of Chapter 5 of this Thesis.
   Igo, Z., Merloni, A., Georgakakis, A., Buchner, J., Arcodia, R., Salvato, M., Nandra, K., Aird, J., Trakhtenbrot, B., Comparat, J., Lamer, G., Laloux, B., Kluge, M., Krumpe, M., Kyritsis, E., Roster, W., Bulbul, E., Balzer, F., Dwelly, T., Ni, Q., Bauer, F., Assef, R., Brandt, W. N., Gelfand, J., Anderson, S. F., Morrison S.

2. **The global energetics of radio AGN kinetic feedback in the local Universe**
   (2025)
   Forms the basis of Chapter 4 of this Thesis.
   Igo, Z. & Merloni, A.
   A&A, 697, A196. doi: 10.1051/0004-6361/202452888

3. **The LOFAR - eFEDS survey: The incidence of radio and X-ray AGN and the disk-jet connection**
   (2024)
   Forms the basis of Chapters 2 and 3 of this Thesis.
   Igo, Z., Merloni, A., Hoang, D., Buchner, J., Liu, T., Salvato, M., Arcodia, R., Bellstedt, S., Brüggen, M., Croston, J. H., de Gasperin, F., Georgakakis, A., Hardcastle, M. J., Nandra, K., Ni, Q., Pasini, T., Shimwell, T., & Wolf, J.
   A&A, 686, A43. doi: 10.1051/0004-6361/202349069

Second-author publications

4. **The first X-ray look at SMSS J114447.77-430859.3: the most luminous quasar in the last 9 Gyr**
(2023)
Forms part of Chapter 1 of this Thesis.
Kammoun, E. S., **Igo, Z.**, Miller, J. M., Fabian, A. C., Reynolds, M. T., Merloni, A., Barret, D., Nardini, E., Petrucci, P. O., Piconcelli, E., Barnier, S., Buchner, J., Dwelly, T., Grotova, I., Krumpe, M., Liu, T., Nandra, K., Rau, A., Salvato, M., Urrutia, T., & Wolf, J.
A&A, 522, 5217-5237. doi: 10.1093/mnras/stad952

Co-authored publications

5. **Shooting for the stars: Jet-mode feedback and AGN jet deceleration from stellar mass-loading**
(2025)
O'Shea, T. M., Heinz, S., Soares-Furtado, M., **Igo, Z.**, Merloni, A.
Accepted for publication in *The Astrophysical Journal (ApJ)*
doi: 10.48550/arXiv.2510.26881

6. **Retrieving the hot circumgalactic medium physics from the X-ray radial profile from eROSITA with an IlustrisTNG-based forward model**
(2025)
Shreeram, S., Comparat, J., Merloni, A. et al. (including **Igo, Z.**)
A&A, 703, A137. doi: 10.1051/0004-6361/202554508

7. **X-ray selected broad absorption line quasars in SDSS-V: BALs and non-BALs span the same range of X-ray properties**
(2025)
Hiremath, P, Rankine, A. L., Aird, J. et al. (including **Igo, Z.**)
Monthly Notices of the Royal Astronomical Society (MNRAS), 542, 3.
doi: 10.1093/mnras/staf1352

8. **The average soft X-ray spectra of eROSITA active galactic nuclei**
(2025)
Chen, S.-J., Buchner, J., Liu, T. et al. (including **Igo, Z.**)
A&A, 701, A144. doi: 10.1051/0004-6361/202554737

9. **Searching for supermassive black hole binaries within SRG/eROSITA-De: I. Properties of the X-ray selected candidates**
(2025)
Tubín-Arenas, D., Krumpe, M., Homan, D. et al. (including **Igo, Z.**)
A&A, 698, A192. doi: 10.1051/0004-6361/202554749

10. **Genuine Retrieval of the AGN Host Stellar Population (GRAHSP)**
(2024)
Buchner, J., Starck, H., Salvato, M. et al. (including **Igo, Z.**)
A&A, 692, A161. doi: 10.1051/0004-6361/202449372

11. **The eROSITA Final Equatorial Depth Survey (eFEDS): the hard X-ray selected sample**
(2024)
Nandra, K., Waddell, S. G. H., Liu, T. et al. (including **Igo, Z.**)
A&A, 693, A212. doi: 10.1051/0004-6361/202449416

12. **The SRG/eROSITA all-sky survey: Hard X-ray selected Active Galactic Nuclei**
(2024)
Waddell, S. G. H., Nandra, K., Buchner, J. et al. (including **Igo, Z.**)
Resubmitted to A&A after first positive referee report
doi: 10.48550/arXiv.2401.17300

13. **Winds of change: The nuclear and galaxy-scale outflows and the X-ray variability of 2MASS 0918+2117**
(2024)
Baldini, P., Lanzuisi, G., Brusa, M., et al. (including **Igo, Z.**)
A&A, 686, A217. doi: 10.1051/0004-6361/202349071

14. **The SRG/eROSITA all-sky survey. First X-ray catalogues and data release of the western Galactic hemisphere**
(2024)
Merloni, A., Lamer, G., Liu, T., et al. (including **Igo, Z.**)
A&A, 682, A34. doi: 10.1051/0004-6361/202347165

15. **The X-ray invisible Universe. A look into the haloes undetected by eROSITA**
(2024)
Popesso, P., Biviano, A., Bulbul, E., et al. (including **Igo, Z.**)
A&A, 527, 895-910. doi: 10.1093/mnras/stad3253

16. **O Corona, where art thou? eROSITA's view of UV-optical-IR variability-selected massive black holes in low-mass galaxies**
(2024)
Arcodia, R., Merloni, A., Comparat, J., et al. (including **Igo, Z.**)
A&A, 681, A97. doi: 10.1051/0004-6361/202347531

17. **Stellar-wind variability in Cygnus X-1 from high-resolution excess variance spectroscopy with Chandra**
(2023)

Härer, L. K., Parker, M. L., El Mellah, I., R., et al. (including **Igo, Z.**)
A&A, 680, A72. doi: 10.1051/0004-6361/202346669

18. **A new discovery space opened by eROSITA. Ionised AGN outflows from X-ray selected samples**
(2023)
Musiimenta, B., Brusa, M., Liu, T. et al. (including **Igo, Z.**)
A&A, 679, A84. doi: 10.1051/0004-6361/202245555

Chapter 1

Introduction

"[The black hole] teaches us that space can be crumpled like a piece of paper into an infinitesimal dot, that time can be extinguished like a blown-out flame, and that the laws of physics that we regard as 'sacred,' as immutable, are anything but."
– John Wheeler (1911–2008)

Black holes were first theorised around 250 years ago, first described mathematically over 100 years ago, first discovered around 60 years ago and first imaged 6 years ago; yet many years of research still lie ahead to shed light on their darkness. This Thesis contributes to this body of research on the mechanisms that make black holes shine very brightly (accretion) and not appear black at all, contrary to popular belief. It also delves into the ways in which black holes impact the galaxies they live in (feedback), despite being up to a thousand times less massive and a trillion times smaller in size. Overall, it aims to describe observations of black holes at different wavelengths, ranging from radio to optical and X-rays —similar to looking through differently tinted sunglasses— in order to piece together a picture of black hole accretion and feedback for black holes living in the smallest and largest galaxies in the local Universe.

1.1 Black holes throughout history

The term ‘black hole’ was originally coined in 1967 by John Wheeler, who was the PhD advisor of well-known physicist Richard Feynman. However, the early theories in this field date back to the 18th century to the predictions of ‘dark stars’, stars which had such a strong surface gravity that not even light could escape (Michell 1784; de Laplace 1796). The true breakthrough came with the publishing of the series of papers by Albert Einstein on the theory of General Relativity (GR; Einstein 1916). This theory no longer thought of gravity as an ‘invisible tether’ between objects, like Isaac Newton described centuries ago, but as the invisible curvature of space-time. As Wheeler elegantly put it: “Space-time tells matter how to move; matter tells space-time how to curve.” It was then Karl Schwarzschild who, from the trenches along the Russian front-lines during World War I, derived the solution to the field equations in Einstein’s theory of GR for how a spherically symmetric object with mass M curves space-time (Schwarzschild 1916):

$$ds^2 = -\left(1 - \frac{2GM}{Rc^2}\right) c^2 dt^2 + \left(1 - \frac{2GM}{Rc^2}\right)^{-1} dR^2 + R^2 d\theta^2 + R^2 \sin^2\theta \, d\phi^2, \tag{1.1}$$

where ds^2 is the invariant space-time interval measuring the separation between two nearby events in space-time that is defined in four-dimensions, three from spherical coordinates (R, ϕ, θ) and one from time t; G is the Gravitational constant and c is the speed of light. This Schwarzschild metric was an important discovery because it not only highlighted the true singularity present at $R = 0$, where the curvature of space-time itself diverges, but also a coordinate singularity at

$$R = R_s = \frac{2GM}{c^2}, \tag{1.2}$$

where the time and radial terms vanish and diverge, respectively. The latter is named the Schwarzschild radius, R_s, and defines the radius of the spherical boundary, or ‘Event Horizon’, beyond which not even light can escape and all paths lead inevitably toward $R = 0$.

This was such a strange finding that the physics community at the time thought that there must be some natural phenomenon that prevents these solutions in the real Universe. For example, it was known that stars consume some form of fuel to shine so brightly (although the details of nuclear fusion in stellar cores were only discovered in the late 1930s) and at some point this fuel could run out, meaning that there would be no more radiation pressure to prevent gravitational collapse. However, with the dawn of the quantum physics era in the 1920s, Wolfgang Pauli discovered the ‘Pauli Exclusion Principle’ (later generalised by Jordan & Wigner 1928), which was a theorem stating that no two fermions (particles with half-integer spin, such as electrons) can occupy the same quantum state simultaneously. This meant that in the cores of dying stars, electrons started moving around faster, thereby providing an additional force, in the form of electron degeneracy pressure, to counteract the force of gravity. In fact, electron degeneracy pressure is the main support against gravity for white dwarf stars, which are compact stellar remnants formed from progenitors with a mass less than around 8 times the mass of the Sun ($\lesssim 8M_\odot$).

However, Subrahmanyan Chandrasekhar realised that electrons are also subject to Einstein's laws of special relativity, stating that nothing can travel faster than the speed of light, which led him to derive the maximum mass threshold for a white dwarf to be around $1.44M_\odot$. Beyond this limit, electrons are forced into protons, creating neutrons (also fermions), and provide one last source of resistance to prevent the gravitational collapse through neutron degeneracy pressure. Nevertheless, for stellar progenitors above a mass of $\sim 25M_\odot$ or remnant stellar cores of $> 2.5M_\odot$, nothing can prevent a collapse into a stellar-mass black hole (Oppenheimer & Snyder 1939; Osterbrock 1989).

With this strong theoretical basis, the first observational X-ray astronomers began looking at the sky to find these exotic objects using Geiger counters launched in sounding rockets that spent just a few minutes above the Earth's atmosphere to observe celestial X-ray emitters (as the atmosphere absorbs X-rays). The first detection came from the region of the sky near the constellation Cygnus, and thus it was named Cygnus-X-1 (Cyg X-1; Bowyer et al. 1965), but it was only after the launch of the first X-ray satellite *Uhuru* in 1970, that its position could be better constrained. Webster & Murdin (1972) found the location of Cyg X-1 to align with a blue supergiant star that appeared to be orbiting with a period of 5.6 days around an unknown companion. Knowing that such stars have a surface brightness temperature of $T \sim 30 - 40,000\ K$ and so cannot be responsible for the X-ray emission (which requires $T \sim 1 \times 10^6\ K$) and using Kepler's law of orbital motion, Webster & Murdin derive the mass of the unknown companion to be between $2.5-6M_\odot$ (later revised to be $\sim 15M_\odot$ by Orosz et al. 2011). This pointed to the companion being a black hole, and it became the first high mass X-ray binary system (HMXRB) ever discovered. Since then, only a few hundred more systems, also with low-mass stellar companions (LMXRBs), have been found, but owing to their relatively nearby nature, have provided unparalleled datasets to study the properties of their compact object companions, be it neutron stars or black holes (e.g. Grimm et al. 2003; Fabbiano 2006; Antoniou & Zezas 2016).

Parallel to the search for the end-state of massive stars as stellar-mass black holes, in the beginning of the 1960s, Maarten Schmidt began observing the sky using radar (at radio wavelengths), a technique much advanced during World War I and II[1]. He was surprised to see many radio sources in the sky because stars typically do not emit at this wavelength, given that their black body spectrum peaks in the optical part of the electromagnetic (EM) spectrum. Unfortunately, given the long wavelength of the radio waves, it was difficult to pinpoint the location of the radio emission of a particularly bright object, 3C 273, up until an occultation event by the moon revealed a more precise search region. Within this region, the host was identified and an optical spectrum of the object was taken, only to reveal a striking discovery: the Balmer lines (originating from atomic transitions from higher excited states, $n \geq 3$, of Hydrogen to the first excited state, $n = 2$) were shifted to longer wavelengths by a factor of 15.8% compared to lab measurements (Schmidt 1963). Given previous discoveries of an expanding Universe by Hubble (1929), a larger shift between the observed and emitted wavelength ($\lambda_{\mathrm{obs,em}}$) meant a larger recessional velocity, v, as per the

---

[1] In particular, it was Karl Janksy who first discovered diffuse radio emission, peaking in the direction of the Galactic Center, using a low-frequency radio antenna (Jansky 1933).

relativistic Doppler effect (Eq. 1.3),

$$\frac{\lambda_{\rm obs}}{\lambda_{\rm em}} = \sqrt{\frac{1+v/c}{1-v/c}} = 1+z, \tag{1.3}$$

where z is the redshift ($z = (\lambda_{\rm obs} - \lambda_{\rm em})/\lambda_{\rm em}$), effectively a measure of the distance to an astronomical object. For small $v << c$, this equation reduces to $z \approx v/c$, meaning the redshift of 3C 273 was around 0.158, or that it was moving away from us at a velocity equal to 15.8% of the speed of light. This was orders of magnitude higher than previously discovered stars, moving at $v << 1\%\,c$ in our galactic neighbourhood. 3C 273 is also incredibly bright, $\sim 10^{14}$ times brighter than our Sun, a typical star, and ~ 5000 times brighter than our Milky Way, a typical spiral galaxy (Greenstein & Schmidt 1964; Courvoisier 1998). As its origin at the time was unknown, but it was very bright and point-like, just like stars in the night sky, it was named a quasi-stellar object (QSO), or quasar. This extreme brightness was then hypothesised to be associated with accretion of interstellar gas onto a massive compact object (Salpeter 1964; Lynden-Bell 1969), such as a super-massive black hole (SMBH), with mass upwards of $M_{\rm BH} \sim 10^6 M_\odot$. However, it was only much later that the mass of the black hole could be estimated through dynamical methods mapping the motion of the gas surrounding the black hole to be greater than $10^9 M_\odot$ (Courvoisier 1998; Li et al. 2022). This confirmed the nature of 3C 273 as the first extragalactic radio galaxy hosting an accreting super-massive black hole (active galactic nucleus, AGN[2]) at its centre. Upon further investigation, 3C 273 also revealed extended radio emission stretching over ~ 60 kpc, alluding to relativistic jets emanating from the central black hole, as shown in Figure 1.1. This was in contrast to other point-like radio sources also found with early surveys (e.g. Edge et al. 1959; Bennett 1962). Sandage (1965) later broadened the definition of QSOs to include not only quasi-stellar *radio* objects, such as 3C 273, after finding a sample of extragalactic objects that were optically luminous, but radio-faint.

It was still very rare to find AGN at the time, so the community did not give them much importance in the context of their potential impact on their surrounding environments. However, this view was revolutionized by the launch of the Hubble Space Telescope (HST), an optical telescope no longer limited by Earth's turbulent atmosphere, providing astronomers with the sharpest view of the cosmos to date. Importantly, HST could resolve the inner cores of galaxies with its sub-arcsecond ($\sim 0.05''$) angular resolution, providing a clear view into the centres of galaxies, near and far. Such studies, together with several key ground-based observations, revealed the presence of very fast moving stars and gas on stable bound orbits, that could only be kept there if some massive compact object, such as a SMBH, was providing a deep enough potential well (e.g. Kormendy & Richstone 1995; Ferrarese et al. 1996; Genzel et al. 1997; Ghez et al. 2000, 2008; Genzel et al. 2010, a discovery already primed from the early studies of Carl K. Seyfert finding bright nuclei with broad emission lines in what are now commonly known as 'Seyfert galaxies'; Seyfert 1943). Soon after, key relations were found between the stellar mass of the galaxy and the

---

[2] By common nomenclature, QSOs are the most luminous type of AGN, so all QSOs are AGN but not vice versa.

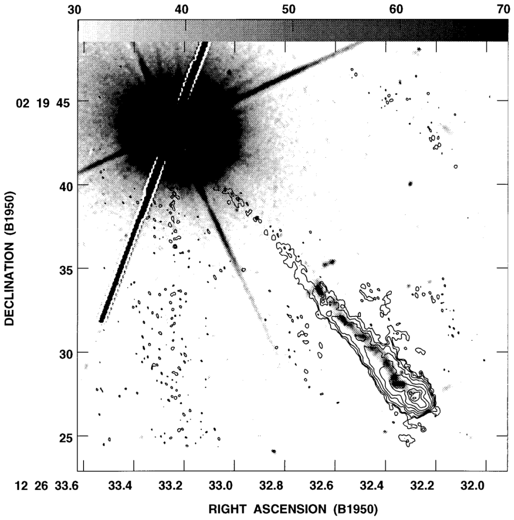


Figure 1.1: The bright QSO 3C 273, as seen by the Hubble Space Telescope at optical wavelengths (grey scale) and by the MERLIN array at radio wavelengths (contours). The radio jet is also visible at a position angle of $\sim$ 220°. Figure taken from Bahcall et al. (1995).

mass of the central SMBH, and the velocity dispersion of stars around the core, or bulge, of the galaxy and the SMBH mass, indicating some fundamental relationship between the SMBH and its host galaxy (e.g. Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Kormendy & Ho 2013).

Nowadays, many millions of AGN are known, thanks to multi-wavelength surveys over all wavelengths of the EM spectrum from radio, optical/ultra-violet (UV), infrared (IR), X-ray and gamma ray surveys. Additionally, even though AGN constitute only around $\sim 1-10\%$ of all SMBHs, with the remaining population being dormant (i.e. not accreting at a detectable level), the vast amount of energy they release has been hypothesised to significantly impact the evolution of galaxies or larger-scale environments such as the gas in between galaxies and beyond. A simple thought experiment taking an average SMBH of $M_{\rm BH} = 10^8 M_{\odot}$, converting just 10% of its rest-mass into energy, via Einstein's mass-energy equivalence equation, $E = mc^2$, provides $\sim 2 \times 10^{55} J$ of energy. This is approximately equivalent to the energy radiated by 100 billion stars over 13.8 billion years, the age of the Universe.

After decades of indirectly inferring the existence of black holes, the Event Horizon Telescope (EHT) combined signals from several radio dishes scattered across the Earth, through a process called radio interferometry, to produce the first ever images of a SMBH's shadow (Event Horizon Telescope Collaboration et al. 2019, 2022). Figure 1.2 shows these two SMBHs, one in the heart of the galaxy *Messier* 87 (M87) and the other in our own Milky way (Sagittarius A*, Sgr A*). Recently, the Laser Interferometer Gravitational-Wave Observatory (LIGO) also opened up a new, non-electromagnetic window to detect

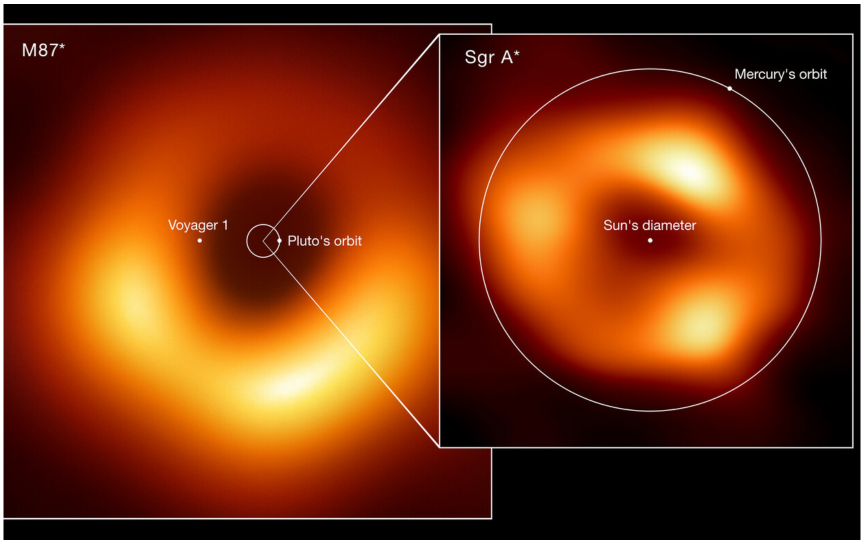


Figure 1.2: First images of the two SMBHs, Virgo A* in M87 (left) and Sgr A* (right), taken by the Event Horizon Telescope (EHT) Collaboration. Orbits of Pluto and Mercury, as well as Voyager 1, the furthest man-made object (spacecraft) from Earth, are marked for scale. Image credit: EHT collaboration (Event Horizon Telescope Collaboration et al. 2019, 2022) with acknowledgment to Lia Medeiros (xkcd).

black holes through the gravitational waves, effectively ripples in the fabric of space-time, caused by the merger of black holes (or other compact objects, such as neutron stars; Abbott et al. 2016).

As can be seen from the above overview of black holes throughout history, a multi-wavelength approach has been vital to grasp the physics at play in these mysterious objects. It also evolved into a multi-scale endeavour, not only in looking for synergies between stellar-mass and super-massive black holes, but also looking at the impact of the latter type on larger scale environments, such as the host galaxy and beyond.

The subsequent parts of the Introduction elaborate on the concepts mentioned in this section, including: (i) accretion onto compact objects and AGN (Ch. 1.2); (ii) the physical structure of AGN and their multi-wavelength emission (Ch. 1.3); (iii) black holes across the mass scale (Ch. 1.4); (iv) detecting AGN with multi-wavelength surveys (X-ray, radio, optical; Ch. 1.5); (v) important discoveries with such surveys regarding the way AGN populate their host galaxies and the way they feed back energy to their surroundings (Chs. 1.6, 1.7, 1.8); (vi) black hole seeding and growth in the early Universe (Ch. 1.9); and finally (vii) an outline of this Thesis (Ch. 1.10).

1.2 Active galactic nuclei (AGN) and their accretion modes

AGN are the most luminous persistent emitters of EM radiation in the entire Universe and this is possible via the process of 'accretion'. Conservation of angular momentum prevents the direct accretion of infalling matter towards the black hole, forcing it to circularise and form an accretion disk. Then, in order for the matter to accrete onto the central black hole, it must first lose this angular momentum. It does so by passing it outward to material farther from the centre, usually via viscous transport in the disk due to its differential rotation. A cornerstone of accretion theory is the optically thick, geometrically thin disk model of Shakura & Sunyaev (1973), which provides the basis for much of our current understanding and explains many, though not all, of the observed properties of accretion disks. If we consider this 'thin disk' solution and the conservation of energy, but neglect, for the time being, the conservation of angular momentum, we can get an intuitive understanding of this accretion process (Done 2010).

Material accreting at a rate $\dot{M}$ and moving inward from radius R to $R - dR$ releases gravitational potential energy at a rate

$$\frac{dE}{dt} = L_{\text{pot}} = \frac{GM\dot{M}}{R^2} dR. \tag{1.4}$$

Given the virial theorem, only half of this energy can be radiated, meaning

$$dL_{\text{rad}} = \frac{GM\dot{M}}{2R^2} dR. \tag{1.5}$$

Since the disk is optically thick, the photons are trapped and thermalise locally, so the disk radiates approximately as a blackbody. The luminosity of each disk annulus can be expressed as: $dL = dA\, \sigma_{\text{SB}} T^4$, where σ_{SB} is the Stefan–Boltzmann constant and the radiating area of the annulus is $dA = 2 \times (2\pi R\, dR)$, with the factor of two accounting for both the top and bottom surfaces of the disk. Equating the two expressions yields the luminosity from the annulus:

$$dL_{\text{rad}} = \frac{GM\dot{M}}{2R^2}\, dR = 4\pi R\, dR\, \sigma_{\text{SB}} T^4 \tag{1.6}$$

or equivalently,

$$\sigma_{\text{SB}} T^4(R) = \frac{GM\dot{M}}{8\pi R^3}. \tag{1.7}$$

A more complete treatment that accounts for angular momentum transport and postulates that no torque is exerted at the inner boundary introduces only a correction factor of $3(1 - \sqrt{R_{\text{ISCO}}/R})$, where R_{ISCO} is the innermost stable circular orbit (ISCO), in this case, of the disk. Therefore, the disk can be described as a multi-colour (multi-temperature) blackbody spectrum with $T \propto R^{-3/4}$.

Additionally, the luminosity of the disk can be parameterised in terms of the Eddington Luminosity, $L_{\rm Edd}$:

$$L_{\rm Edd} = \frac{4\pi G M m_p c}{\sigma_{\rm T}} \simeq 1.26 \times 10^{38} \left(\frac{M}{M_\odot}\right) \text{ erg s}^{-1}. \tag{1.8}$$

This sets a theoretical limit, the 'Eddington limit', on the accretion onto a compact object. It is defined by the balance between the outward radiation pressure (acting via Thomson scattering with cross-section, $\sigma_{\rm T}$, on electrons in a fully ionized, spherically symmetric hydrogen plasma) and the inward gravitational force exerted by the object of mass M on the protons of mass m_p. In practice, this limit can be exceeded as accretion flows are not such idealised, spherically-symmetric plasmas. However, it is still a very useful metric that will be used throughout this Thesis, notably when defining the 'Eddington ratio', $\lambda_{\rm Edd}$, which is the ratio of the bolometric luminosity $L_{\rm bol}$ (luminosity emitted over all wavelengths) to $L_{\rm Edd}$.

$$\lambda_{\rm Edd} = \frac{L_{\rm bol}}{L_{\rm Edd}}. \tag{1.9}$$

In fact, for a disk luminosity that is at a fixed fraction (f) of Eddington $L_{disk} = f L_{\rm Edd} \propto M$, we can derive the maximum temperature of the disk ($T_{\rm max}$; occurring at $R \sim 1.36 R_{\rm ISCO}$), to scale as $T_{\rm max} \propto M^{-1/4}$. This means that for stellar-mass black holes, such as Cyg X-1 (recall Ch. 1.1), with black hole mass $M_{BH} \sim 10 M_\odot$, the temperature of the disk should peak in the soft X-ray part of the EM spectrum, meanwhile for SMBHs, such as 3C 273, with $M_{BH} \gtrsim 10^6 M_\odot$, it should peak in the ultra-violet (UV) regime ('big blue bump', see Ch. 1.3). Observationally, this is found to be the case (e.g. Shields 1978; Elvis et al. 1994; Done et al. 2007), which then begs the question: does the fact that we observe emission from AGN and XRB peaking in the expected energy range also imply that the main mechanisms of accretion are mass-invariant? This is a core research area that Chapters 3 and 5 will further explore.

It can be seen above that $R_{\rm ISCO}$ is a key factor in determining the properties of accretion. In fact, $R_{\rm ISCO}$ depends on the black hole angular momentum (J), which is one of the three fundamental parameters of a black hole (according to the 'no hair theorem'; Ruffini & Wheeler 1971), along with the black hole mass and electric charge. For astrophysical black holes, the charge is expected to be zero, so effectively the important parameters to uniquely describe a black hole are its mass and angular momentum[3], usually characterised using the dimensionless spin parameter,

$$|a| = cJ/GM^2. \tag{1.10}$$

The black hole spin parameter can be in the same (prograde) or opposite (retrograde) direction as the accretion disk, and can mathematically range from zero (non-spinning, often called 'Schwarzschild' black holes) to unity (maximally spinning, often called 'Kerr' black holes). For a non-spinning black hole, $R_{\rm ISCO} = 3R_S = 6GM/c^2 = 6R_g$, where R_g is the gravitationally radius. However, the ISCO can shrink down to $0.5R_s = R_g$

---

[3] It seems easy enough, but these are two notoriously difficult parameters to constrain observationally!

for a maximally-spinning prograde black hole, and extend to $4.5R_S = 9R_g$ for retrograde spin (see review by e.g. Reynolds 2021). The location of the ISCO determines the inner radius at which a stable disk can exist and thereby how much gravitational potential energy can be converted into radiation: the closer in the ISCO, the deeper matter can sink adiabatically into the gravitational potential well and the greater the radiative efficiency, η. The accretion luminosity is then given by

$$L = \eta \dot{M} c^2. \tag{1.11}$$

For maximally spinning black holes $\eta \sim 42\%$, whereas for non-spinning black holes $\eta \sim 6\%$[4].

So far we have only considered Shakura-Sunyaev disks, which are typically present in black holes accreting at moderate rates, with $\lambda_{\rm Edd} \gtrsim 1-10\%$. In the AGN nomenclature, we call these 'radiatively efficient' sources, as a substantial fraction of the energy is being released radiatively at the location where it is produced, or in the form of wide-angle winds, as shown in Figure 1.3 (e.g. Fabian 2012; Harrison & Ramos Almeida 2024, and references therein). They are predominantly detected in the optical/UV and at X-ray wavelengths (see Ch. 1.3 for further details of the geometry and associated multi-wavelength emission of AGN). A small fraction of radiatively efficient, luminous accretion disks have also been associated with radio jets, which is the case for 3C 273, for example.

On the other hand, at low accretion rates, $\lambda_{\rm Edd} \ll 1\%$, the disk cannot efficiently radiate energy away and may develop an advection-dominated inner accretion flow (ADAF), becoming a 'puffed up', optically thin, geometrically thick, hot disk as shown on Figure 1.3 (e.g. Narayan & Yi 1994, 1995; Yuan & Narayan 2014). Due to the low density in the ADAF flow, the ions, which are heated by viscous stresses, cannot transfer this energy via Coulomb interactions to the electrons (responsible for radiating away the energy) and this results in a 'two-temperature' accretion flow. In this classical ADAF model there are no significant outflows; the excess energy is simply stored in the ions' thermal energy and advected into the black hole. However, alternative solutions propose kinetic energy as a sink of energy, for example in the form of a relativistic particle jet (Begelman et al. 1984; Blandford & Begelman 1999; Blandford et al. 2019), which is the case for some low luminosity AGN (LLAGN; e.g. Merloni & Heinz 2007). In general, these accretion modes are characterised by being 'radiatively inefficient', meaning that the energy generated by accretion is largely trapped within the inflowing gas and either carried into the black hole or released mechanically through jets and outflows.

Lastly, for very high accretion rates ($\lambda_{\rm Edd} \gtrsim 100\%$), the Shakura–Sunyaev disk model is expected to break down. Although the accretion flow remains optically thick, it should become geometrically thicker due to the dominance of radiation pressure in the inner regions and the necessity to accommodate an increased mass supply to the AGN. At the same time, as shown by Shapiro et al. (1976), radiation pressure–dominated disks can become thermally and viscously unstable, leading to strong variability or structural transitions in the flow. These instabilities are alleviated in the so-called 'slim disk' regime

---

[4] This vastly exceeds the measly 0.7% efficiency of proton-proton chain that powers our bright Sun.

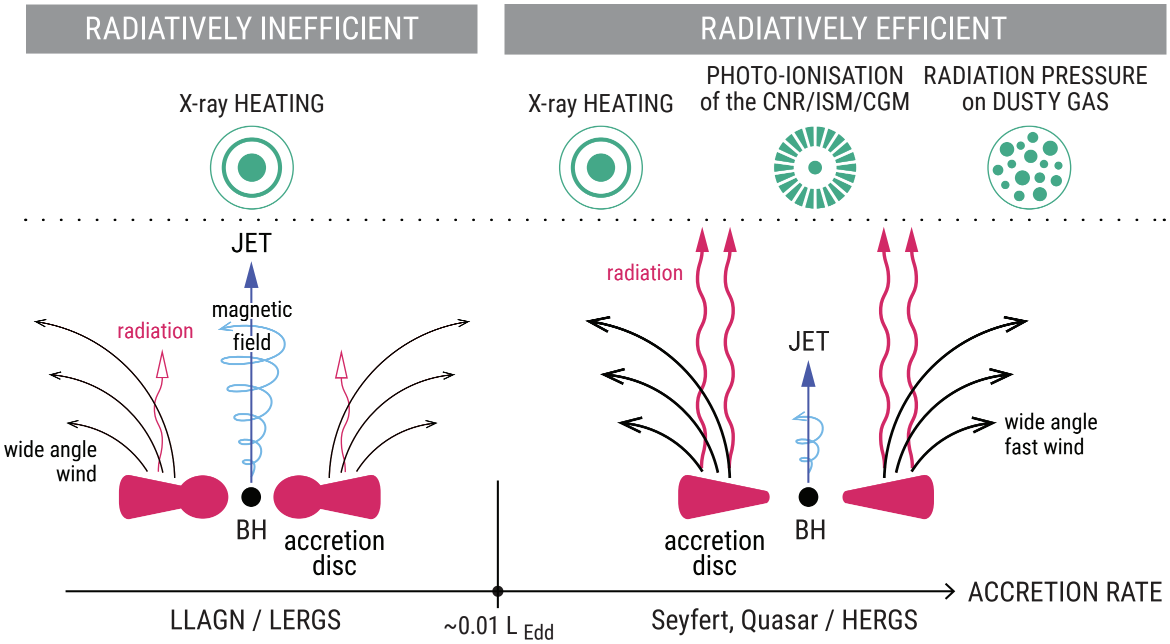


Figure 1.3: Diagram showing the geometry and main emission processes from AGN in the radiatively inefficient (left) and radiatively efficient (right) modes. The former includes sources like low luminosity AGN (LLAGN) and low-excitation radio galaxies (LERGS), whereas the latter includes Seyferts, Quasars and high-excitation radio galaxies (HERGs). However, observationally it is not always trivial to determine the accretion mode; for example, both modes can have a radio jet. Figure adapted from Harrison & Ramos Almeida (2024).

(Abramowicz et al. 1988), where the photon diffusion time exceeds the accretion time, causing a significant fraction of the dissipated energy to be advected inward with the flow rather than radiated locally. This results in a radiative efficiency lower than that of thin disks, even though the mass accretion rate is much higher.

Observationally, it is difficult to ascertain the accretion mode of an AGN, as we rarely directly measure the physical accretion rate $\dot{M}$ and there are several common properties among the modes, such as winds and jets, as is clear from Figure 1.3 (see also e.g. Sądowski et al. 2014; Massonneau et al. 2023; Lowell et al. 2024). However, it is an essential concept for this Thesis, as we attempt to observationally disentangle these accretion modes using large statistical samples in order to investigate the coupling between the disk and the jet, and to place their energetic output in a broader context by examining their impact on the large-scale host-galaxy and host-halo environments.

1.3 Physical structure and multi-wavelength signatures of AGN

Figure 1.4 shows the various physical structures of an AGN, each at different scales and each responsible for a part of the emission we observe (e.g. Padovani et al. 2017). These include the accretion disk, broad and narrow line region, torus, corona and jets. The emission from these components is visible over the entire EM spectrum, from the radio to X-ray and gamma ray wavelengths, as depicted by the the spectral energy distribution (SED) of a typical unobscured AGN (black curves) in Figure 1.5 (e.g. Alexander & Hickox 2012; Heckman & Best 2014; Hickox & Alexander 2018; Hardcastle & Croston 2020a). It is important to note, however, that the AGN SED can drastically differ for different types of AGN, or even the same AGN at different times (e.g. Hickox et al. 2014) and AGN should be thought of as 'events' in the lifetime of a galaxy rather than 'objects' (e.g. Harrison & Ramos Almeida 2024). Unlike stars, AGN spectra cannot be represented by a single temperature black body. Instead, their emission in different wavelengths can often be described by different power-laws, where the flux in a given wavelength band is given by

$$f_\nu \propto \nu^{-\alpha}, \tag{1.12}$$

with α being a constant (Frank et al. 2002; Beckmann & Shrader 2012). Observationally, we typically measure the AGN emission in one waveband at a time and then use a bolometric correction factor, $k_{\rm bol}$, to convert to $L_{\rm bol}$. Figure 1.5 also shows an SED of a typical starburst galaxy, Messier 82 (grey curve), that peaks strongly at IR wavelengths.

The primary emission of an AGN originates from the accretion disk, which can have various geometries depending on the accretion mode. In the case of optically think, radiatively efficient accretion, the disk emits predominantly in the UV, with gas temperature of $T \sim 10^4 - 10^5$ K. This manifests itself as a 'big blue bump', as a result of the multi-colour blackbody spectrum from the accretion disk annuli, as shown by the blue dot-dashed curve in Figure 1.5, typically with a power-law rest-frame optical/UV continuum ($\sim 10 - 700$ nm) of the form $f_\nu \propto \nu^{1/3}$ (Beckmann & Shrader 2012). This primary emission is then reprocessed by the various components of an AGN. For example, at a distance of $\sim 0.01 - 1$ pc from the SMBH, dense broad line region (BLR) gas clouds with electron densities $n_e \sim 10^9$ cm^{-3} (Osterbrock 1989) get photoionised by the disk's extreme UV (EUV) radiation. Subsequent recombination produces emission lines that, due to the high velocities of the BLR gas, appear Doppler-broadened with full widths at half maximum (FWHM) of $\gtrsim 1000$ km s^{-1}. Common broad lines observed in AGN spectra include Hydrogen α (Hα; 6563Å), Hydrogen β (Hβ; 4861Å), Helium II (He II; 4686 Å), the Magnesium II doublet (Mg II; 2796Å, 2803Å), and the Carbon IV doublet (C IV; 1548Å, 1551Å). Similarly, at a distance of hundreds of parsec, the narrow line region (NLR), with lower electron densities of $n_e \sim 10^{3-6}$ cm^{-3} and moving at lower velocities, imprints narrow emission lines on the power-law continuum emission. Common narrow lines include the forbidden lines of the Oxygen III ([O III]; 4959Å, 5007Å) and the Nitrogen II ([N II]; 6548Å, 6583Å) doublets with FWHM $\lesssim 1000$ km s^{-1}. Depending on the redshift of the

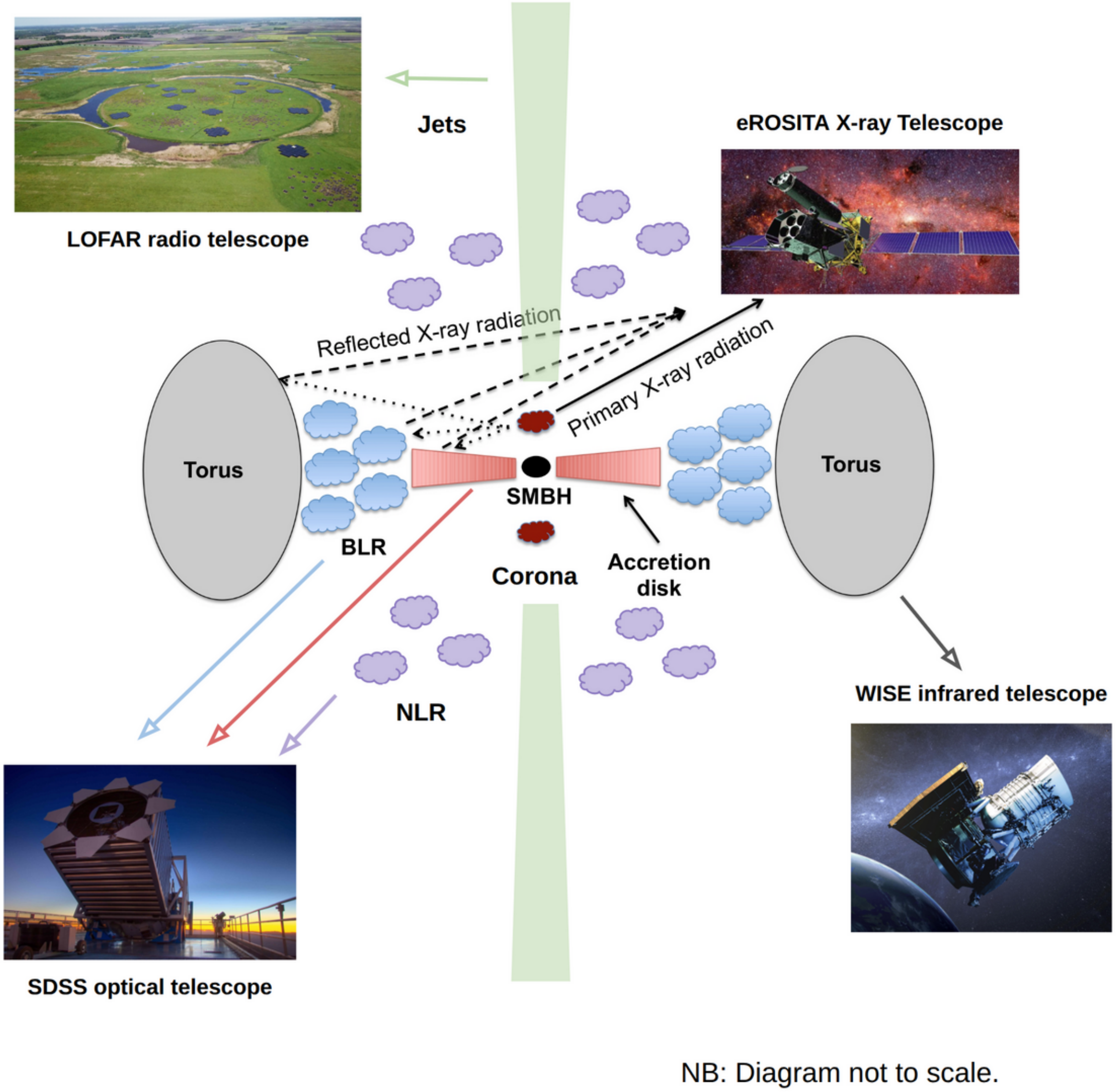


Figure 1.4: Commonly accepted schematic of an AGN, featuring the central black hole, accretion disk, corona, broad line region (BLR), narrow line region (NLR), torus, and radio jets. Arrows show the origin of different wavelength emission from each of these components and an example telescope used to detect them (see Ch. 1.5). Figure adapted from images credited to C. Ricci, Roscosmos Space Corporation, ASTRON, NASA/JPL-Caltech, SDSS/David Kirkby.

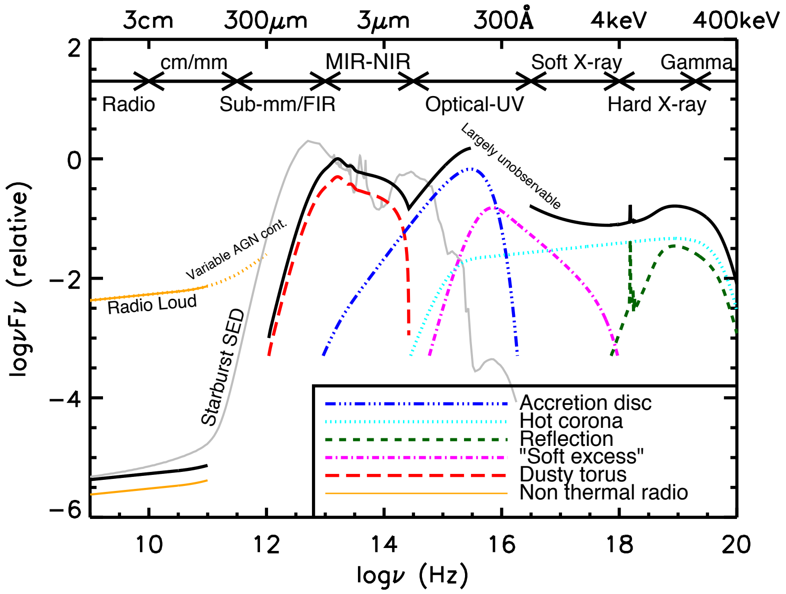


Figure 1.5: SED of an unobscured AGN (black curves) with emission components from the radio jet (yellow), dusty torus (red), accretion disk (blue), along with primary X-ray coronal emission (cyan) and reprocessed X-ray emission in the form of a soft-excess (pink) and reflection components (green). A typical starburst galaxy SED is also shown by the grey curve. Figure adapted from Harrison (2014); Hickox & Alexander (2018).

source, different optical/UV features are visible in the spectrum. These spectra can be obtained by various optical/UV ground- and space-based surveys, such as the *Sloan Digital Sky Survey* (SDSS; see Ch. 1.5 for further information on the different telescopes used to detect AGN emission at different wavelengths).

Outside of the BLR, at around $1-10$ pc, lies the torus, which is an optically thick, geometrically thick dusty absorber (e.g. Nenkova et al. 2008; Buchner et al. 2019). Unlike what is shown in Figure 1.4, the torus is not a smooth, static, well-defined toroidal structure, but rather a clumpy, irregular region, with an inner boundary set by the dust sublimation temperature, that may potentially be outflowing (e.g. Ramos Almeida & Ricci 2017; Hönig 2019; Ananna et al. 2022b). The primary emission from the AGN heats the micrometer-sized dust grains in the torus, which absorb and isotropically re-radiate the emission at IR wavelengths (700 nm $\lesssim \lambda \lesssim$ 1 mm) with an overall intensity which depends on the torus covering factor (e.g. Hickox & Alexander 2018). We can observe this emission from space-based telescopes such as the *Wide-field Infrared Survey Explorer* (WISE). It is important to note that star formation in the galaxy can also heat surrounding dust and contribute to the IR emission of the source, so it must be properly accounted for (see Chapter 5.2.1). Since dust is the main absorber of optical/UV radiation, it is clear from Figure 1.4 that if an AGN is viewed edge-on, the torus blocks the view to the BLR and only narrow emission lines can be observed in the optical spectrum. We refer to these sources as Seyfert 2 galaxies, or generically as Type 2 AGN. At a more face-on view, both the BLR and NLR are visible and AGN are classed as Seyfert 1 galaxies (Type 1 AGN; e.g. Seyfert 1943; Antonucci 1993; Urry & Padovani 1995).

At scales of $R \lesssim 10R_g$, lies the optically thin, hot corona, consisting of electrons forming a plasma with $k_bT \sim 100$ keV (e.g. Haardt & Maraschi 1993). These electrons up-scatter the cooler UV accretion disk 'seed' photons, via inverse Compton scattering (see Ch. 1.3.2), to soft ($E \lesssim 2$ keV) and hard ($E \gtrsim 2$ keV) X-ray energies (see cyan dotted line in Figure 1.5). This results in power-law emission of the form $f_{intr} \propto E^{-\Gamma}$, where Γ is the photon index and has typical values of $\sim 1.8 - 2$ for AGN (e.g. Nandra & Pounds 1994; Beckmann & Shrader 2012; Brandt & Alexander 2015). Photon indices $\gtrsim 2$ indicate a 'soft spectrum', with the bulk of the emission being emitted in the soft band, whereas $\Gamma \lesssim 2$ indicate a 'hard spectrum'. The exact geometry of the corona is still debated, although recent work using the *Imaging X-ray Polarimetry Explorer* (IXPE) favours a geometry elongated in the plane of the disk (slab-like; Tagliacozzo et al. 2023; Gianolli et al. 2023, 2024), rather than the point-like 'lamp-post model', as shown in Figure 1.4. Regardless of the geometry, the corona must be located very close to the central SMBH due to light-travel-time arguments from variability studies and micro-lensing measurements (e.g. Risaliti et al. 2007; Chartas et al. 2009; Fabian et al. 2015).

As shown in Figure 1.4, the coronal X-ray emission can be further reprocessed by surrounding components, resulting in additional X-ray spectral features of AGN. The first of which is the 'soft excess' (pink dot-dashed curve in Figure 1.5), a surplus of X-ray emission < 2 keV with respect to primary coronal power-law emission of debated origin (e.g. Ross et al. 1999; Gierliński & Done 2004; Crummy et al. 2006; Waddell et al. 2024, see also recent work by Chen et al. 2025, incl. Z. Igo, who are among those authors favouring a model involving 'soft Comptonisation' of seed photons by a secondary, optically thick 'warm corona' sandwiched over the inner-disk, instead of reprocessing from the hot corona). Then reflection off dense structures such as the accretion disk and torus can result in a strong iron Kα fluorescence emission line (Fe Kα; 6.4 keV) and the Compton hump, peaking at $E \sim 20 - 30$ keV (green dashed line in Figure 1.5; e.g. George & Fabian 1991; Nandra & Pounds 1994; Bauer et al. 2015). Interestingly, the Fe Kα emission line profile is rich in information on the properties of the innermost regions of the accretion flow, for example the R_{ISCO} and thus the black hole spin (Tanaka et al. 1995; Reynolds 2021; Fabian et al. 1989). Additionally, the torus and various forms of ionised outflowing gas cause photoelectric absorption at soft X-ray energies. For example 'warm absorbers' (see Ch. 1.7) can imprint narrow absorption lines and edges from partially ionised species such as O VII and O VIII (e.g. Crenshaw et al. 2003; Blustin et al. 2005). X-ray photons are also absorbed by the Earth's atmosphere and so we can only observe them with space-based telescopes such as the *extended ROentgen Survey with an Imaging Telescope Array* (eROSITA; Sunyaev et al. 2021; Predehl et al. 2021). It is again important to separate X-ray emission coming from the AGN and other sources of galactic X-ray emission, such as unresolved XRBs, whose disks emit in the soft X-ray regime, and hot gas, emitting Bremsstrahlung radiation (see Chs. 2.2.2 and 5.3.3).

Lastly, in the radio regime ($\lambda \gtrsim 0.01 - 1$ m), bi-polar jets can be observed in around 10% of AGN, emitting synchrotron radiation (see Ch. 1.3.1) due to the radial acceleration

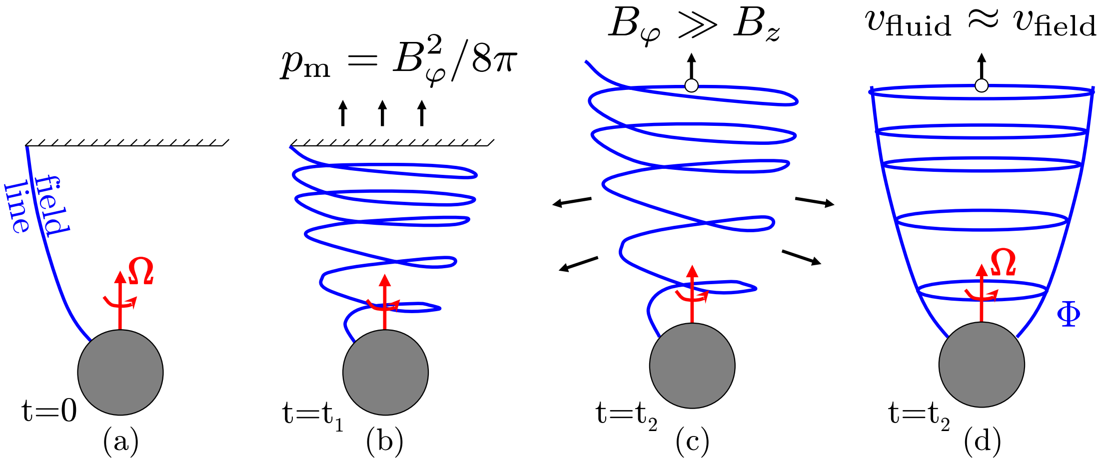


Figure 1.6: Schematic showing the launching of a relativistic jet, tapping the spin power of the black hole via the Blandford-Znajek mechanism (Blandford & Znajek 1977, see text for details). Figure taken from Davis & Tchekhovskoy (2020).

of relativistic particles (e.g electrons), following helical paths set by magnetic fields (e.g. Condon 1992; Saikia 2022; Hardcastle & Croston 2020a). Jets can be launched from sub-pc scales near the SMBH and extend to kiloparsec or even megaparsec scales. The current record holder for the jet with the largest physical size stands at over 7 Mpc (or 23 million light-years[5]; Oei et al. 2024) and is thought to be a potential way to permeate the Cosmic Web of the Universe with magnetic fields. Magnetic fields are thought to be crucial for jet-launching, as shown by Figure 1.6, and so is accretion onto the black hole (e.g. Davis & Tchekhovskoy 2020, recall Figure 1.3). However, it is not yet fully understood if it is the black hole spin (Blandford & Znajek 1977) or accretion disk rotation (Blandford & Payne 1982) that mainly contributes to the powering of jets. Figure 1.6 shows a simple schematic of the launching of a jet from a spinning black hole (Tchekhovskoy 2015; Davis & Tchekhovskoy 2020). First, we can consider a purely poloidal magnetic field line attaching a perfectly conducting sphere to the perfectly conducting 'ceiling' (panel a). The sphere has a spin set by angular frequency, Ω, and begins to twist the field lines, forming a toroidal component which exerts an effective pressure, $p_m = B_\varphi^2/8\pi$ (panel b). Effectively, this forms a 'magnetic spring' which pushes away the 'ceiling' and any plasma in its vicinity, thereby forming a jet (panel c and d)[6]. In practice, the poloidal magnetic field lines are anchored to the accretion flow and are advected towards the black hole. Unlike the sphere, the black hole does not have a physical surface, but the 'magnetic spring' can nevertheless be created by the rotation of space-time which twists the field lines (frame dragging). We can use dimensional analysis to derive the jet power, Q (Blandford & Znajek 1977), by taking the product of the magnetic energy density, $\propto B^2$, the cross-section of the jet base,

---

[5] Note that the diameter of the Milky Way is around 100,000 light-years.

[6] The toroidal component of the magnetic field gives rise to the Lorentz force $\vec{F} \propto \vec{v} \times \vec{B}$ as now the field has a perpendicular component to the velocity of the jet.

$\propto R_g^2$, and the speed at which the energy flows through the jet, $v \sim c$, and considering that it has to depend on an even power of the spin (as positive prograde and negative retrograde values should produce equivalent jet powers, by symmetry) giving

$$Q \propto a^2 B^2 R_g^2 c. \tag{1.13}$$

Such radio jets in AGN have been observed to show a variety of morphologies. This includes compact, unresolved cores of radio emission or resolved, complex, extended structures such as the canonical double-lobed jet shown in Figure 1.4 or bright hotspots, as these collimated streams of energy and particles collide with the ISM (see Chapters 1.5, 1.7 and 2). Historically, resolved jetted structures with core-dominated morphologies are named 'Fanaroff-Riley I' (FRI) systems, whereas lobe-dominated morphologies are classed as FRIIs (Fanaroff & Riley 1974a). We can observe these various radio sources with, for example, the Low Frequency Array (LOFAR; van Haarlem et al. 2013).

1.3.1 Synchrotron radiation

Since much of Chapters 2, 3, and 4 deal with jetted AGN, it is worth expanding on the main emission process that we use to probe the physics of AGN jets: synchrotron radiation (see Condon & Ransom (2016) for a full derivation). Fundamentally, Larmor's equation states that EM radiation P is produced by accelerating an electric charge q:

$$P = \frac{2q^2\dot{v}^2}{3c^3}, \tag{1.14}$$

where $\dot{v} = dv/dt$ is the acceleration of the charge. Substituting in for $\dot{v} = qv\gamma B \sin\theta/(m_e c)$ given the Lorentz force experienced by the charge (in CGS units), we can write P as,

$$P = 2\sigma_T \beta^2 \gamma^2 c\, U_B \sin\theta, \tag{1.15}$$

where $U_B = B^2/8\pi$ is the energy density of the magnetic field and θ is the pitch angle that accounts for the angle of motion of the charge with respect to the magnetic field. The Lorentz factor $\gamma = 1/\sqrt{1-\beta^2}$, where $\beta = v/c$, quantifies how strongly time, length, and mass are modified for an object moving at a relativistic speed.

In order to calculate the synchrotron spectrum of an optically thin radio source consisting of a distribution of electrons, we can assume that each electron radiates all of its average power (Eq. 1.16) at a single frequency.

$$\langle P \rangle = \left\langle \frac{dE}{dt} \right\rangle = \frac{4}{3}\sigma_T \beta^2 \gamma^2 c U_B, \tag{1.16}$$

This single frequency is equal to the peak (or critical) frequency, ν_c,

$$\nu = \nu_c \approx \gamma^2 \nu_g \approx \gamma^2 \frac{qB}{2\pi m_e}, \tag{1.17}$$

where ν_g is the gyrofrequency of an electron (with mass m_e, charge q) following helical path set by a magnetic field with field strength B.

Then we can use the fact that the energy distribution of cosmic-ray electrons in most synchrotron sources follows roughly a power-law of the form

$$n(E)\,dE \propto E^{-p}\,dE, \tag{1.18}$$

where $n(E)\,dE$ is the number of electrons per unit volume with energies E to $E + dE$. Therefore, we can express the emission coefficient, j_ν (i.e. the power radiated per unit volume, per unit frequency, per unit solid angle) as the product of the energy loss per electron and $n(E)\,dE$:

$$j_\nu\,d\nu \propto \left(-\frac{dE}{dt}\right) n(E)\,dE, \tag{1.19}$$

taking into account the assumption from Eq. 1.17. Each electron has energy:

$$E = \gamma m_e c^2 \approx \left(\frac{\nu}{\nu_g}\right)^{1/2} m_e c^2. \tag{1.20}$$

Differentiating Eq. 1.20 with respect to frequency gives:

$$dE \approx \frac{m_e c^2 \nu^{-1/2}}{2\nu_g^{1/2}}\,d\nu, \tag{1.21}$$

so

$$j_\nu \propto \left(\frac{4}{3}\sigma_T \beta^2 \gamma^2 c U_B\right)(E^{-p})\left(\frac{m_e c^2 \nu^{-1/2}}{2\nu_g^{1/2}}\right). \tag{1.22}$$

This can be simplified in terms of ν and B only, given that $\nu_g \propto B$:

$$j_\nu \propto B^{(p+1)/2} \nu^{(1-p)/2}. \tag{1.23}$$

Therefore, the synchrotron radiation follows a power-law distribution, $f_\nu \propto j_\nu \propto \nu^{-\alpha}$ with spectral index $\alpha = (p-1)/2$ (note that the negative convention of α is adopted throughout this Thesis). Typical optically thin synchrotron emitters show spectral indices $\alpha \sim 0.7$, which is in agreement with the typical cosmic ray spectra energy distribution with $p \sim 2.4$ (Rybicki & Lightman 1986).

Furthermore, radiative losses preferentially remove the highest-energy electrons so that the optically-thin slope steepens above the synchrotron critical frequency

$$\nu_c \propto B^{-3} t^{-2} \tag{1.24}$$

(c.f. Eqs. 1.16 and 1.17). This encodes information on the synchrotron age t of the electron population and is the reason why observing at different radio frequencies produces different integrated flux maps: higher or lower frequencies probe more recent or more ancient synchrotron injections, respectively.

Then, if the synchrotron source is sufficiently bright, due to high electron densities and/or magnetic field strengths, the low-frequency emission cannot freely escape and the source becomes optically thick to its own radiation through synchrotron self-absorption. The same population of relativistic electrons that emit the synchrotron radiation also re-absorbs it. This happens below a turnover frequency, determined by the condition where the effective temperature of the relativistic electrons

$$T_e = \frac{E}{3k_b} = \frac{\gamma m_e c^2}{3k_b}, \tag{1.25}$$

equals the brightness temperature, T_b, in the Rayleigh-Jeans limit,

$$T_b = \frac{I_\nu c^2}{2k_b \nu^2}, \tag{1.26}$$

where I_ν is the specific intensity. Substituting for γ in Eq. 1.25, we find $T_e \propto \nu^{\frac{1}{2}}$, and so the observed spectrum scales as $I_\nu \propto f_\nu \propto \nu^{5/2}$. In practice, optically thick compact radio cores are synchrotron self-absorbed, but typically show flat or inverted spectra ($\alpha \sim 0$ to -0.5) because the observed emission comes from a superposition of many self-absorbed components with different turnover frequencies.

1.3.2 Inverse Compton scattering

A further important physical mechanism to understand for Chapters 2, 3 and 5 is the main X-ray emission mechanisms in the corona, that is inverse Compton scattering (Rybicki & Lightman 1986).

Let us consider the interaction between a photon and electron, as shown in Figure 1.7, where the four-momenta before the interaction are $\vec{P}_{e_1} = [\gamma_1 m_e, \gamma_1 m_e \vec{v_1}]$, $\vec{P}_{\gamma_1} = [\frac{\hbar\omega_1}{c^2}, \frac{\hbar\omega_1}{c}\vec{n}_{\gamma_1}]$, and after are $\vec{P}_{e_2} = [\gamma_2 m_e, \gamma_2 m_e \vec{v}_2]$, $\vec{P}_{\gamma_2} = [\frac{\hbar\omega_2}{c^2}, \frac{\hbar\omega_2}{c}\vec{n}_{\gamma_2}]$. Under the assumption that the four-momentum is conserved:

$$\vec{P}_{e_1} + \vec{P}_{\gamma_1} = \vec{P}_{e_2} + \vec{P}_{\gamma_2}. \tag{1.27}$$

Squaring this equation and using Lorentz invariance ($\vec{P}_{e_1} \cdot \vec{P}_{e_1} = \vec{P}_{e_2} \cdot \vec{P}_{e_2} = m_e c^2$ and $\vec{P}_{\gamma_1} \cdot \vec{P}_{\gamma_1} = \vec{P}_{\gamma_2} \cdot \vec{P}_{\gamma_2} = 0$) to substitute out some terms, we get:

$$\vec{P}_{e_1} \cdot \vec{P}_{\gamma_1} = \vec{P}_{e_2} \cdot \vec{P}_{\gamma_2}. \tag{1.28}$$

Multiplying through by $\vec{P}_{\gamma_2}$ and simplifying, given $\vec{P}_{e_2} \cdot \vec{P}_{\gamma_2} = \vec{P}_{e_1} \cdot \vec{P}_{\gamma_1}$ and $\vec{P}_{\gamma_2} \cdot \vec{P}_{\gamma_2} = 0$:

$$\vec{P}_{e_1} \cdot \vec{P}_{\gamma_2} + \vec{P}_{\gamma_1} \cdot \vec{P}_{\gamma_2} = \vec{P}_{e_1} \cdot \vec{P}_{\gamma_1}. \tag{1.29}$$

Using the four-product and the geometry of the interaction to simplify dot products as $\vec{v}_1 \cdot \vec{n}_{\gamma_2} = v_1 \mathrm{cos}\theta_2$, $\vec{v}_1 \cdot \vec{n}_{\gamma_1} = v_1 \mathrm{cos}\theta_1$, $\vec{n}_{\gamma_1} \cdot \vec{n}_{\gamma_2} = \mathrm{cos}\theta_3$, we can write the ratio of photon energies as:

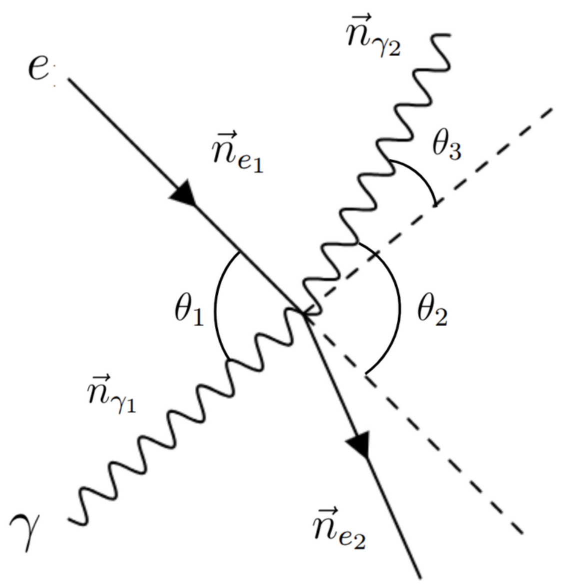


Figure 1.7: Feynman diagram of Compton scattering on an electron and photon.

$$\frac{\omega_2}{\omega_1} = \frac{(1 - \frac{v_1}{c}\mathrm{cos}\theta_1)}{(1 - \frac{v_1}{c}\mathrm{cos}\theta_2) + \frac{\hbar\omega_1}{\gamma_1 \mathrm{m_e c^2}}(1 - \mathrm{cos}\theta_3)}. \tag{1.30}$$

For the case of a stationary target ($v_1 = 0$; $\gamma = 1$), this reduces to standard Compton scattering where the Compton wavelength shift follows $\Delta\lambda = \lambda_2 - \lambda_1 = \frac{h}{m_e c}(1 - \mathrm{cos}\theta_3)$. This can be recast in terms of the change in energy of the photon as $\Delta\varepsilon = \varepsilon_2 - \varepsilon_1 = -\frac{\varepsilon_1^2\,(1-\mathrm{cos}\theta_3)}{m_e c^2}$, meaning the photon transfers some of its energy to the electron.

For the case of a relativistic electron, where the incoming photon is less energetic than the electron rest energy in the electron rest frame ($\hbar\omega' = \gamma_1\hbar\omega_1(1 - \frac{v_1}{c}\mathrm{cos}\theta_1) \ll \mathrm{m_e c^2}$), the recoil term in the denominator is negligible and Eq. 1.30 simplifies to

$$\frac{\omega_2}{\omega_1} \approx \frac{1 - \frac{v_1}{c}\mathrm{cos}\theta_1}{1 - \frac{v_1}{c}\mathrm{cos}\theta_2}. \tag{1.31}$$

As θ_1 and θ_2 can take on many different values, the energy of the photon can either decrease or increase. However, when properly averaged over angles with relativistic aberration and Doppler boosting taken into account, there is on average a net increase in the photon energy. This process is called ‘inverse Compton scattering’. In the special scenario where the photon and electron collide head-on and the electron recoils back the way it came from ($\theta_3 = 180°, \theta_1 = 180°, \theta_2 = 0°$), we can derive a maximal boosting factor of: $\frac{\omega_2}{\omega_1} = \frac{1+(v_1/c)}{1-(v_1/c)} \approx 4\gamma_1^2$, where $v_1/c = (1 - \gamma_1^{-2})^{\frac{1}{2}}$ has been Taylor expanded.

In an astrophysical context, there is a distribution of electrons and photons that result in multiple Compton scattering events. It can be shown that Eq. 1.30, averaged over an isotropic distribution of both photons and electrons, reduces to an average fractional energy per scattering (Rybicki & Lightman 1986) of

$$\left\langle \frac{\Delta\varepsilon}{\varepsilon_1} \right\rangle = \frac{4}{3}\gamma^2\beta^2 - \frac{\varepsilon_1}{m_e c^2}. \tag{1.32}$$

In the case of a (quasi-)thermal distribution of electrons, the electrons will have a typical root-mean-square (rms) velocity of $v^2 = \frac{3k_B T_e}{m_e} = 3\Theta c^2$, where Θ is the dimensionless average electron temperature. A thermal distribution is a good assumption for the electrons in the hot corona as electron-electron Coulomb interactions allow for the electrons to efficiently thermalise (but see discussion in e.g. Lightman & Zdziarski 1987; Malzac & Belmont 2009). Simplifying Eq. 1.32, in the limit of $\gamma \sim 1$ and $\frac{\varepsilon_1}{m_e c^2} << \Theta$, which is valid for the seed accretion disk photons, we get

$$\left\langle \frac{\Delta\varepsilon}{\varepsilon_1} \right\rangle \approx 4\Theta, \tag{1.33}$$

meaning that on average the seed photons are up-scattered by a factor of $(1 + 4\Theta)$ per scattering in the hot corona. Such scatterings happen multiple times to produce a 'Comptonised spectrum' with a power-law distribution (Pozdnyakov et al. 1976; Rybicki & Lightman 1986) described by

$$I(\varepsilon) = I(\varepsilon_1)\left(\frac{\varepsilon}{\varepsilon_1}\right)^{-\alpha'}, \tag{1.34}$$

where ε_1 is the average energy of the UV photons and

$$\alpha' = -\ln\tau/\ln \mathrm{A}_1, \tag{1.35}$$

where τ is the optical depth and A_1 is the mean amplification factor of the photon energy in a single scattering, $A_1 = 1 + \langle\varepsilon_2/\varepsilon_1\rangle$, from Eq. 1.33. To find the value of α', Haardt & Maraschi (1991) proposed a 'two-phase' disk–corona model, where the corona lies above the optically thick accretion disk and a fraction fraction f of the gravitational power, P_G, is dissipated in the hot corona, whilst $(1 - f)P_G$ is dissipated within the cold disk. Under the assumption of radiative equilibrium, Haardt & Maraschi (1991) derived a set of energy balance equations that determined the value of A_1 required for the Comptonisation process. This in turn defined τ and T_e, and thereby α'. Interestingly, they found that there is a natural adjustment of the T_e and τ to keep a relatively constant α'. For typical AGN coronae, with $k_B T_e \sim 100$ keV, this results in the observed photon indices $\Gamma = \alpha' + 1$ of around $1.8 - 2$ (e.g. Fabian et al. 2015; Brandt & Alexander 2015; Ricci et al. 2018).

1.4 Black holes across the mass scale

In Ch. 1.2, we described the basics of how matter is accreted by a black hole, but the question remains: is this accretion mechanism mass-invariant? For example, stellar mass

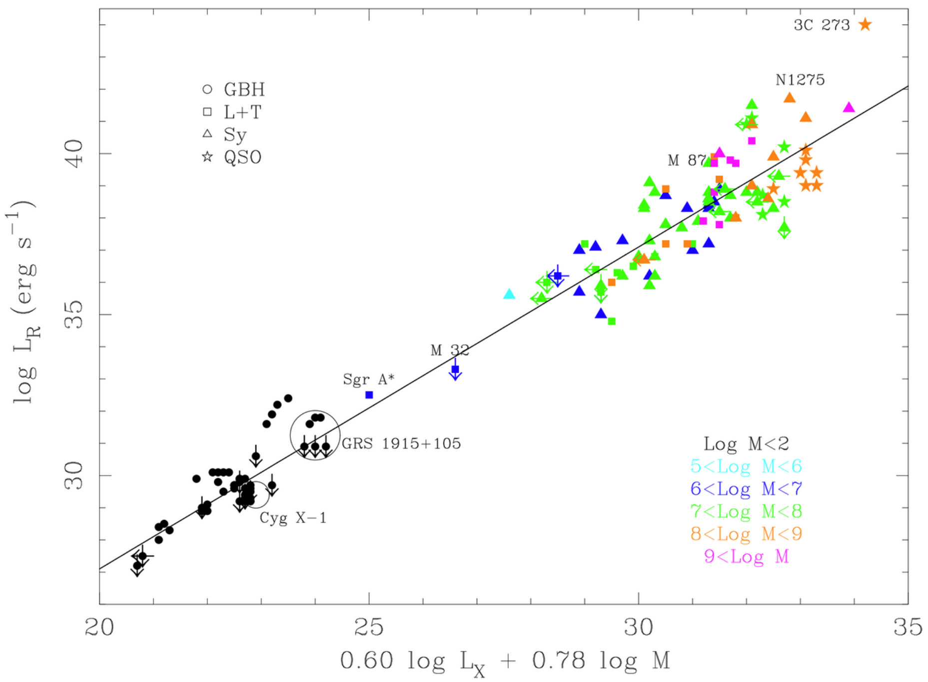


Figure 1.8: Fundamental plane of black hole accretion from Merloni et al. (2003) that unifies radiatively inefficient stellar and supermassive black holes by their radio and X-ray emission.

black holes in binary systems have been found to go through hysteresis cycles of quiescent and outburst stages (e.g. Fender et al. 2004). This is interesting in the context of this Thesis, as in the quiescent state the luminosity is low, the emission is hard ($\Gamma \lesssim 2$) and there are persistent radio jets, whereas in the outburst state the luminosity is high, the emission is soft and there are no persistent jets (only intermittent ones, some of the time). These are very similar properties to the AGN accretion modes presented in Figure 1.3 (see also e.g. Svoboda et al. 2017), except for the fact that the week to month long quiescent-outburst cycles in XRBs would correspond to millions of years for an AGN (as given by a comparison of the respective viscous timescales).

Intrigued by these striking similarities across the mass scale, Merloni et al. (2003), investigated the radio and X-ray properties of a sample of XRBs and AGN, accreting in a radiatively inefficient mode, only to find that these followed a tight correlation, as shown in Figure 1.8. This 'fundamental plane of black hole activity' is underpinned by theory from Heinz & Sunyaev (2003), who propose that the dynamics of the inner accretion disk is scale invariant, given the assumption that it is only dependent on gravity, and therefore the jets launched from such inner regions should have similar properties. They derive a jet flux which scales as $f_{\nu} \propto M_{\rm BH}^{1.42-0.33\alpha}$, where α is the radio spectral index defined in Ch. 1.3.1. This non-linear relationship shows that AGN jets are much more powerful than XRB jets. Similarly, Falcke et al. (2004); Körding & Falcke (2005); Körding et al. (2006)

establish tight radio–X-ray correlations that hold for black holes across the mass scale from X-ray binaries to AGN, along with extending the fundamental plane of black hole activity to encompass different accretion modes. They ultimately conclude that even with the large range in black hole masses, jet formation may be a universal, mass-invariant process.

In Chapters 2 and 3, we aim to test if the mass-invariant accretion (to first order; see Ch. 1.6) translates into a similarly mass-invariant mechanisms that powers jets. Additionally, in Chapter 5 we further vet this idea of mass-invariant accretion by looking at a previously elusive parameter space of black holes with masses in between that of stellar mass and supermassive black holes.

1.5 Detecting AGN with multi-wavelength surveys

As we have seen, AGN produce emission spanning the full EM spectrum. We now turn to some of the methods by which this radiation can be observed and quantified, namely via X-ray, radio and optical surveys (recall Figure 1.3). Surveys provide the statistical power needed to unveil robust trends and uncover new sources with potentially new physics. This power was already tapped into in the late 1950s, with the Cambridge Catalogues of Radio Sources, for example its third edition (3C), in which the radio quasar 3C 273 from Ch. 1.1 was discovered. In fact, every survey is shaped by the need to address a set of focused and innovative scientific questions, and its design evolves in parallel with ongoing technological advancements and lengthy feasibility studies.

1.5.1 X-ray Surveys with eROSITA

The science driver behind eROSITA is to study the large scale structure of the Universe and characterise cosmological models including Dark energy, the mysterious entity that makes up almost 70% of our observable Universe (Merloni et al. 2012; Predehl et al. 2021; Sunyaev et al. 2021; Bulbul et al. 2024; Ghirardini et al. 2024). This is done via the eROSITA All Sky Surveys (eRASS; Merloni et al. 2024) to detect hundreds of thousands of galaxy clusters (the largest gravitationally bound structures in the Universe, containing 10-100s of galaxies) out to high redshifts $z > 1$. As a by product of this main science goal, millions of AGN have also been detected, in what is now the deepest all sky X-ray survey to date.

eROSITA is one of two instruments on the *Spektrum Roentgen Gamma* (SRG) mission launched in 2019, which orbits around the second Lagrange point (L2) of the Sun-Earth system, 1.5 million km from Earth. The other instrument is a hard X-ray detector, called *ART-XC*, which surveys the sky in the $\sim 5{-}30$ keV range (Pavlinsky et al. 2021). eROSITA observes the full sky in the $0.2 - 8$ keV band, but it is most sensitive to emission in its 'main band' of $0.2 - 2.3$ keV. Figure 1.9 shows a schematic of eROSITA, including the seven telescope modules (TMs), each with their own grazing-incidence X-ray mirror and pn charge-coupled device (pn-CCD) camera (Meidinger et al. 2020), similar to those that were originally developed for *XMM-Newton* (Strüder et al. 2001). Each TM has an X-ray baffle

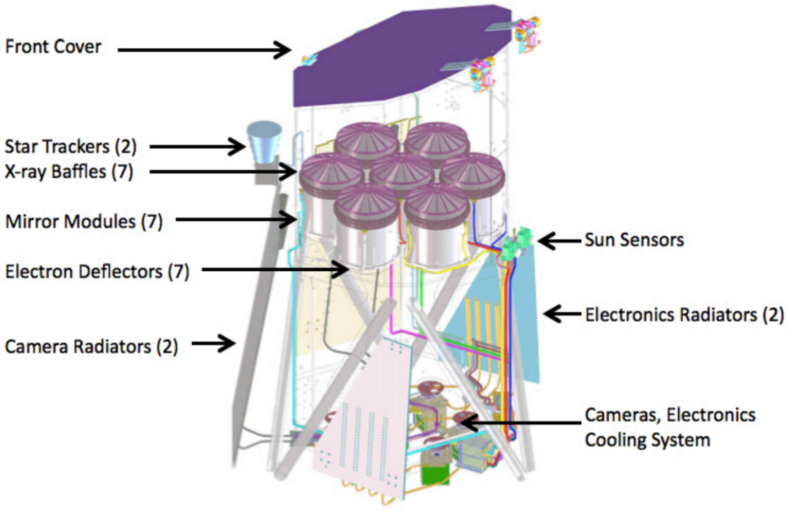


Figure 1.9: Schematic of the eROSITA telescope, with the seven mirror modules observing the X-ray sky in the range $0.2 - 8$ keV. Figure taken from Merloni et al. (2012).

to prevent unwanted X-ray photons (straylight) from entering the optical system (Friedrich et al. 2014). The mirrors have a Wolter Type 1 design with 54 concentric mirror shells with parabolic/hyperbolic shape to focus X-rays, via grazing incidence, to the detector. Their gold coating further enhances the reflection capabilities of the mirrors and imprints a characteristic gold edge in the effective area curve around 2.3 keV, as shown in Figure 1.10. In the same Figure, it is also clear that the combination of a wide field of view of eROSITA and sensitive soft effective area (i.e the grasp) out-performs prior focusing X-ray missions, enabling efficient detection of both point-like and diffuse, extended sources. The field-of-view average half energy width (HEW) of the point spread function, as measured from the source stacking method applied to the eRASS data, is $\sim 30''$ (Merloni et al. 2024). Additionally, the pn-CCDs have good spectral resolution with $75 - 82$ eV at 1.49 keV (Predehl et al. 2021). This allows for detailed X-ray spectroscopy to be carried out with eROSITA, although it is mostly limited to $E < 2$ keV (e.g. Waddell et al. 2024; Yeung et al. 2024).

In November 2019, as part of its Performance Verification program, eROSITA scanned the sky for four days to produce the eROSITA Final Equatorial Depth Survey (eFEDS; Brunner et al. 2022), a $\sim$140 deg^2 pilot survey of eRASS. eFEDS has a uniform exposure of ~ 1.2 ks (after correcting for vignetting) and the point source catalogue includes 21,952 candidate AGN detected in the main $0.2 - 2.3$ keV band (Liu et al. 2022b). Chapters 2 and 3 exploit the eFEDS survey and its excellent multi-wavelength coverage.

For its full survey-mode, eROSITA scans the entire sky in a period of six months, with a scan rate of ~ 0.025 deg s^{-1} (Merloni et al. 2024). It does so by drawing out great circles on the sky, intersecting at the ecliptic poles, and shifting the rotation axis by around one degree per day following the motion of the Earth around the Sun. During

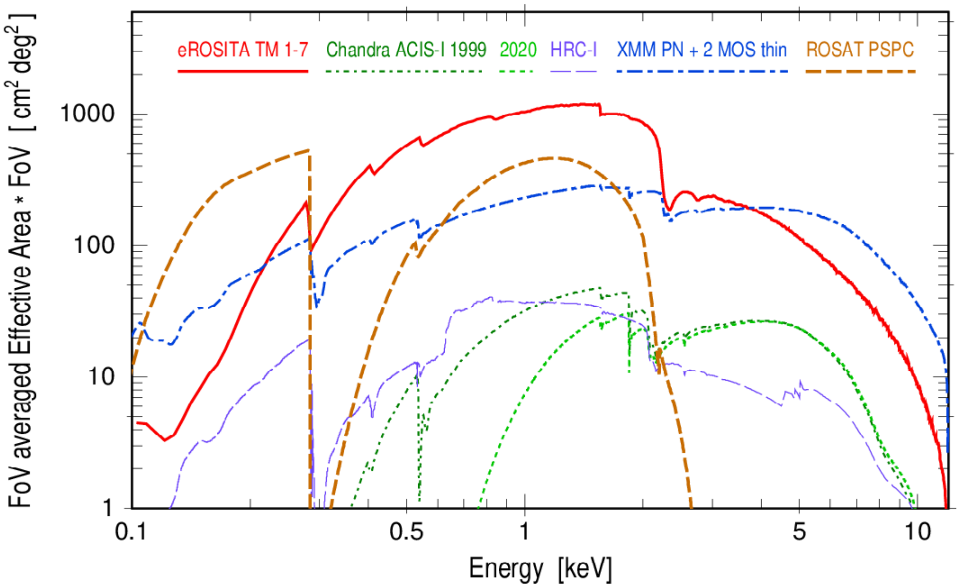


Figure 1.10: The product of the field of view and (averaged) effective area ('grasp'), as a function of energy for recent X-ray missions. eROSITA is shown by the solid red curve and has higher grasp than any other mission in the $0.2-2.3$ keV band. The grasp of Chandra ACIS-I in 1999 and 2020 is plotted dark and light green, respectively, along with Chandra HRC-I in purple, XMM-Newton in blue, and ROSAT in brown. Figure taken from Predehl et al. (2021).

each 6-month scan, a source near the ecliptic equator is visited typically six times in intervals of 4 h, for up to 40 seconds per visit, whereas sources near the ecliptic poles are visited $70-1080$ times (Bogensberger et al. 2024), offering very deep exposures in these regions. This variety of different timescales has allowed eROSITA to be a powerful X-ray time-domain discovery machine, finding new types of quasi-periodic-eruptions and tidal disruption events (e.g. Arcodia et al. 2021; Malyali et al. 2023; Liu et al. 2023; Baldini et al. 2025). From December 2019 to February 2022[7], eROSITA completed 4.4 All Sky Surveys, totalling to an average uniform exposure time of ~ 800 s (with up to 120 ks at the ecliptic poles) and ~ 1.7 billion $0.2-5$ keV calibrated photons (~ 380 Gb of telemetry). In the $0.2-2.3$ keV band, a single pass (eRASS1, 2, 3, 4) has a typical point source sensitivity of $\sim 5\times10^{-14}$ erg s^{-1} cm^{-2}, which improves to $\sim 2\times10^{-14}$ erg s^{-1} cm^{-2} in the cumulative eRASS:4 survey. Chapter 5 uses the deep eRASS:4 data. A comparison of the number of sources detected in just the first eRASS, already outnumbering previous X-ray missions, is shown in Figure 1.11.

---

[7] On the 26th of February 2022, eROSITA was put into safe mode: `https://www.mpe.mpg.de/7912248/news20221117`.

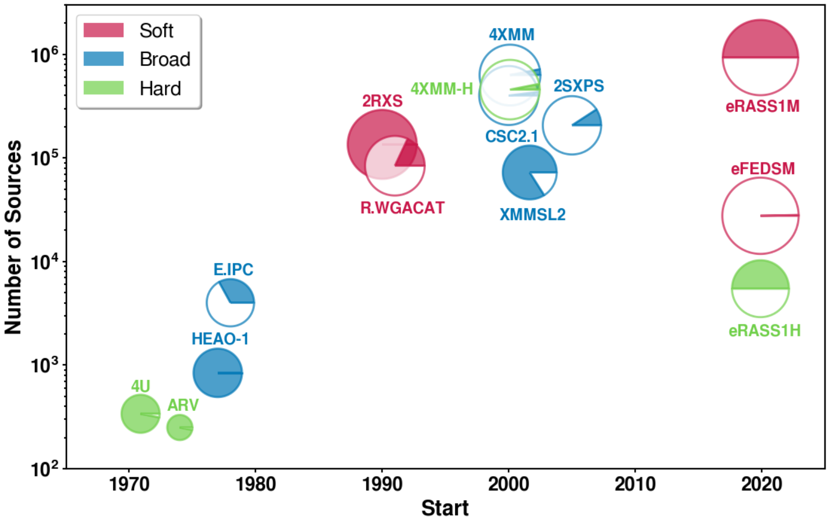


Figure 1.11: Graphical comparison of selected X-ray catalogues based on the data presented in Table 9 of Merloni et al. (2024). For each catalogue the total number of objects vs. the time of the start of the corresponding data-taking period are plotted. Each catalogue is represented by a circle, whose radius is proportional to the logarithm of the 'discovery rate' (number of objects divided by the number of months of operations), while the shaded portion of the circle represents the fraction of the sky covered by the catalogue. The colour coding distinguishes telescopes operating mainly in the soft (< 2.5 keV; red), hard (> 2.5 keV; green) or broad (both soft and hard; blue) X-ray band. Figure adapted from Merloni et al. (2024) with credits to Z. Igo.

eROSITA X-ray spectral and timing case study: SMSS J114447.77-430859.3

To illustrate the power of eROSITA, we show the X-ray spectra and long-term lightcurves of SMSS J114447.77-430859.3, or J1144 for short, which is the brightest quasar in the last 9 Gyr of the Universe ($z = 0.83$). This is work published in Kammoun et al. (2023), where I was responsible for the eROSITA data analysis.

J1144 is an object at intermediate Ecliptic latitudes meaning that it passes in the eROSITA field of view around $8 - 10$ times every 6 months, with those passes (each less than 40 s) all occurring within 1–2 days. J1144 was observed in all eRASS1–5. Data were reduced in the standard way using the eROSITA Science Analysis Software System (eSASS) eSASSusers_211214 pipeline version c020 and c947 for eRASS1–4 and 5, respectively (Brunner et al. 2022; Merloni et al. 2024).

Figure 1.12 shows the observed $0.5 - 10$ keV flux for eROSITA, *Swift*, and *XMM-Newton* (black squares, blue circles, and red diamond, respectively), based on the best-fit spectral model (see below for the case of eROSITA and refer to Kammoun et al. 2023, for further details). The source shows a clear variability by factors of ~ 10 and ~ 2.7 over timescales

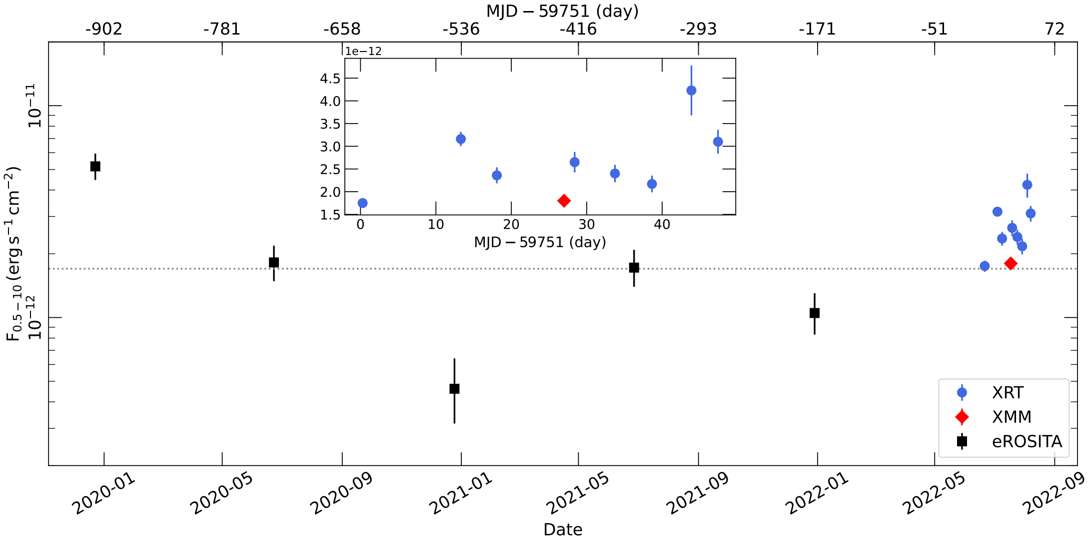


Figure 1.12: The long-term X-ray light curve of J1144 in the 0.5-10 keV band. The black squares, blue circles, and red diamond correspond to eROSITA, *Swift*, and *XMM-Newton*, respectively. The inset shows a zoom-in on the recent monitoring of the source in 2022, with the start being the first *Swift* observation (MJD = 59751). The grey dotted line corresponds to the 3σ upper detection limit by ROSAT obtained in 1990.

of a year and of a few days, respectively.

Light curves were also extracted using eSASS and analysed using `bexvar` to search for variability (Buchner et al. 2022). Variability in `bexvar` is quantified by the intrinsic scatter (σ_{bexvar}) on the assumed log-normal distribution of count rates in any given time bin (see Sec. 3 of Buchner et al. 2022). This log-scatter on the log-count rate is similar in concept to the excess variance on the linear count rate, usually quoted as the normalised excess variance (NEV; e.g., Vaughan et al. 2003). However, low values of $\sigma_{\text{bexvar}} \ll 0.1$ are consistently found for all eRASS1–5 over several energy bands. Overall, the long term (inter-eRASS) X-ray variability, clearly seen in Fig. 1.12, is much more significant than the short term (intra-eRASS) variability. For that reason, in the following, we model the time-averaged spectrum from each eRASS.

Spectral fitting was performed using the `pyXspec` X-ray analysis environment of `XSPEC v12.12.0` (Arnaud 1996; Gordon & Arnaud 2021) coupled with Bayesian X-ray Analysis (`BXA v4.0.0`; Buchner et al. 2014; Buchner 2021a), a Bayesian parameter estimation and model comparison package using the nested sampling algorithm UltraNest (Buchner 2021b). The spectra, grouped by each single-pass eRASS, from the source and background regions were jointly fitted using a source model, consisting of an absorbed power law (see below), plus background model. The background model was calculated following Simmonds et al. (2018), by applying Principal Component Analysis (PCA) on an existing parametric model for eROSITA background spectra and adding Gaussian lines until it no longer improved the fit, as judged by the Akaike Information Criterion (see also Section 3.1

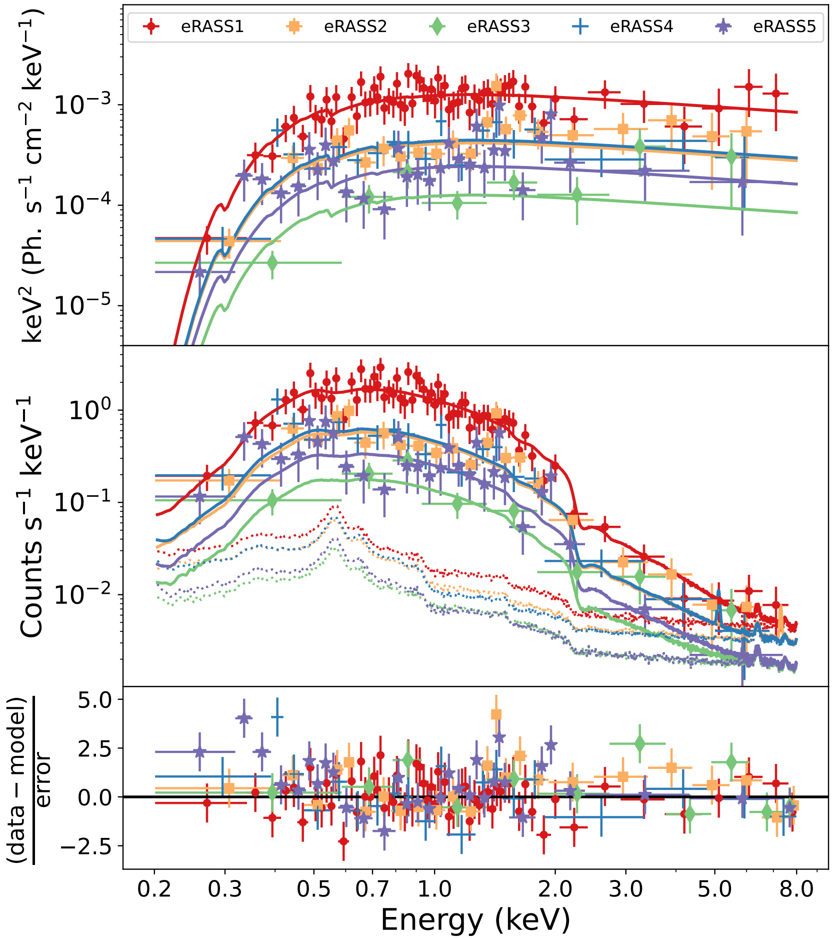


Figure 1.13: Observed spectra from the different eROSITA observations (eRASS1−5: red circle, yellow square, green diamond, blue cross, purple star). Top: Deconvolved spectra. Middle: Convolved spectra, folded through detector response but not effective area, along with background spectra for each eRASS (dotted, same colour scheme). Bottom: normalised residuals obtained by fitting the spectra.

of Liu et al. 2022c, for more details). The shape of background component was therefore already fixed in the joint modelling, but the normalisation was left free to vary so it could adjust to the required background flux level appropriate for this source.

We use a simple absorbed power law to model the source: `TBabs` × `zTBabs` × `zpow` in XSPEC formalism, where `TBabs` and `zTBabs` (Wilms et al. 2000) correspond to the Galactic absorption in the line of sight and the intrinsic neutral absorption at the redshift of the source, respectively. We fixed the Galactic column density at $N_{\mathrm{H}} = 7.9 \times 10^{20}\,\mathrm{cm}^{-2}$ (HI4PI Collaboration et al. 2016) but the intrinsic column density, the power law photon index, and power-law normalisations (“norm1-5”) were left free for each eRASS spectrum. Upon modelling each spectrum individually, it is found that the intrinsic column density

Table 1.1: Best-fit parameters obtained by modelling the eROSITA spectra. Note that the intrinsic column density (N_{H}) is a 3σ upper limit. Normalisations (norm1–5) are in units of Photon $\mathrm{keV^{-1}\,cm^{-2}\,s^{-1}}$.

$N_{\mathrm{H}}[\mathrm{cm^{-2}}]$	Γ	log(norm1)	log(norm2)	log(norm3)	log(norm4)	log(norm5)
$< 2.7 \times 10^{21}$	$2.21^{+0.14}_{-0.12}$	$-2.27^{+0.06}_{-0.05}$	$-2.72^{+0.07}_{-0.06}$	$-3.31^{+0.09}_{-0.10}$	$-2.75^{+0.06}_{-0.06}$	$-2.97^{+0.08}_{-0.06}$

and photon index are consistent with each other across eRASS1–5, albeit with large uncertainties, due to the quality of the data. Yet, the values of the normalisations show a clear variability across the five spectra. For this reason, we re-fit the spectra simultaneously by tying the column density and the photon index for all observations, and letting the normalisations free to vary. The prior set on the column density, the photon index, and the normalisations was a log-uniform prior ranging from $(10^{-3} - 10^{3}) \times 10^{22}\,\mathrm{cm^{-2}}$, a uniform prior from $1 - 4$, and a log-uniform prior from $10^{-6} - 10$ Photon $\mathrm{keV^{-1}\,cm^{-2}\,s^{-1}}$, respectively.

Figure 1.13 shows the results of this simultaneous fitting of eRASS1–5 spectra. The top panel shows the deconvolved spectra, to better compare with observations from different instruments. The middle panel shows the convolved spectra, folded through detector response but not effective area, along with background spectra (dotted lines). It is clear that the background becomes dominant past $2 - 3$ keV. The normalised residuals are also displayed in the bottom panel of the same figure. The apparent difference in the levels of background are caused by the varying extraction region sizes that scale proportional to the source flux. Quantitatively, the eRASS1 (highest flux state) source extraction region area is a factor of ~3 larger than that of eRASS3 (lowest flux state), and this is reflected in the difference in normalization between the red and green dotted lines. The best-fit (posterior median) photon index and 3σ upper limit on the column density for eRASS1–5 are $\Gamma = 2.21^{+0.14}_{-0.12}$ and $N_{\mathrm{H}} < 2.7 \times 10^{21}$ $\mathrm{cm^{-2}}$, respectively. Moreover, we find a factor > 10 decrease in the normalisation of the power law between eRASS1 and eRASS3, after which it steadily increases again up till eRASS5. Table 1.1 lists the best-fit parameters for this simultaneous fitting of eRASS1–5 spectra.

Overall, this very bright quasar has otherwise relatively ‘normal’ AGN properties. Its long-term light-curve reveals variability by factors of several over timescales of a year and of a few days. Its X-ray spectra are characterised by a relatively constant photon index of ~ 2 during this time, but variable intrinsic emission and hints of radiatively driven outflows manifesting as absorption features. Spectral energy distribution (SED) modelling reveals a sub-Eddington accretion mode with standard thin disk and a low temperature corona.

1.5.2 Radio surveys with LOFAR

LOFAR is a radio telescope array which has among its main science goals: (i) to study the epoch of reionisation using the 21 cm emission line from neutral hydrogen at high redshifts $z \gtrsim 6$; (ii) to survey the sky at low frequencies to understand of the formation and evolution of galaxies, AGNs and galaxy clusters over cosmic time; and (iii) to explore the transient radio sky (see van Haarlem et al. 2013, for more details). LOFAR is unique as it explores the low-frequency range from $10 - 240$ MHz, features automated processing pipelines to deliver fully calibrated science products to the public and has a large field of view (FoV) in the northern hemisphere, along with high angular resolution of $\sim 6\,''$ across its Two Meter Sky Surveys (LoTSS; Shimwell et al. 2017, 2019, 2022).

This is achieved by a set of antennas at various stations distributed across Europe, with the core station (i.e the 'Superterp', shown in Figure 1.4) located in the Netherlands. Rather than using steerable dishes, LOFAR comprises phased arrays of simple dipole antennas grouped into stations, each equipped with low-band (LBA) and high-band (HBA) elements. The basics of such a radio interferometer are shown in Figure 1.14, where each antenna pair, separated by baseline $\vec{B}$, measures an interference pattern from planar EM waves originating from a distant celestial source in direction $\hat{s}$ (Condon & Ransom 2016). Due to the separation between antenna 1 and 2, the EM signal arrives with a geometric

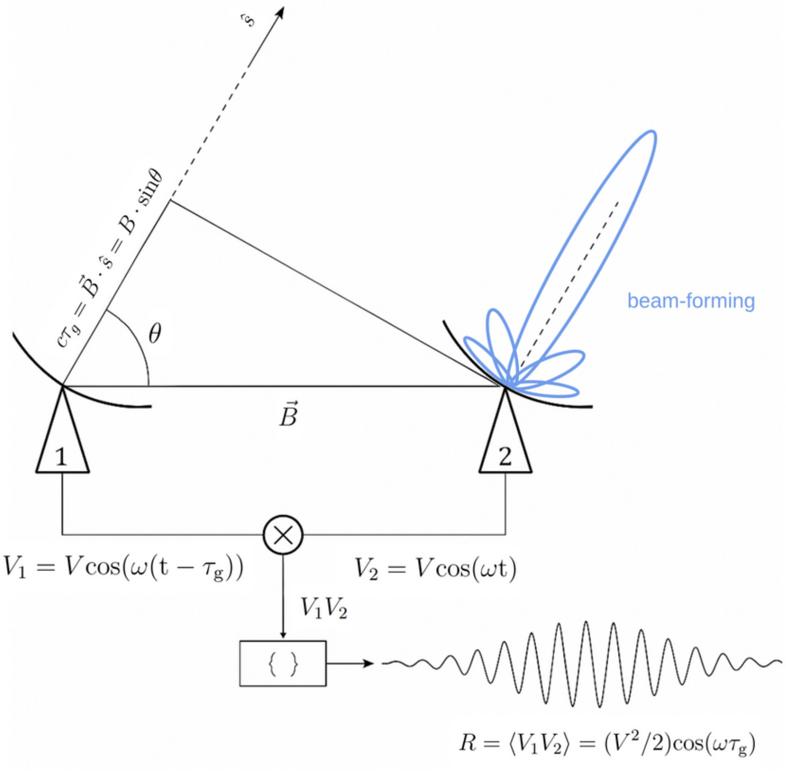


Figure 1.14: Schematic of a radio interferometer with two radio antennas separated by a baseline, $\vec{B}$, observing a celestial source in direction $\hat{s}$, which introduces a geometric time delay τ_g that depends on the angle θ to the source. The amplitude and phase of the signals are combined in a correlator. Figure adapted from Condon & Ransom (2016).

time delay of

$$\tau_g = \frac{\vec{B} \cdot \hat{s}}{c}, \tag{1.36}$$

leading to voltages, V_1 and V_2, that are phase-shifted, as shown on Figure 1.14. Each LOFAR antenna is effectively sensitive to the entire sky, so 'beam-forming' is used to focus the sensitivity in a particular direction defined by the beam (as is done with single dish radio telescopes, see blue lobes in Figure 1.14). In the case of LOFAR, analogue methods (such as increased cable lengths) enable the electronic pointing of the array by adjusting signal delays rather than physically moving components.

The signals are sent to the correlator to be multiplied together and averaged in time, producing a distinct interference (fringe) pattern. The resulting complex quantity, known as the visibility, encodes both the amplitude and phase information corresponding to a particular spatial frequency (Fourier component) of the sky brightness distribution. This forms the u–v plane, where the coordinates u and v are defined by the baselines between two antennas. As the Earth rotates, more spatial frequencies from each baseline are sampled in this plane. A 'dirty image' can then be constructed via an inverse Fourier transform of these visibilities. Since the $u-v$ plane is only partially sampled, the PSF in the dirty image contains a complex (beam) sidelobe structure, leading to artifacts. These can be removed by deconvolving the PSF (which is known from the $u-v$ coverage) from the dirty image to produce a sky model. Then it can be convolved again with a clean Gaussian beam to produce a 'clean image'. However, before the final image is obtained, the measured visibilities need to be corrected for various effects that alter the signal along its path, including varying gains, direction-dependent ionospheric phase distortions and clock drifts between stations (Shimwell et al. 2017, 2019; Tasse et al. 2021). Initial corrections are applied using the *Prefactor* pipeline for direction-independent effects. This is followed by direction-dependent calibration via the DDF-pipeline using bright compact calibrator sources distributed across the field. Finally, source detection algorithms (e.g. 'Python Blob Detector and Source Finder', PyBDSF; Mohan & Rafferty 2015) identify and characterise discrete radio sources across the cleaned radio image.

The maximum baseline length, $B_{\max}$, sets the achievable angular resolution (Eq. 1.37), which in the case of LOFAR is around 2000 km for the full international array and around 100 km for the core stations.

$$\theta \text{ [radian]} \approx \frac{\lambda}{\mathrm{B}_{\max}}. \tag{1.37}$$

Radio interferometry offers a significant advantage in angular resolution compared to single dish radio telescopes, where $B_{\max}$ is instead set by the diameter of the dish (Condon & Ransom 2016). This makes LOFAR, with its excellent baseline coverage, a powerful survey instrument as it enables the detection of structures even in compact sources thanks to its high $6''$ angular resolution (for LoTSS using the core stations), as well as the identification of larger scale diffuse emission (Shimwell et al. 2022), as shown in Figure 1.15. Recall from Eq. 1.24 that lower radio frequencies can also probe older populations of electrons, which tend to live in the lobes of radio galaxies.

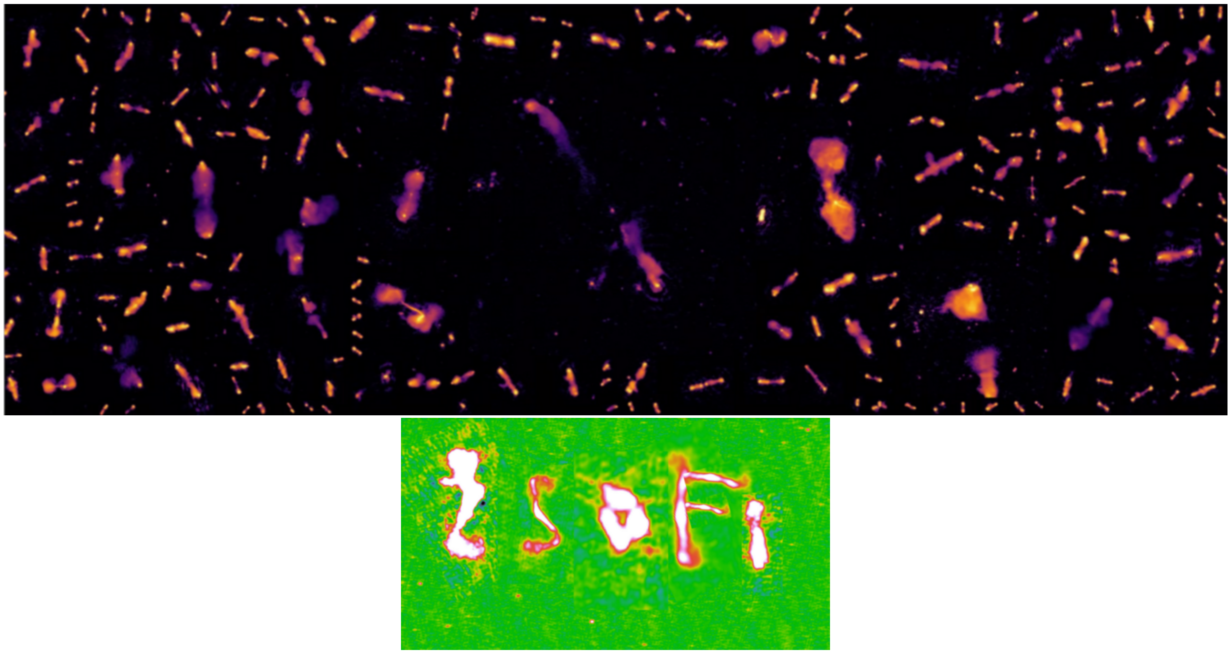


Figure 1.15: Top: Example radio galaxies in the LoTSS DR2 survey showing the diversity of morphologies that can be resolved with LOFAR. Figure credits to M. Hardcastle and the LoTSS Survey (Shimwell et al. 2022). Bottom: Example radio galaxies from the LOFAR-eFEDS survey that take on the form of my name (note the ‘F’ is an artistic combination of several radio galaxies).

The co-spatial LOFAR field to the eFEDS pilot survey forms the basis of Chapters 2, 3 and 4. This LOFAR-eFEDS field has HBA observations at 144 MHz frequencies with root-mean-square (rms) noise levels of $\sim 135\ \mu$ Jy beam^{-1} (Pasini et al. 2022). As it is a low declination field, the resolution is slightly compromised due to increased ionospheric effects compared to LoTSS at $8'' \times 9''$ angular resolution. PyBDSF was run on the mosaic images created for the full LOFAR-eFEDS field in order to model the radio emission with Gaussian components and produce the final source catalogue. A peak flux detection threshold of 5σ, where sigma is the local rms noise, was imposed for a source to be detected, calculated via sliding a 150 × 150 pixel box in 15 pixel steps. The diversity of radio galaxies that can be observed by LoTSS and LOFAR-eFEDS surveys are shown in Figure 1.15.

1.5.3 Optical surveys with DESI Legacy Imaging Survey and SDSS

Radio and X-ray surveys by themselves contain no information about the host galaxy from which AGN emission originates. To help with this, we need optical surveys of photometric or spectroscopic nature. Two important optical surveys for this Thesis are the *Dark Energy Survey Instrument* Legacy Imaging survey (henceforth LS, in particular the 9th and 10th data releases: LS9 and LS10; Dey et al. 2019) and SDSS. The former is a photometric all-sky survey with its main science driver being to identify and select targets for the 5-year spectroscopic DESI Survey, which released over 18 million spectroscopic targets just this year (DESI Collaboration et al. 2025a). It is advantageous because of its large, homogeneous sky coverage, photometric depth and accurate astrometry. Optical photometry in the *g, r, z* bands is included, with limiting AB magnitudes of 23.95, 23.54 and 22.50, respectively (LS10 also includes observations in the *i*-band for portions of the sky). The LS catalogues also include WISE forced photometry at the optical source coordinates, following Lang (2014); Lang et al. (2016), at 3.4 μm, 4.6 μm, 12 μm and 22 μm. This method ensures matched aperture photometry, making use of the higher optical angular resolution compared to that of WISE. The positional accuracy of these optical sources ($\sim$0.1$''$) exceeds that of radio and X-ray astrometry from LOFAR and eROSITA.

SDSS is a spectroscopic survey that has been active for over 20 years (York et al. 2000) and its multi-object spectrograph enables tens to hundreds of objects per square degree to be simultaneously observed. Its core science drivers include reliably mapping the large scale structure, which requires accurate astrometry and redshift determination, and delivering high-quality spectroscopic catalogues to the community. The most recent public release is SDSS-V DR19 (Kollmeier et al. 2025). Figure 1.16 shows some example stacked optical spectra from X-ray AGN in the eFEDS field (Aydar et al. 2025), highlighting the power-law continuum, along with the broad and narrow emission lines (with some prominent host galaxy absorption lines, e.g. Ca H and K) mentioned in Ch. 1.3.

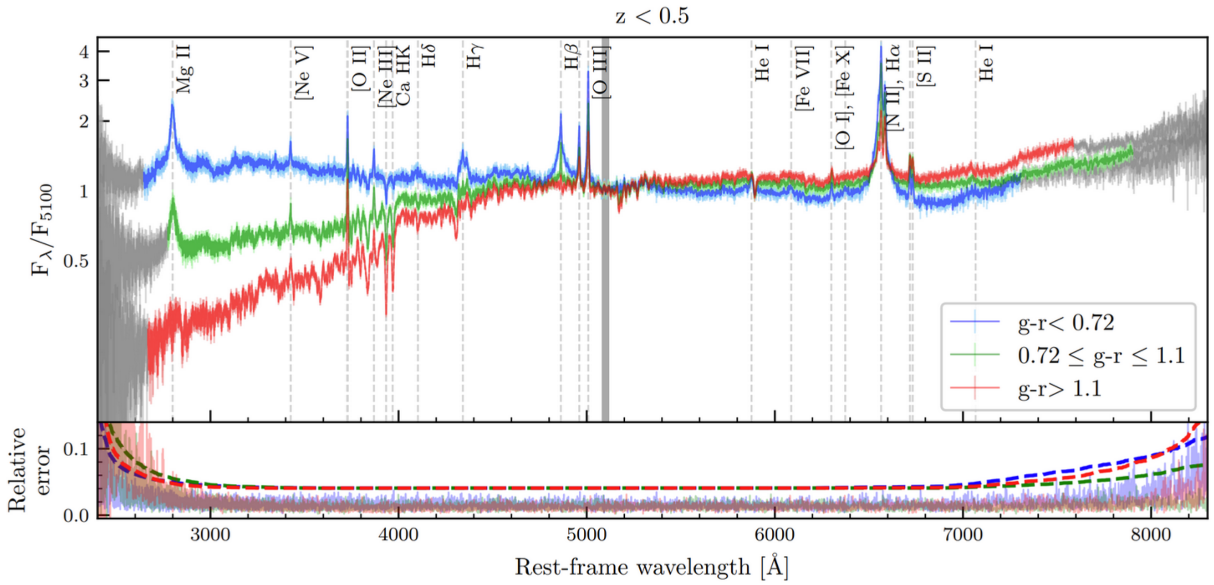


Figure 1.16: Median stacked spectra for X-ray AGN in the eFEDs field at $z < 0.5$, with the flux normalised at 5100Å. They are divided into three $g - r$ colour bins, where for example the redder spectrum is the stack for more dust-obscured AGN with redder optical colours. Some key emission and absorption lines are also marked with dashed vertical lines. Figure taken from Aydar et al. (2025).

The Legacy Survey is used throughout this Thesis to determine host galaxy counterparts for radio and X-ray selected sources and SDSS provides thousands of spectroscopic redshifts. For example, Chapter 2.3.2 elaborates on the details to find host galaxy counterparts in LS9 to the LOFAR-eFEDS radio sources. The photometry from these surveys is also vital to build SEDs for the galaxies in order to determine their physical properties, such as stellar masses and star-formation rates (see e.g. Ch. 5.2).

1.6 Statistical studies of the incidence of AGN

Obtaining a complete census of AGN is key to understanding their accretion history across cosmic time, their feeding–feedback cycle, and the physical conditions that trigger AGN activity. One approach is to explore how AGN, detected using multi-wavelength surveys, populate galaxies of different physical properties. In other words, we can probe the 'incidence' of AGN, or the fraction of galaxies hosting AGN as a function of different galaxy and black hole properties. Such studies require the compilation of large, statistically complete parent samples of galaxies, from which one can define a target sample (e.g. radio or X-ray AGN) to probe their incidence and to characterise the galaxy and AGN properties.

Previous works in this field include those of Aird et al. (2012), Bongiorno et al. (2012), Georgakakis et al. (2017), and Birchall et al. (2022), who selected AGN from deep extragalactic fields to measure the incidence of X-ray AGN as a function of host stellar mass (as a proxy of black hole mass) and specific black hole accretion rate. X-ray surveys allow pure and complete statistical samples of AGN to be compiled as X-rays trace black hole accretion while being less affected by contamination and by dust obscuration, compared to other wavelengths. Importantly, these studies found that the probability of a galaxy hosting an X-ray AGN is described, to first order, by a universal specific black hole accretion rate distribution independent of the host galaxy stellar mass. This lead to the conclusion that the same physical mechanisms are in charge of triggering and fuelling AGN activity in all moderately massive galaxies (e.g. Aird et al. 2012). However, such studies have struggled to push down to the low-mass regime of AGN in dwarf galaxies due to the lack of wide-area, sensitive X-ray surveys and detailed multi-wavelength information for robust characterisation of the hosts.

Additionally, it remains unknown as to whether or not a similar mass-invariant relation can be found for radio AGN as a function of mass-normalised jet power in a given accretion mode. This could be expected from the scale-invariant jet formation theories described in Ch. 1.4. Previous radio AGN studies have found that the incidence of radio AGN is a strongly increasing function of host galaxy stellar mass (Best et al. 2005; Smolčić et al. 2009), populating 'red and dead' galaxies (Best & Heckman 2012). Sabater et al. (2019) has since expanded on this using a larger sample of LOFAR radio AGN in the local Universe ($z < 0.3$) to find that the most massive galaxies ($M_* > 10^{11}\ M_\odot$) are always 'switched on' as radio AGN. Kondapally et al. (2022) further investigated the role of galaxy type on jet powering, by selecting a sample of LERGs (considered to be accreting in a radiatively inefficient mode) in quiescent and star-forming galaxies. They found a significant population of LERGs in bluer star-forming galaxies whose incidence displays a flatter stellar mass dependence compared to quiescent LERGs, implying different fuelling mechanisms at play.

Nevertheless, there have not been studies of the radio AGN incidence as a function of mass-normalised jet power, which would probe the complementary aspect to the X-ray AGN incidence studies as a function of mass-normalised radiative power. These mass-scaled power indicators are vital to find true underlying trends and avoid the known degeneracies of flux-limited survey being unable to discriminate between to high accretion rate (high jet power), small mass black holes and low accretion rate (low jet power), large mass black holes (e.g. Aird et al. 2012). This Thesis aims to fill these gaps in the current literature by building new samples of radio and X-ray AGN, using the recent eFEDS-LOFAR and eRASS:4 surveys, to explore their incidence in the local populations of dwarf to massive galaxies.

1.7 Observational evidence of AGN feedback

As mentioned in Ch. 1.1, key relations between certain black hole and host galaxy properties opened a new avenue to explore the co-evolution of these systems across cosmic time. Namely, the black hole mass was found to scale as $M_{\rm BH} \propto \sigma^{4-5}$, where σ is the stellar velocity dispersion, and $M_{\rm BH} \propto M_{\rm bulge}$, where $M_{\rm bulge}$, is the stellar mass of the bulge (e.g. Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Kormendy & Ho 2013). These scaling relations seem to suggest that the SMBH grows until it is limited by feedback mechanisms which act to expel gas from the bulge potential well, and that this SMBH growth happens in tandem with galaxy growth, through mergers and star formation. In particular, King (2003) and King (2005) show that the power four or five scaling for the $M - \sigma$ relation can naturally be attributed to energy- or momentum-conserving outflows. However, Jahnke & Macciò (2011) point out that the $M_{\rm BH} - M_{\rm bulge}$ relation does not necessarily imply co-evolution as the hierarchical assembly of black hole and stellar mass through galaxy merging can equally reproduce this roughly 1:1 relation.

Regardless, numerous observational studies have confirmed the presence of AGN feedback, establishing it as one of the key advances in our understanding of AGN and galaxy evolution. As shown in Figure 1.3, the main forms of AGN feedback are through jets and outflows. The former showcases one of the most direct pieces of evidence for AGN feedback in action. This is shown in Figure 1.17 where such radio jets have carved out clear cavities in the X-ray emitting intra-cluster medium (ICM; see review by e.g. Hlavacek-Larrondo et al. 2022; Hardcastle & Croston 2020b). A calculation of the energy required to inflate the cavities also gives us a direct indication of the power of the jetted feedback (e.g. Willott et al. 1999; Cavagnolo et al. 2010). This is an important concept that is used throughout Chapters 2, 3 and 4, in particular when assessing, quantitatively, the impact of jetted feedback on the host galaxy or larger-scale dark matter halo. For example, previous works have found that the kinetic power in jets can efficiently offset local cooling flows in galaxy groups and clusters (e.g. Fabian 1994; Peterson et al. 2004; McNamara et al. 2005; Mc-

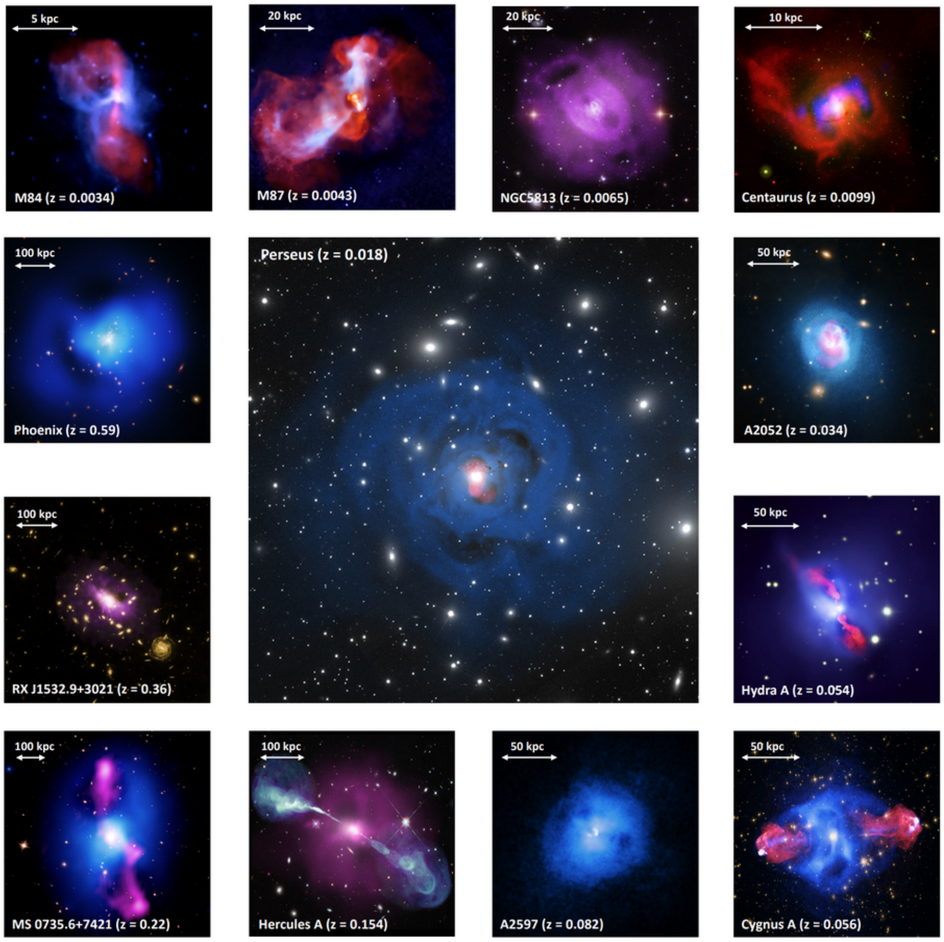


Figure 1.17: Radio jetted feedback shown to carve out cavities in the ICM gas of well-known clusters. X-ray emission is in blue from *Chandra*, radio emission is shown in pink from the *Very Large Array* (VLA) and the background image is from SDSS, illustrating the galaxies. Notably, the famous Perseus cluster (central image) even shows numerous cavities, possibly from past jetted activity. Figure taken from Hlavacek-Larrondo et al. (2022).

Namara & Nulsen 2012), thereby preventing local star-formation. Yet, much of the past literature has been focused on the large, powerful FRI or FRII radio sources, much like those seen in clusters as shown in Figure 1.17. Chapter 4 offers a new perspective on this field by using the high sensitivity LOFAR-eFEDS survey to statistically assess how even low-luminosity, compact radio sources influence their surrounding environments.

Similarly, AGN outflows have been observed in a variety of gas phases, velocities and spatial scales (e.g see reviews by Laha et al. 2021; Harrison & Ramos Almeida 2024). On the largest host-galaxy scales, we find molecular gas outflows, evidenced by broad CO, OH, [C II] line wings, which have relatively slow velocities of $v \sim 100 - 1000$ km s^{-1},

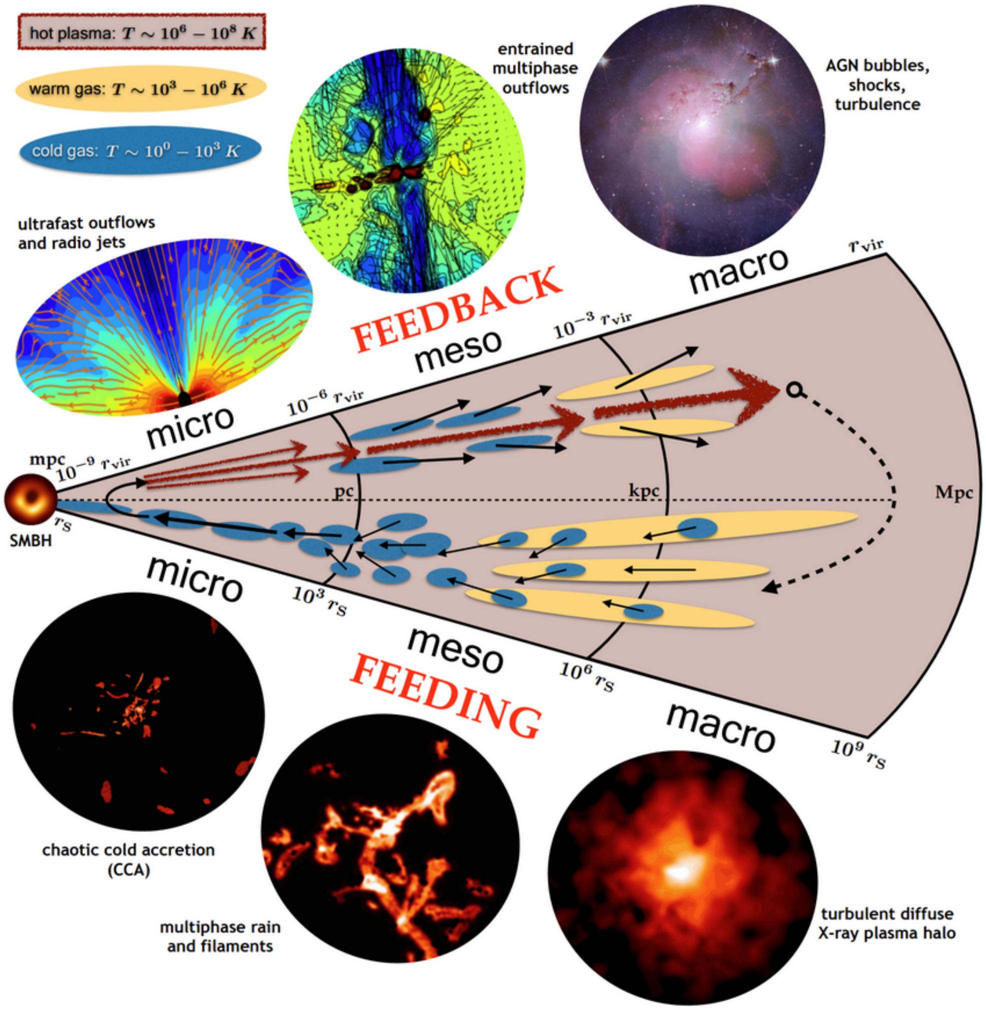


Figure 1.18: The self-regulated multi-phase, multi-scale AGN feeding and feedback cycle. Figure taken from Gaspari et al. (2020).

but show significant mass outflow rates of $\dot{M}_{\rm out} \sim 10 - 1000\ M_\odot\ {\rm yr}^{-1}$ (e.g. Sturm et al. 2011; Cicone et al. 2014, 2018; Ramos Almeida et al. 2022). On NLR scales, we have ionised outflows showing up, for example, in broadened [O III] emission or the blue wings of emission lines such as Hα (e.g. Fiore et al. 2017). On BLR and torus scales, we detect warm absorber winds (e.g. Crenshaw et al. 2003; Mizumoto et al. 2019) from the soft X-ray and/or UV absorption features they imprint in O VII, O VIII and C IV, for example. Finally, at a scale of a few hundred R_g, ultra-fast outflows are observed to reach speeds of $\sim 0.03 - 3\,c$ as found by blue-shifted Fe XXV–XXVI absorption (e.g Tombesi et al. 2010; Gofford et al. 2013; Igo et al. 2020), with more uncertain mass outflow rates ranging on the order of $0.01 - 10 M_\odot\ {\rm yr}^{-1}$. Using these values and those for a slower molecular outflow above, we can estimate the range of kinetic powers via $\dot{E}_{\rm kin} = \frac{1}{2}\dot{M}_{\rm out}v^2$ to reach values of

$\sim 10^{45-46}\,\mathrm{erg\,s^{-1}}$. Given that these outflows are common in AGN (see e.g. Igo et al. 2020) and the lifetime of an AGN is on the order of several $10-100\,\mathrm{Myr}$ (as given by the Soltan argument Soltan 1982), this results in an energy injection that can significantly alter the gas distribution and dynamics within the galaxy.

Overall, AGN feedback is a multi-phase and multi-scale phenomenon that is thought to self-regulate the galaxy growth and the AGN itself. An illustration of how AGN feedback may be tied to AGN feeding, via accretion, is shown in Figure 1.18.

1.8 The necessity of AGN feedback in cosmological simulations

In the early generations of cosmological and semi-analytic (SAMs) models of galaxy formation, it became clear that AGN feedback processes were necessary to prevent the excessive formation of stars in the largest dark matter haloes, producing an overabundance of massive, blue, and star-forming galaxies (e.g. Croton et al. 2006; Bower et al. 2006; Sijacki et al. 2007). This is shown in Figure 1.19, where AGN feedback has to be invoked to reproduce observations (green curve from Moster et al. 2013) of the stellar mass−halo mass relation at the high mass end (e.g. see review by Harrison 2017). Meanwhile at the low-mass end, stellar winds from massive stars and supernova feedback can effectively drive outflows and reheat gas, making it similarly less efficient to form stars. However, these mechanisms become insufficient for the deep gravitational potential wells in massive halos (as shown by the purple curve in Fig. 1.19).

Building on these early insights, cosmological hydrodynamical simulations such as Illustris (Vogelsberger et al. 2014), EAGLE (Schaye et al. 2015), Horizon-AGN (Dubois et al. 2016) and IllustrisTNG (Weinberger et al. 2018), to name a few, started incorporating physically motivated AGN feedback prescriptions to capture the interaction between black holes and their host galaxies. These models intended to explicitly link black hole accretion rates to the energy and momentum injected into the surrounding gas, with the coupling strength and physical effects varying according to the accretion mode (as described in Ch. 1.2), and trace their impact self-consistently within full cosmological hydrodynamical simulations. Including AGN feedback in this way proved crucial to match key observables such as the galaxy stellar mass function, the bi-modality of galaxy colours, and the thermodynamic properties of hot gas in groups and clusters (e.g. Di Matteo et al. 2005; Springel et al. 2005; McCarthy et al. 2010; Puchwein & Springel 2013; Nelson et al. 2018; Bahar et al. 2024).

Interestingly, observers and simulators approach AGN feedback from complementary directions: observers start from the detectable emission and infer the underlying accretion physics, while simulators begin with the physical accretion properties and aim to predict the resulting observable signatures. As a result, differences in terminology and focus can create a gap between the two communities, limiting the extent to which observations can inform simulations or simulation outputs can be directly compared to measured quantities.

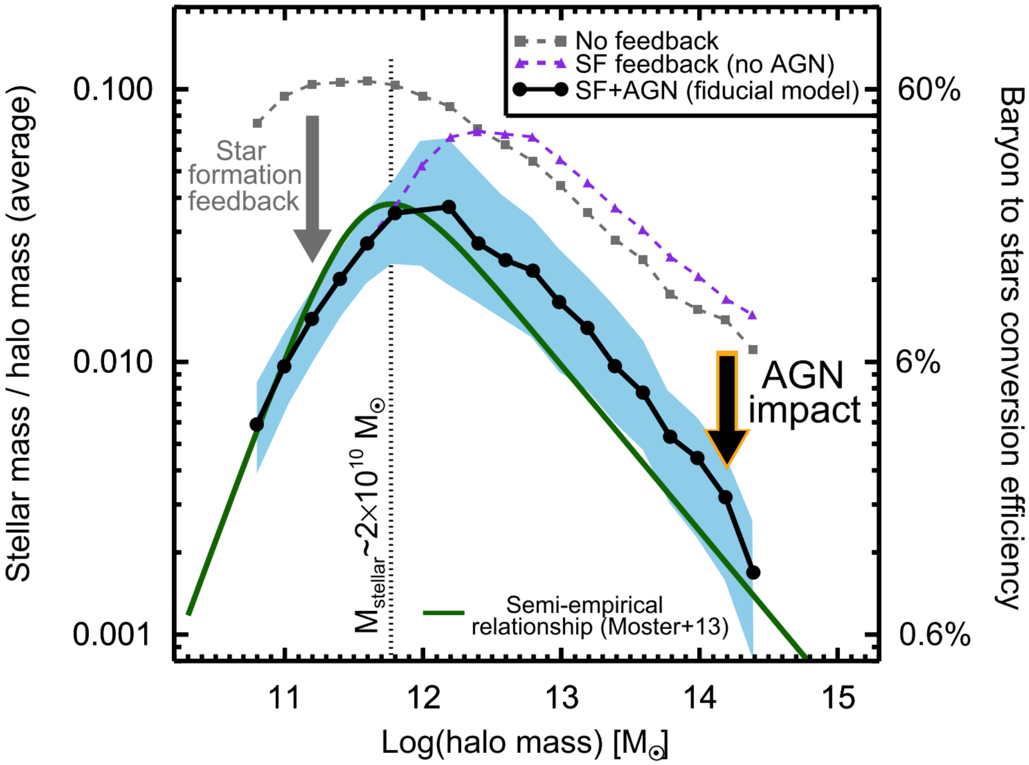


Figure 1.19: The ratio of stellar mass to halo mass as a function of halo mass as probed by observations from Moster et al. (2013) and comparing to simulations with or without AGN and stellar feedback from Somerville et al. (2008). Figure taken from Harrison (2017).

This is what Chapter 4 aims to overcome by rigorously defining all input observables such that the statements on the global energetics of radio AGN kinetic feedback can serve as a benchmark for simulations.

Finally, it is important to bear in mind that the aforementioned cosmological hydrodynamical simulations and SAMs differ greatly in their implementations of their underlying numerical architectures, sub-grid physics, AGN feedback prescriptions and the coupling efficiencies used to transfer feedback energy and momentum to the surrounding medium (e.g Sijacki et al. 2007; Booth & Schaye 2009; Dubois et al. 2012; Costa et al. 2014; Weinberger et al. 2017; Harrison et al. 2018; Costa et al. 2020; Koudmani et al. 2024). This is further complicated by the fact that the scales that must be covered in such simulations range over 14 orders of magnitude (but see recent advancements from the FORGE'd in FIRE simulations; Hopkins et al. 2024). These factors contribute to differing reproduction of observable results, for example in the X-ray luminosity function and $M_{\rm BH} - M_{\odot}$ relation (e.g. Habouzit et al. 2021). Ward et al. (2022) similarly demonstrate that simulations cannot conclusively determine whether the global impacts of AGN feedback are positive (i.e. enhancing star-formation), negative (i.e. inhibiting star-formation) or neutral.

1.9 Early black hole seeding and growth as probed by local analogues

SMBHs in the centres of massive galaxies are now deemed ubiquitous and thought to co-evolve with their host galaxies, as discussed throughout this Chapter. However, at the time of writing, the jury is still out regarding the presence of massive ($M_{\rm BH} \sim 10^{4-7}\ M_\odot$; MBHs) or intermediate-mass ($M_{\rm BH} \sim 10^{2-5}\ M_\odot$; IMBHs) black holes in the centres of smaller galaxies ('dwarfs'; see recent reviews by e.g. Mezcua 2017; Greene et al. 2020).

The answer to this question may hold keys to our understanding of black hole seeding in the early Universe and their subsequent growth across cosmic time. Black holes can be seeded from a variety of processes, such as the death of Population III stars at $z > 15$ leading to $\sim 10^2 - 10^3\ M_\odot$ black holes ('light seeds'; e.g Bond et al. 1984; Madau & Rees 2001; Abel et al. 2002; Bromm & Larson 2004) or the collapse of pristine primordial gas clouds directly into $\sim 10^4 - 10^6\ M_\odot$ black holes ('heavy seeds'; e.g. Loeb & Rasio 1994; Bromm & Loeb 2003; Lodato & Natarajan 2006; Volonteri et al. 2008; Latif et al. 2023). Additionally, recent studies provide growing evidence that black holes can undergo hierarchical growth within high-concentration nuclear star clusters (e.g. Baldassare et al. 2022; Rantala et al. 2024; Partmann et al. 2025; Rantala & Naab 2025; Fujimoto et al. 2025). This process may represent an alternative heavy-seed formation channel, in which stellar-mass black holes undergo unstable runaway growth through tidal captures (e.g. Stone et al. 2017; Natarajan 2021), operating even at later cosmic epochs and thereby blurring the distinction between light and heavy seeds.

Depending on the seeding mechanism, different black hole occupation fractions (BHOF) in low-mass galaxies ($M_* \leq 10^{10} M_\odot$) are expected. Light seeding predicts close to 100% BHOF and the heavy seeding predicts a steeply falling BHOF towards lower masses with around 50% occupation fraction around $\log(M_*/M_\odot) \sim 8 - 9$ (e.g. Ricarte & Natarajan 2018; Burke et al. 2025; Miller et al. 2015, but are also degenerate to the various post-seeding growth channels, see e.g. Chadayammuri et al. 2023.).

Observationally, there are two main approaches to take to tackle this problem: 1) search for high-redshift ($z > 10 - 20$) black hole 'seeds'; and 2) search for the remnants of such seeds in the local universe ($z < 0.2$) that did not grow. The former has recently seen great advancements thanks to observations with the *James Webb Space Telescope* (*JWST*; Gardner et al. 2023) which are pushing the limits to detecting more massive black holes earlier in cosmic time, thereby challenging black hole seeding models as super-Eddington accretion models often need to be invoked to explain the existence of these objects (e.g. a $\sim 10^8 M_\odot$ SMBH observed in the first 470 million years of the universe by Bogdán et al. 2024; Natarajan et al. 2024).

The second approach, and the one adopted in Chapter 5, hinges on the assumption that low-mass galaxies in the local Universe, may be analogous to the first galaxies that formed in the early Universe and can thus be used to test high-redshift black hole growth and seeding models (e.g. Mezcua 2017). In recent years, this field has evolved into a multi-wavelength search and characterisation of black holes in low-mass galaxies, thanks to ever-deeper and ever-wider multi-wavelength surveys (e.g. Greene & Ho 2004, 2007; Nyland et al. 2012; Reines et al. 2013; Sartori et al. 2015; Chilingarian et al. 2018; Kaviraj et al. 2019; Mezcua et al. 2018, 2019; Reines et al. 2020; Eberhard et al. 2024; Pucha et al. 2025). However, as Wasleske & Baldassare (2024) show, different selection methods recover only subsets of the dwarf galaxy population, potentially leading to biased estimates of the BHOF and its lower bound, the 'active fraction', defined as the fraction of low-mass galaxies that contain an accreting AGN.

As mentioned in Ch. 1.6, the X-ray regime offers a relatively pure and complete way to examine the way AGN populate their host galaxies. This is useful, as signatures of black hole seeding and growth should be imprinted in population-level predictions that trace the luminosity and stellar mass distribution of black holes. However, computing the AGN incidence as a function of mass-scaled radiative power in the low-mass regime has thus far been greatly hindered by the lacking sample sizes, especially given the requirement of complete samples to infer statistically reliable X-ray AGN fractions. Some key works in this context are by Birchall et al. (2020), Birchall et al. (2022) and Aird et al. (2018), who compile statistical samples of tens of X-ray AGN in dwarf galaxies to probe the distribution of specific accretion rates. They derive an AGN active fraction between $\sim 0.1-1\%$, similar to other X-ray works using different methods (e.g. Pardo et al. 2016; Mezcua et al. 2018; Pacucci et al. 2021), but struggle to directly connect these results to the high redshift black hole seeding, for three main reasons. Firstly, due to limited sample sizes, past studies have only been able to constrain the incidence distributions over a limited range of specific accretion rate, with relatively large uncertainties, meaning that their derived X-ray AGN active fractions often require extrapolation. Secondly, the X-ray AGN incidence as a function of mass-scaled radiative power traces the black hole growth with respect to the *total galaxy stellar mass*, rather than tracing the underlying Eddington ratio distribution (recall Eq. 1.9). This is particularly important because the scaling relations between stellar mass and black hole mass remain uncertain in the low-mass galaxy regime (e.g. Greene et al. 2010; Kormendy & Ho 2013; Graham & Scott 2015; Reines & Volonteri 2015; van den Bosch 2016; Greene et al. 2020). Thirdly, and arguably most importantly, a strong degeneracy exists between low occupation fractions and high active fractions versus high occupation fractions and low active fractions.

Chapter 5 will significantly expand on these results using the deepest eRASS:4 data to uncover hundreds to thousands of AGN in this new regime of low-mass galaxies, greatly increasing the known sample of such objects and putting the tightest constraints on the specific accretion rate distributions (at least at the high accretion rate end), to date. Even though the eRASS:4 survey depth ultimately does not allow for conclusive statements to be made on the BHOF in the low-mass regime, low-mass galaxies are still the most likely to have retained the imprint of early black hole seeding. Therefore, a detailed study of their accretion mechanisms is vital to begin to disentangle the degeneracy between activation and occupation fractions. Chapter 5 also touches on the relevance of AGN feedback, thus far only discussed in the context of massive galaxies (Chs. 1.7 and 1.8), in the low-mass galaxy regime.

1.10 Outline of Thesis

With this Thesis we aim to better understand the multi-wavelength emission of AGN, particularly in the radio and X-ray wavelength regimes, to constrain feeding and feedback mechanisms by examining how they populate their host galaxies across the mass scale.

Chapter 2 compiles a new sample of radio and X-ray AGN from the LOFAR-eFEDS field after careful consideration of contaminating sources of emission and controlling the completeness of the parent galaxy sample. It further explores the extent to which the radiatively efficient and inefficient accretion modes of the AGN in this sample can be distinguished.

Chapter 3 computes the incidence of X-ray and radio AGN as a function of mass-scaled power indicators. The radio study is performed here for the first time (and the X-ray, for the first time using eROSITA data). Thanks to the detailed parent galaxy and AGN characterisation from Chapter 2, the radio AGN incidence can be investigated as a function of stellar mass, jet power and radio morphology and the results can be placed into a wider context by looking at the disk–jet coupling in these systems.

Chapter 4 builds on the interesting differences found in Ch. 3 regarding the way radio AGN of different radio morphologies, living in galaxies of different masses, distribute their power among the galaxy population. The average jetted energy of the massive galaxy population is computed and compared to the galaxy and halo binding energies, as well as the halo thermal energy, to quantitatively measure the global impact of radio AGN kinetic feedback in the local Universe.

Chapter 5 uses the deepest eRASS:4 data to assemble the largest sample of X-ray AGN in low-mass galaxies to date and studies their incidence as a function of specific accretion rate. It derives estimates of the cumulative AGN fraction and comments on how this can inform us about black hole seeding and growth across cosmic time.

Finally, Chapter 6 concludes upon the main findings of this Thesis and gives an outlook on the future developments in the field.

A standard flat cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$ is used throughout and all magnitudes are AB magnitudes corrected for galactic extinction.

Chapter 2

The LOFAR-eFEDS survey of radio and X-ray AGN

In this Chapter, we exploit the rich multi-wavelength coverage of the eROSITA eFEDS field to build a LOFAR - eFEDS value added catalogue. We later use it in Chapters 3 and 4 to compute the incidence of radio and X-ray AGN as a function of mass-scaled power indicators and derive statements about the global energetics of AGN feedback in the local universe. To achieve these aims, careful considerations must be made to create a pure AGN sample by removing contaminants, characterise the stellar content and star formation rates of the host galaxies, and apply appropriate stellar mass and luminosity completeness cuts.

We first discuss the eFEDS X-ray AGN, for which the optical counterparts and X-ray spectral characterisation has been presented in Salvato et al. (2022a) and Liu et al. (2022d), respectively. We define a sample of 'G9 X-ray AGN', which are those X-ray AGN with counterparts in the 9hr field of the Galaxy and Mass Assembly Survey (GAMA09), in the redshift range $z < 0.4$, and analyse their host galaxy properties. Then we describe the creation of the LOFAR - eFEDS value added catalogue. This includes characterising radio morphologies, obtaining optical counterparts from LS9 and defining a new sample of radio AGN in the GAMA09 field ('G9 radio AGN'), according to a radio-excess relative to their host SFR. Thanks to the well-characterised, spectroscopic GAMA09 galaxies, we can form a highly complete parent galaxy sample from which we derive incidence distributions to learn more about the triggering, fuelling and jet powering of AGN, as presented in Chapters 3 and 4. We conclude this Chapter by investigating the properties of our X-ray and radio detected sources, including AGN, to understand the dominant accretion mode or emission mechanism present in these systems. To do so, we define various mass-scaled power indicators, which will be particularly relevant in Chapters 3 – 5.

We have made the LOFAR - eFEDS catalogue publicly available and encourage the community to use it for their diverse scientific endeavours[1]. The contents of this Chapter are published in Sections 1 – 3 and Appendix A, B, C of Igo et al. (2024).

---

[1] The source catalogue is available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via `http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/` or on the LOFAR Surveys DR website: `https://lofar-surveys.org/efeds.html`

2.1 Overview of data and catalogues

As explained in Chapter 1, the eFEDS field was deliberately chosen to overlap with a multitude of surveys, thereby mapping the universe, in a window covering ∼140 deg^2 of equatorial sky, from radio to optical to X-ray wavelengths. The eFEDS survey, DESI Legacy Imaging Survey, and LOFAR survey in X-rays, optical/infrared and radio respectively, have already been described in Chapter 1.5.

However, for the detailed characterisation of the radio and X-ray sources in the LOFAR-eFEDS field, we further make use of the Galaxy and Mass Assembly (GAMA) DR4 survey (Driver et al. 2022), specifically the 9hr field (hereafter: GAMA09). This is a spectroscopic galaxy survey with complementary broadband coverage from the far-UV (∼1500 Å) to the far-infrared (∼500 μm), which allows for the creation of wide SEDs for individual galaxies and the determination of galaxy properties through SED fitting (Robotham et al. 2020; Bellstedt et al. 2020, 2021). These include stellar mass and star formation rate (SFR) estimates from stellar population synthesis modelling of SEDs, using Bruzual & Charlot (2003) stellar evolution models, taking a Chabrier (2003) initial mass function (IMF) and Calzetti et al. (2000) dust curves. The science-ready source catalogue (`gkvScienceCatv02`) and additional host-galaxy properties from SED fitting (`gkvProSpectv02`) are available to download from the fourth GAMA data release website[2].

Figure 2.1 shows a sky plot of these four surveys: eFEDS (blue shaded), LOFAR (red outline), Legacy Survey (grey shaded) and the GAMA09 (purple hatched). The accessibility of these catalogues is the following. The eFEDS X-ray catalogue, along with its optical counterparts and X-ray spectroscopy results are available on the eROSITA Early Data Release website[3].The full field 144MHz LOFAR-eFEDS source catalogue containing 45,207 entries, produced using PyBDSF at a resolution of $8'' \times 9''$ is made available on the LOFAR Surveys Data Releases website[4]. The columns are the same as in the LoTSS DR1 from Shimwell et al. (2019), but more details can be found in the PyBDSF documentation[5]. The Legacy Survey data and documentation is found on its homepage[6] and GAMA09 catalogue is described above.

The LOFAR-eFEDS value-added catalogue, which underpins the work described in Chapters 2−4 is provided for the 36,631 LOFAR sources in the eFEDS region where the X-ray exposure time exceeds 500s (see Fig. 2.1). Included in this VAC is the radio morphology classification for all sources; optical host galaxy identification using Legacy DR9 data for those sources for which a counterpart is found; host-galaxy properties and spectroscopic redshifts for the subset of sources with optical counterparts in the GAMA09 field; and several science and quality flags that are developed and explained in the subsequent sections of this Chapter. Importantly, the calculation of the incidence of radio and X-ray AGN ('target sample') in Chapter 3 is done only on this 'parent sample' of galaxies within the

---

[2] `https://www.gama-survey.org/dr4/`

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

[4] `https://lofar-surveys.org/efeds.html`

[5] `https://pybdsf.readthedocs.io/en/latest/write_catalog.html`

[6] `https://www.legacysurvey.org/`

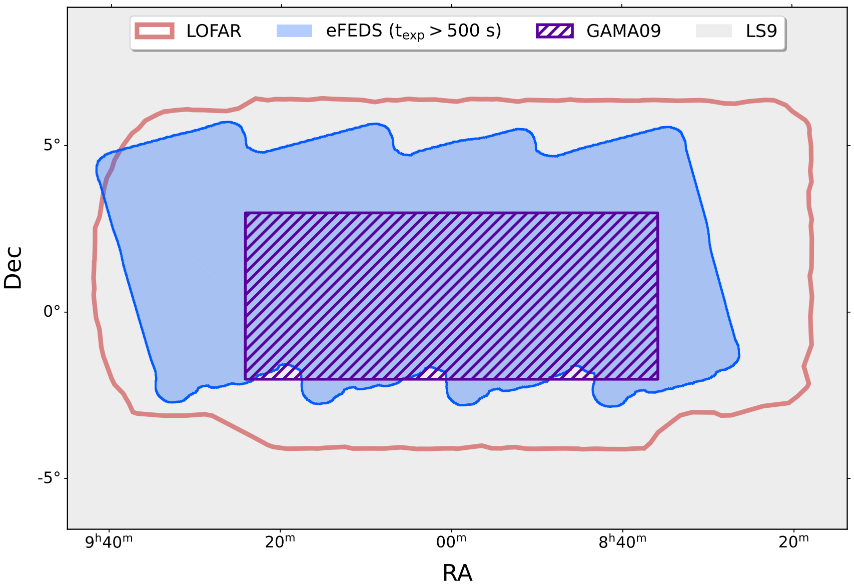


Figure 2.1: Sky area coverage of the radio (LOFAR; light red), X-ray (eROSITA eFEDS; blue) and optical/UV (GAMA09; purple, LS9; grey) surveys in this equatorial field. We note that the eFEDS area has been cut to the region where the vignetted exposure time exceeds 500 s.

GAMA09 field. In particular, we consider only the GAMA09 subsample[7] with quality flag `SC` ≥ 6 and $0 < z < 0.4$, resulting in a cleaned catalogue with 90% completeness limit for $r < 19.77$ (more detailed discussion on completeness follows in Chapter 2.2.3).

We summarise the different catalogues and subsamples of this work in Table 2.1, which serves as a reference to the reader throughout this Chapter and explains the detailed characterisation of X-ray and radio AGN sources (Ch. 2.2 and 2.3). Additionally, the complete column descriptions and notes about the LOFAR - eFEDS value added catalogue can be found in Chapter 2.6.

2.2 Characterisation of the X-ray AGN sample

This section describes the X-ray AGN sample, their optical host galaxies and their properties, along with the considerations taken to control for stellar mass and X-ray luminosity completeness. For a summary of the number of sources at each step, see Table 2.1.

2.2.1 Optical counterparts of the X-ray sources

The identification of the optical counterparts for the X-ray sources is crucial to later classify the host galaxy properties, including stellar mass and star-formation rate. This is done

---

[7] `http://www.gama-survey.org/dr4/schema/table.php?id=684`

Table 2.1: Summary table of the number of sources in each multi-wavelength catalogue and the different subsets explained in Chapter 2.

Catalogue	**Number**	**Comments**
eFEDS Main	27,021*	* = cut to eFEDS MOC $t_{\mathrm{exp}} > 500$s
LS DR9	11,255,466*	duplicate ("DUP") sources removed
LOFAR	36,631*	24,613 compact, 12,018 complex
GAMA09	48,190	`SC` ≥ 6; $z < 0.4$; 21,462 mass-complete
eFEDS-LS9	20,696	Salvato et al. (2022a); `p_any` > 0.035, `CTP_quality` > 2
eFEDS-LS9-GAMA09	584	
LOFAR-LS9	22,759[a]	`p_any` > 0.06, S/N > 5; 16130 compact, 6629 complex.
LOFAR-LS9-GAMA09	2,619[b]	1901 compact, 718 complex
GAMA09 (G9) X-ray AGN	523	325 mass-complete
G9 X-ray AGN in Quiescent gal.	147	124/147 mass-complete
G9 X-ray AGN in Star-forming gal.	376	201/376 mass-complete
G9 Radio AGN	764	682 mass-complete (404/445 compact; 278/319 complex)
G9 Radio AGN in Quiescent gal.	646	595/646 mass-complete (354/385 compact; 241/261 complex)
G9 Radio AGN in Star-forming gal.	118	87/118 mass-complete (50/60 compact; 37/58 complex)
G9 Radio + X-ray sources	121	74 (32) /92 mass-complete X-ray (radio) AGN; 24 Radio + X-ray AGN

[a] The LOFAR-LS9 NWAY match yields 22,754 sources, to which 5 (out of 6) large FRII sources are added manually (the remaining large FRII source with LOFAR Source id 8153 was already present among the matched sources, albeit the optical CTP had to be corrected). [b] The mass-complete sources were visually inspected to confirm correct counterpart association and classify radio morphology. Two radio sources (LOFAR Source id: 10347, 27051) were found to be consistent with noise fluctuations of the background and are excluded from this point on (see more details in Appendix 2.3.2).

using a Bayesian cross-matching algorithm called NWAY (Salvato et al. 2018), which uses not only source sky density, distance priors and positional accuracy, but also additional priors based on observable characteristics (e.g. magnitudes, colours). Notably, it provides a best match flag (`match_flag` = 1), a probability for the match being the correct one (`p_i`) and a probability of the source in question having any counterpart at all in the search region (`p_any`).

The counterpart identification of the eFEDS X-ray sources has already been presented by Salvato et al. (2022a), who used external pre-constructed priors, trained on X-ray sources with secure counterparts from Legacy Survey DR8 (Dey et al. 2019) in 3XMM and Chandra catalogues (adjusted to have 'eFEDS-like' source properties). An updated version of their catalogue has been used in this work, where the eROSITA positional error (`RADEC_ERR`) was divided by $\sqrt{2}$ (1-dimensional positional error). This new version (V18) is consistent with the originally released V17 catalogue, with minimal (5%) change in the counterparts, of which only 0.3% have `CTP_quality`>2 (sources with secure counterparts). There is also improved counterpart association as most of the original sources with `CTP_quality`=2 (i.e. with more than one possible counterpart) now have a secure and unique match (Saxena et al, in prep.).

Since the release of the counterpart catalogue in Salvato et al. (2022a), a new data release of the DESI Legacy Survey (DR9), with improved flux calibration and source detection near bright sources, also became available. Therefore, we use the DR9 catalogue to update the optical photometry for the eFEDS X-ray sources by performing a simple $1''$ positional cross-match between Legacy Survey DR8 and DR9. We do not recompute any new counterparts and we note that all original DR8 optical counterparts to the eFEDS sources have a match within $1''$ to a DR9 source. Considering the above, we proceed to apply the same thresholds of `p_any`>0.035 and `CTP_quality`>2, as in Salvato et al. (2022a).

Liu et al. (2022b) provides X-ray spectroscopic results for all eFEDS sources. The absorbed power law models are used to calculate intrinsic $2-10$ keV luminosities (see below). An updated version of this X-ray spectroscopic catalogue using the improved spectroscopic redshifts from SDSS-V DR18 (Almeida et al. 2023), keeping all else equal, is adopted for this Chapter.

2.2.2 X-ray AGN among the GAMA09 galaxies

As the parent sample is the GAMA09 galaxies, the X-ray AGN catalogue is necessarily also limited to the GAMA09 region. GAMA09 counterparts to the X-ray sources in the field are identified using a simple $2''$ (to account for the fibre sizes) positional match between the LS9 optical counterpart coordinates and those of the GAMA09 galaxies. Five sources (eROIDs: 584, 1730, 6498, 9305, 14520) with discrepant redshifts between GAMA and SDSS drop out of the sample as their redshifts are identified, via robust and systematic visual inspection of the SDSS spectra, to exceed the upper bound ($z < 0.4$) of the sample. This leaves a total of 584 X-ray sources associated with GAMA09 galaxies.

Not all of these X-ray sources are AGN, because at low luminosities the sample starts to be dominated by the collective (unresolved) X-ray emission from XRBs and emission from hot diffuse gas within the hosts. The X-ray luminosity of the sources is calculated in the standard way using the absorption corrected flux in the rest-frame $0.5-2$ keV from the

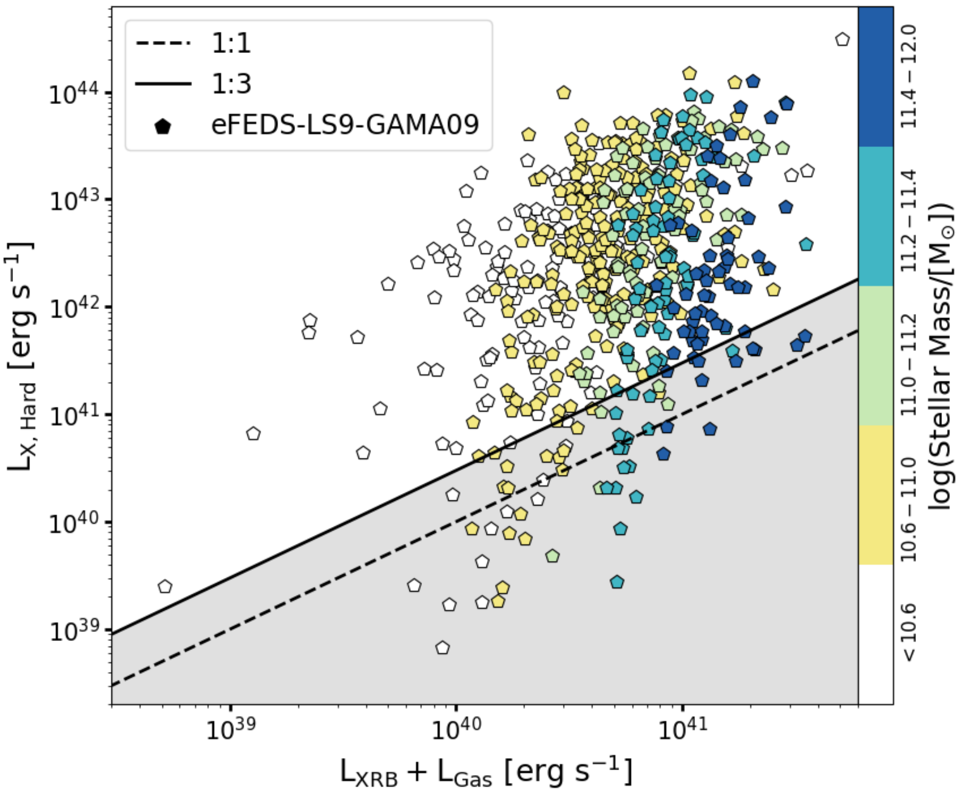


Figure 2.2: Intrinsic $2-10$ keV luminosity of each eFEDS X-ray source versus the X-ray luminosity expected from X-ray binaries and hot gas of the GAMA09 matched galaxies, colour coded according to the stellar mass. Sources lying above the 1:3 solid black line have X-ray emission securely dominated by AGN processes and constitute our X-ray AGN sample; sources in the shaded area are compatible with non-AGN emission processes, and excluded from the analysis.

work of Liu et al. (2022b) (with the updated redshifts, as mentioned above). These were then converted to hard $2-10$ keV luminosities using the modelled photon index. To select out the X-ray AGN, the relations from Lehmer et al. (2016), Eq. 2.1 below, and Mineo et al. (2012), Eq. 2.2 below, are used to estimate the corresponding X-ray emission from X-ray binaries and hot gas (in erg s^{-1}), respectively, for a given M_* and SFR.

$$L_{\mathrm{XRB}} = \alpha_0(1+z)^{\gamma}\ M_*/[M_\odot] + \beta_0(1+z)^{\delta}\ \mathrm{SFR}/[M_\odot\ yr^{-1}]. \tag{2.1}$$

The parameters have the following values for $2-10$ keV X-ray luminosity: $\log_{10}(\alpha_0) = 29.37 \pm 0.15$, $\gamma = 2.03 \pm 0.60$, $\log_{10}(\beta_0) = 39.28 \pm 0.03$ and $\delta = 1.31 \pm 0.13$ (from Sect. 6.3.2. of Lehmer et al. 2016).

$$L_{\mathrm{Gas}} = (8.3 \pm 0.1) \times 10^{38}\ \mathrm{SFR}/[M_\odot\ yr^{-1}]. \tag{2.2}$$

Figure 2.2 shows the intrinsic $2-10$ keV luminosity of each eFEDS X-ray source versus the X-ray luminosity expected from XRBs and hot gas, colour coded by the host galaxy

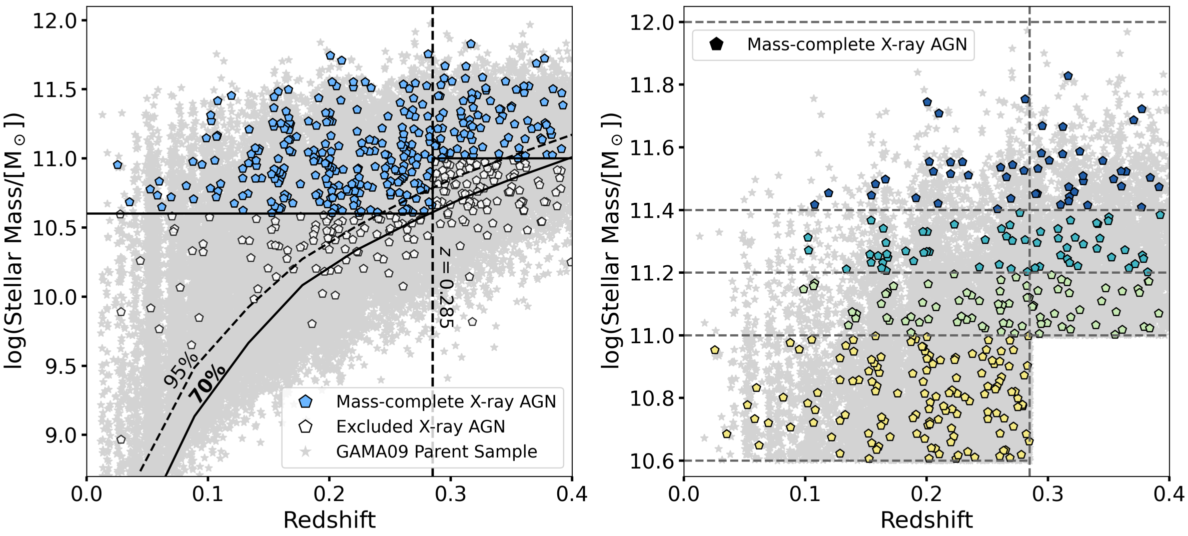


Figure 2.3: Stellar mass versus redshift distribution of the galaxies in GAMA09 (grey points) and of the X-ray AGN (blue filled pentagons). A vertical line divides the sample into two redshift bins. Completeness curves (70%, 95% with solid, dashed black lines, respectively) and horizontal thresholds are used to exclude sources incomplete in stellar mass (unfilled markers). A zoom-in of the mass-complete sample, split into four stellar mass bins, is presented in the right panel.

stellar mass (explained in the next section). Black dashed and solid lines mark the 1:1 and 1:3 levels, respectively. The 523/584 sources (∼90%) that lie above the grey-shaded region defined by the 1:3 line, are the sources in which the X-ray emission is dominated by AGN processes and is henceforth referred to as the X-ray AGN among the GAMA09 galaxies, or ‘G9 X-ray AGN’ (see Table 2.1). This AGN sample is free of `CLUSTER_CLASS`=5 sources, meaning that the AGN X-ray luminosities are not biased by additional emission potentially coming from hot cluster gas (`CLUSTER_CLASS`=5 sources are defined as those with unreliable optical counterparts, likely lying in regions of identified overdensities of red passive galaxies, with their X-ray emission most likely associated with a cluster, see Salvato et al. 2022a, and details therein).

2.2.3 Stellar mass and X-ray luminosity complete X-ray AGN samples

Figure 2.3 (left) shows the stellar mass versus redshift distribution of the GAMA09 parent sample (grey points) and of the X-ray AGN among them (blue pentagons). The stellar masses are taken from the `StellarMass_50` GAMA catalogue entry, the median of the posterior distribution from the Bayesian SED fitting (with Markov Chain Monte Carlo;

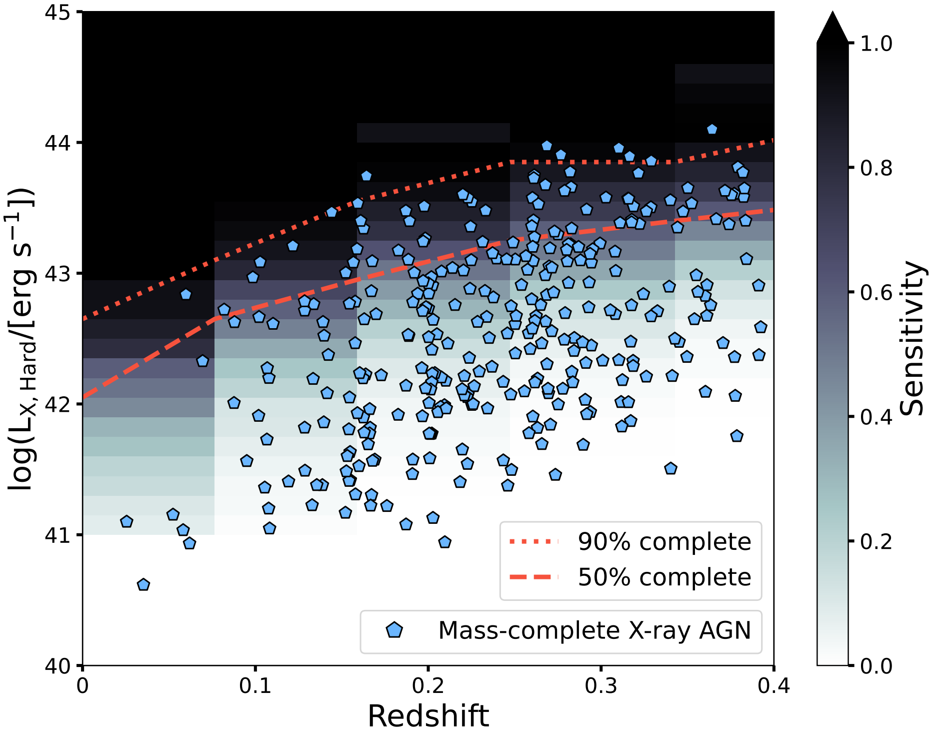


Figure 2.4: Intrinsic $2-10$ keV X-ray luminosity versus redshift of the G9 X-ray AGN (blue pentagons). A sensitivity grid, for both obscured and unobscured sources, is plotted in the background from simulations done by Liu et al. (2022e) and used to compute the 50% (orange, dashed) and 90% (orange, dotted) X-ray luminosity completeness limits, respectively.

MCMC). Although no AGN component was used in the SED fitting to determine the host galaxy stellar mass, we show in Section 3.5 that these measurements are robust and in agreement with later works that do account for AGN (Thorne et al. 2022; Aihara et al. 2018; Li et al. 2023). A vertical line at $z = 0.285$ divides the sample equally into two (low and high) redshift bins (see Ch. 3.1).

As this investigation deals with fractions of galaxies hosting AGN, it is vital to ensure that these samples are complete in both stellar mass and AGN luminosity. Firstly, the stellar mass completeness limits are calculated using the limiting stellar mass method, $M_{*,\mathrm{lim}}$ (e.g. Pozzetti et al. 2010; Moustakas et al. 2013; Mountrichas et al. 2022), shown in Eq. 2.3, where $M_{*,\mathrm{lim}}$ is the stellar mass a given galaxy would have if its r-band magnitude (r_{mag}) was equal to the limiting r-band magnitude of the survey ($r_{\mathrm{lim}} = 19.8$):

$$\log_{10} M_{*,\mathrm{lim}} = \log_{10} M_{*} + 0.4\ (r_{\mathrm{mag}} - r_{\mathrm{lim}}). \tag{2.3}$$

Then, for each redshift interval ($\Delta z = 0.04$), the cumulative distribution of $M_{*,\mathrm{lim}}$ was used to calculate the 70% (solid, black) and 95% (dashed, black) completeness limits and plot this against the maximum redshift in each given interval. As shown in Fig. 2.3 (left), the completeness function is rather steep compared to the change in stellar mass value, and therefore the 70% limit is used for this work, in order to maximise source numbers.

Solid black horizontal lines at $\log(M_*/M_\odot) = 10.6$ and 11.0 mark the stellar mass completeness limits for the low and high redshift bins, respectively and the white-filled markers are the X-ray AGN which are excluded as a result of this cut. Overall, there are 325 X-ray AGN and 21,462 GAMA09 galaxies included in these ‘mass-complete’ (to 70%), volume-limited samples in the redshift range $0 < z < 0.4$. Therefore, the total percentage of these galaxies hosting X-ray AGN detected by eROSITA is about 1.5%. Fig. 2.3 (right) shows a zoom-in of this ‘mass-complete’ sample, splitting up the data into four mass bins (yellow, green, teal, blue), which is elaborated upon in Ch. 3.1.

Figure 2.4 shows the distribution of intrinsic $2 - 10$ keV X-ray luminosities ($L_{\mathrm{X,Hard}}$ in erg s^{-1}) versus redshift for the G9 X-ray AGN (blue pentagons). An X-ray luminosity sensitivity grid from detailed simulations (Liu et al. 2022e, their Fig. 8 and Section 4.1) is also plotted in the background. X-ray sensitivity functions are a complex combination of redshift, absorbing column density (N_H), spectral shapes and k-correction factor dependent parameters. This is rigorously taken into account, as the mock eFEDS AGN catalogue (Comparat et al. 2019) used for these simulations is highly representative of the real eFEDS data. Thus, these parameters, along with any correlation of N_H with M_* (Buchner et al. 2017) for example, have been folded into the X-ray luminosity completeness curves shown on Fig. 2.4. They are valid for the entire N_H distribution of the sample, including both unobscured and obscured sources. The fact that only a few lie above the orange dashed, 50% and orange dotted, 95%, limits, is a result of the soft X-ray response of eROSITA (Predehl et al. 2021). This is because gas preferentially absorbs softer X-ray photons, leading to sources with higher obscuration being more likely to be undetected (see Fig. 3.3 for how this changes when using the purely unobscured sensitivity grid).

In order to compare with past work dealing with X-ray AGN incidences using a hard X-ray band selection (e.g. *XMM-Newton*), these full X-ray completeness correction functions can be used to implement a weighting per bin in stellar mass, redshift, luminosity and λ_{Edd} when computing the incidences (see details in Ch. 3.1).

In Chapter 3.2.1 the purely unobscured (soft X-ray) eFEDS sensitivity functions are presented and applied to the results, in order to show the impacts of obscuration on the measured X-ray AGN incidence.

2.2.4 Host galaxy properties of X-ray AGN

Having defined a clean sample of G9 X-ray AGN and shown the stellar mass and X-ray luminosity distributions as functions of redshift, this section discusses the properties of the host galaxies of these AGN.

AGN with star-forming host galaxies tend to scatter around the so-called main sequence (MS) of star forming galaxies, a well-studied relation in the literature (e.g. Noeske et al. 2007; Speagle et al. 2014; Popesso et al. 2023b, and references therein). There are variances in derived MS relations, especially at the high stellar mass end, which is why two well-founded, yet analytically different ones, namely those presented in Speagle et al. (2014)

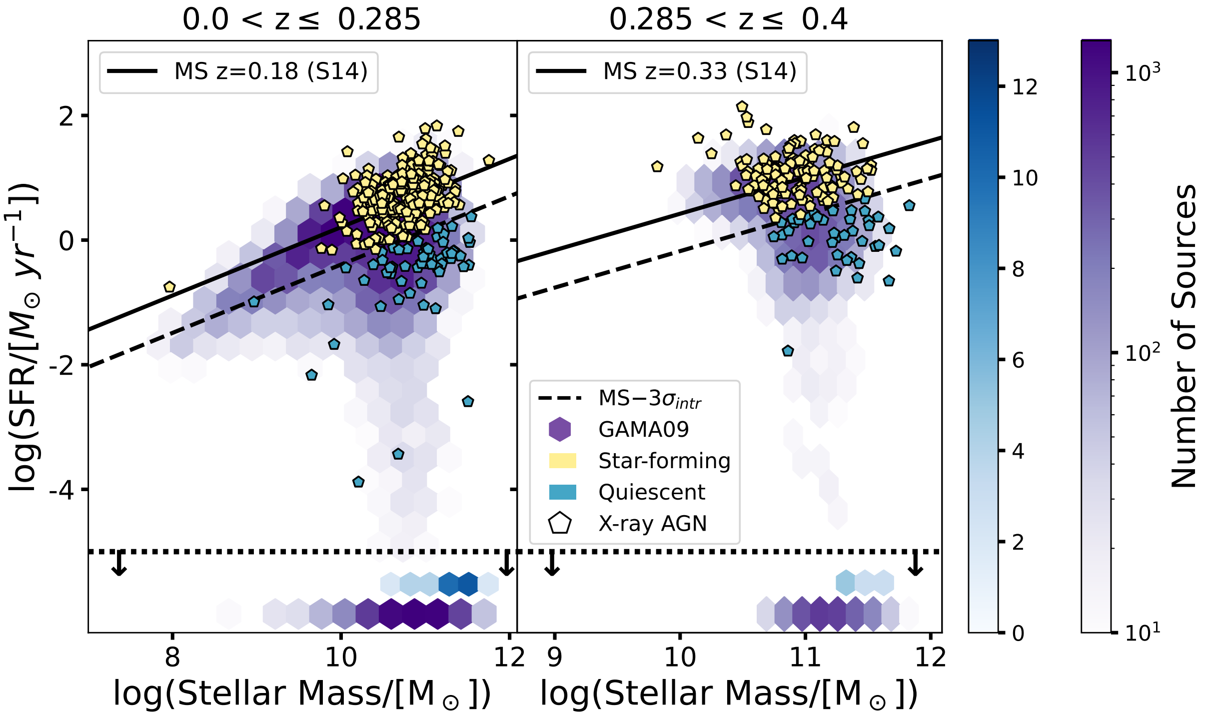


Figure 2.5: SFR versus stellar mass for the GAMA09 galaxy sample (purple hexbins) and for the X-ray AGN (pentagons). The solid black line in each redshift panel marks the star-forming galaxy main sequence from (S14; Speagle et al. 2014). Sources 3σ below this line (black, dashed) are considered to be quiescent galaxies (blue), otherwise they are classified as star-forming (yellow). Quiescent sources below $\log(\mathrm{SFR}/[\mathrm{M}_\odot\ \mathrm{yr}^{-1}]) = -5$ are marked as upper limits and their distributions are shown in the bottom of each panel.

and Popesso et al. (2023b), were tested during the radio and X-ray incidence analysis. Ultimately, both produced similar results and thus the simplest of the two was chosen, for which the best fit MS relation is Eq. 28 from Speagle et al. (2014) with intrinsic scatter $\sigma_{\mathrm{intr}} = 0.2$ (see also their Fig. 8 for a visual representation of the relation).

Figure 2.5 shows the stellar mass versus star-formation rate (`SFR_50`) of the GAMA09 parent sample (purple hexbins), X-ray AGN (yellow, blue) and this MS relation in black for the two redshift bins introduced above, along with a dashed line marking SFRs $3\sigma_{\mathrm{intr}}$ below the MS. The MS is calculated using the mean redshifts of the GAMA09 sample within that redshift bin. The dashed line is used as a demarcation between the quiescent (blue) and star-forming (yellow) X-ray AGN. All sources with $\log(\mathrm{SFR}/[\mathrm{M}_\odot\ \mathrm{yr}^{-1}]) \leq -5$ (below the black dotted line) are marked as upper limits to indicate their quenched nature (Bellstedt et al. 2020). Such low SFRs are possible because a skewed Normal parameterisation is used to fit the star formation history of each GAMA galaxy, from which the SFR is obtained by averaging over the past 100 Myr. Therefore, galaxies with SFRs peaking in the early universe could have $\log(\mathrm{SFR}/[\mathrm{M}_\odot\ \mathrm{yr}^{-1}]) \leq -5$ at present day.

Their stellar mass distributions are plotted as blue (X-ray AGN) and purple (GAMA09 galaxies) hexbins, shifted to an arbitrary low SFR value to avoid overlap. In the low (high) redshift bin, there are a total of 33 (11) X-ray AGN and 5453 (2121) GAMA09 galaxies, which have $\log(\mathrm{SFR}/[\mathrm{M}_\odot\ \mathrm{yr}^{-1}]) \leq -5$

2.3 Characterisation of the radio AGN sample

Having presented the X-ray AGN in the G9 field, we now move to the analysis of the radio AGN sample. We first discuss the distinction between compact and complex radio morphologies among radio sources; then proceed with the identification of the optical counterparts to the radio sources using LS9 and GAMA09; then to the definition of radio AGN as 'radio-excess' sources; and finally to the classification of the optical host galaxies into quiescent and star-forming. For a summary of the number of sources at each step, see Table 2.1.

2.3.1 Compact versus complex radio morphology

Using the LOFAR catalogue consisting of the 36,631 radio sources in the eFEDS $t_{\mathrm{exp}} > 500$s region, the first step is to distinguish between radio sources with different morphologies (compact versus complex/extended). This is because they may be governed by different physical processes, either local to the source or on larger scales. For example, several recent studies are investigating the effects of differing accretion modes (see Ch. 2.4 and Ch. 3.5.3) and small- or large-scale environment (see Ch. 3.5.2) on the radio properties of AGN, such as jet power and morphology (see also e.g. Best 2004; Ching et al. 2017; Shabala 2018; Gaspari et al. 2020; Mingo et al. 2022). Radio sources with markedly different morphologies may also require different cross-identification procedures, so they need to be classified first.

Similarly to Williams et al. (2019), we define a set of four criteria that must be fulfilled for a LOFAR source to be classified as 'compact'. Firstly, we consider the fact that perfect 'compact' (point-like, i.e. unresolved) sources have a ratio of the total integrated flux density to peak flux ($R = F_{\mathrm{Tot}}/F_{\mathrm{Peak}}$) equal to unity and reside completely within the size of the restoring beam (Shimwell et al. 2022). Considering that calibration is not perfect, we fit a Gaussian to the distribution of $F_{\mathrm{Tot}}/F_{\mathrm{Peak}}$ to determine the correct threshold to isolate compact galaxies; see Fig. 2.6, left. Compact sources are defined as those below the threshold marked by the vertical black dashed line at $R < 3.6$ (8σ).

Secondly, compact emitters tend to have smaller sizes, or more specifically, projected extents on the sky, measured for example by the full width half maximum (FWHM) of the major axis (`Maj`) of the source. The right panel of Fig. 2.6 shows the distribution of major axes fit by a Gaussian for the LOFAR sample. Compact sources are defined as those below the threshold of `Maj`$< 19.1''$ (6σ; vertical black dashed line).

The third criterion is that a compact source must be fit with only a single Gaussian by PyBDSF, that is `S_Code`=S. This excludes those sources fit with multiple Gaussians,

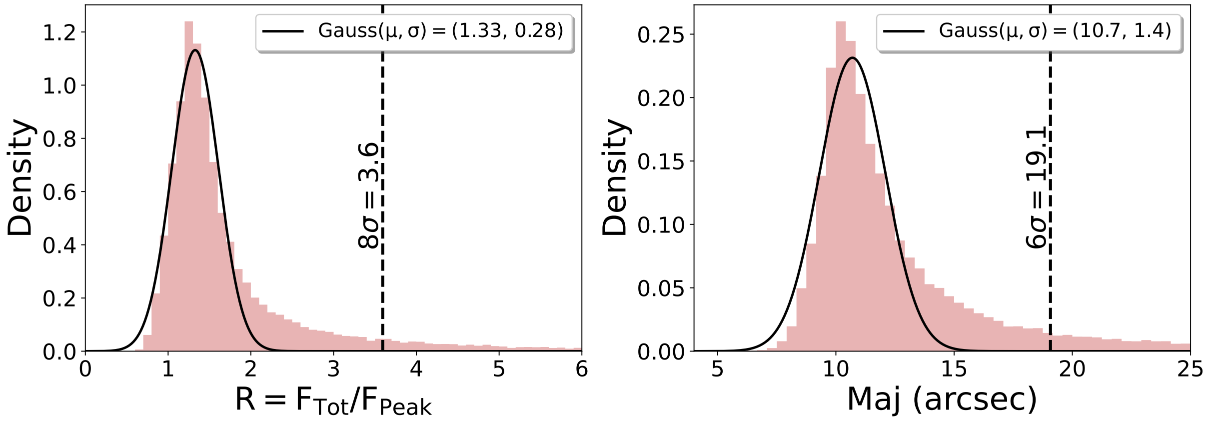


Figure 2.6: Histograms showing the flux ratio (total to peak flux; left) and major axis (right) distributions for the LOFAR sample of 36,631 sources in the eFEDS field, each fit by a Gaussian (black curves) to determine the thresholds for being a compact radio emitter. These are shown as black dashed vertical lines (see text for more details).

`S_Code`=M, or sources fit with a single Gaussian but being located in the same island as other radio sources, `S_Code`=C (Mohan & Rafferty 2015).

Lastly, compact sources must be in an isolated region without any other catalogued LOFAR sources (no nearest neighbours) within 45″. This is to remove cases where far-away lobes/hotspots, associated with the same host galaxy, are erroneously catalogued as two distinct compact radio sources; or the case of dense cluster regions with multiple nearby radio emitters.

We consider that all four criteria have to be simultaneously fulfilled in order for a source to be considered compact, having `LOFAR_compact_flag` set to `True` in the catalogue (see Table 2.2). As a validation of this approach, Fig. 2.7 plots the signal-to-noise ratio, defined in this case as total flux divided by the error on the total flux (note that in the rest of the Chapter, F_{Peak} is used to calculate S/N), versus the natural logarithm of flux ratio, $\ln(R)$. Compact or unresolved sources are likely to lie under the black dashed line 99.9% of the time (see Eq. 2 and further discussion in Shimwell et al. 2022). This is indeed the case for the compact LOFAR-eFEDS sample we defined, but note that the inverse is not true and the curve cannot be used to select 'complex' sources. Instead, we simply define as 'complex' all those sources which do not satisfy at least one of our compactness criteria described above.

Overall, 24,613/36,631 (67%) of the LOFAR sources are classified as 'compact', and 12,018 as 'complex' radio emitters. Fig. 2.8 shows two examples of a prototypical compact and complex source in our sample, where the LS9 one-band image is overlaid with radio contours spanning several factors of the local noise rms.

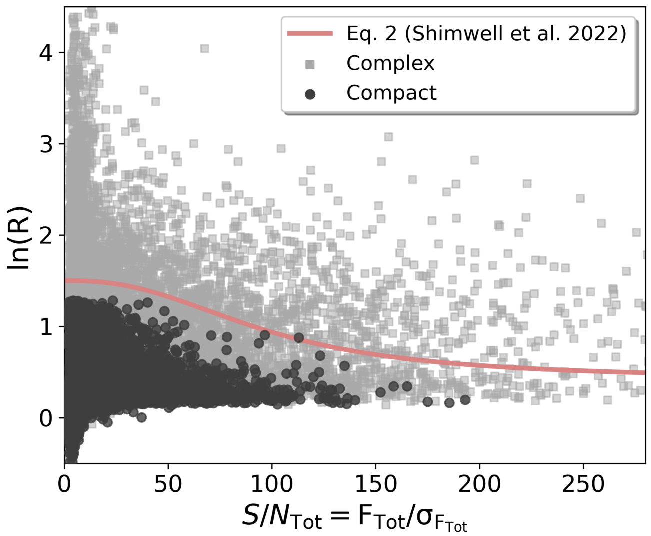


Figure 2.7: The natural logarithm of the ratio of total to peak fluxes versus the signal-to-noise ratio for the LOFAR sample of 36,631 sources. The signal-to-noise ratio is calculated here by dividing the total radio flux by its associated error. Sources classified as compact are shown in black, complex ones in light grey. The light red curve is taken from Shimwell et al. (2022) Eq. 2, below which 99.9% of all compact or unresolved sources lie.

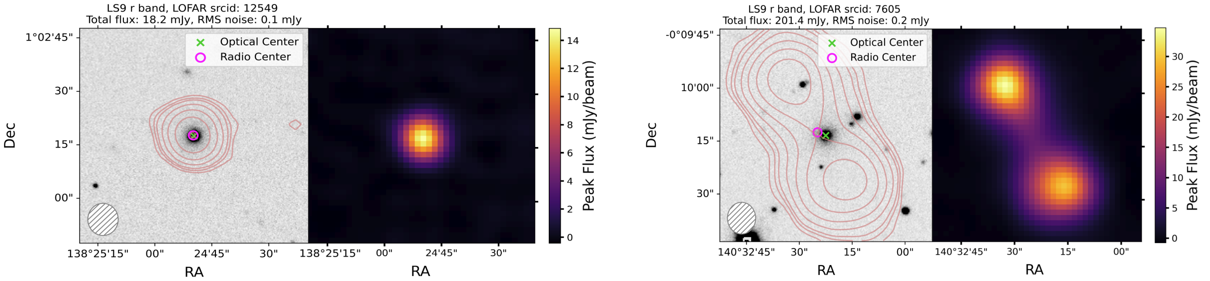


Figure 2.8: Cutouts (60″ × 60″) showing prototypical compact (left) and complex (right) radio morphologies with light red contours marking several factors of the local noise (at rms ×[2, 4, 6, 12, 24, 48]) and magenta circles indicating the radio centres. The beam size is depicted as a hatch-filled circle in the bottom left corner. In grey scale, the LS9 r-band image depicts the host galaxy, with its optical centre marked with a green cross. Radio intensity maps, with a colour bar indicating the peak flux, are also shown for these two examples.

2.3.2 Optical counterparts to the radio sources

The next step is to find optical counterparts to the LOFAR-eFEDS radio sources. Firstly, as noted in Williams et al. (2019), the positional errors in the LOFAR catalogues, as outputted automatically by PyBDSF, are often underestimated. This is because the pipeline returns only the error on the FWHM of the major and minor axes of the Gaussian fits, without considering additional correlated noise known empirically to be present. Following Williams et al. (2019), a $\sqrt{2}$ factor is applied to the catalogue positional error, and an additional astrometric uncertainty of $\sigma_{\rm ast} = 0.6$ is added in quadrature. The latter is common in radio to optical studies (Williams et al. 2019), arising from the directional asymmetries in mapping from major/minor axes to RA/Dec space in the optical, especially considering the low declination of the LOFAR-eFEDS field (meaning the beam is elongated). This step is necessary for NWAY to be able to accurately identify counterparts, especially for the obvious bright radio sources with unrealistically small positional errors. Overall, the positional error ($\sigma_{\rm pos}$; catalogue column `LOFAR_pos_err`) on the radio sources in this work is calculated using the following equation:

$$\sigma_{\rm pos} = \sqrt{2 * \left(\sigma^2_{\rm RA} + \sigma^2_{\rm Dec}\right) + \sigma^2_{\rm ast}} \tag{2.4}$$

where $\sigma_{\rm RA,Dec}$ are the 1σ errors on the RA, Dec position.

Drawing a cumulative histogram of the $\sigma_{\rm pos}$ values of the LOFAR sources revealed that only 10% of sources have $\sigma_{\rm pos} > 1.6''$. Therefore, the maximum search radius to be used in NWAY to find optical counterparts for the radio sources is taken to be $8''$ (five times this value). The LS9 positional uncertainties of the sources in Table 2.1 are much smaller, having an average of $0.1''$, which is taken as the constant $\sigma_{\rm pos}$ for all optical sources in the NWAY procedure.

Adding magnitude and/or colour priors significantly improves the accuracy of matching radio sources to their optical counterparts as radio emitters tend to be found in redder galaxies (e.g. ellipticals; see Williams et al. 2019). Therefore, absorption corrected *g, r, z* and *W1* magnitudes from LR9 (applying the extinction correction using the dust maps of Schlegel et al. 1998) are added as internal priors in the NWAY match, using the ‘auto’ feature, where NWAY learns to differentiate the magnitude or colour (or other source parameter) distributions between target and field sources ‘on the fly’ (for more details see Section B6.1 in Salvato et al. 2018). Lastly, the appropriate sky densities of each catalogue are calculated and a prior completeness fraction of 70% is assumed for the NWAY matching process (see e.g. Williams et al. 2019; Smolčić et al. 2008). The latter can be justified given that the radio sources in the LOFAR Deep Field Kondapally et al. (2021) with total integrated flux > 1 mJy, having optical counterparts above the LS9 r-band limit of 23.54 mag, is 68%. The non-detected sources are also likely high redshift ($z \sim 3-4$) obscured radio quiet quasars and so their absence would not impact the results of this work, based on a local sample of radio AGN (see e.g. Section 6.1 in Kondapally et al. 2021).

In this way, 33,769/36,631 LOFAR sources are matched to LS9 optical sources. To cut those matches which are statistically unlikely to be real matches, whilst keeping as many matches as possible (i.e. finding the balance between purity and completeness), an 'optimal' `p_any` cut is defined. This is done by creating a 'fake' match catalogue where the Declinations of the radio sources are shifted by 60″ (Salvato et al. 2022a). Any real radio sources within 8″ of such 'fake' sources are removed and then the 'fake' catalogue is again matched in the same way to LS9. A reverse cumulative ratio, effectively the 'completeness', is then calculated between the `p_any` distributions of the 'fake' to real matches. The purity is defined as one minus this ratio. Fig. 2.9 depicts the trade-off between purity (purple curve) and completeness (green curve), and the optimal `p_any`, located at the intersection, is 0.06. Applying this cut on the real sample results in 25,806/36,631 matches, or in other words a 70% match fraction. In comparison, Williams et al. (2019) find optical and/or IR counterparts in Pan-STARRS and WISE to 73% of the first data release of the LOFAR Two Meter Sky Survey (LoTSS-DR1). This agrees well with our result considering that the LoTSS-DR1 sky coverage is at a higher declination compared to the equatorial eFEDS field leading to lower rms noise and higher radio sensitivity, but that Pan-STARRS reaches to an overall shallower depth compared to LS DR9. Furthermore, it is in line with the aforementioned extrapolations drawn from the properties of optical counterparts to the LOFAR Deep Field radio sources (Kondapally et al. 2021).

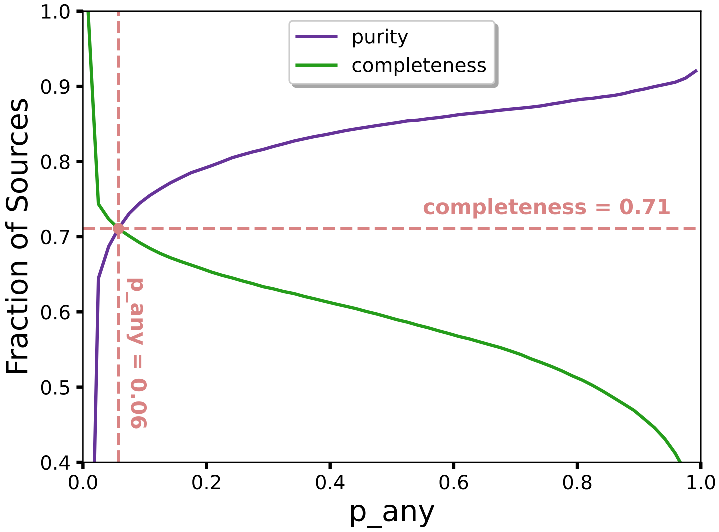


Figure 2.9: Purity (purple) and completeness (green) curves, found by creating a 'fake' catalogue of sources with shifted sky coordinates, to calculate the fractions of wrongly assigned counterparts as a function of `p_any` (see text for details). The optimal `p_any` threshold is equal to 0.06.

Then we choose to limit our analysis to LOFAR sources with S/N > 5, where S/N is defined as the ratio of the peak radio flux to the error in the peak flux[8]. This resulted in 22,754 matches between the LOFAR and LS9 catalogues.

---

[8] We note that 11% of the original 36,631 radio source sample have S/N < 5.

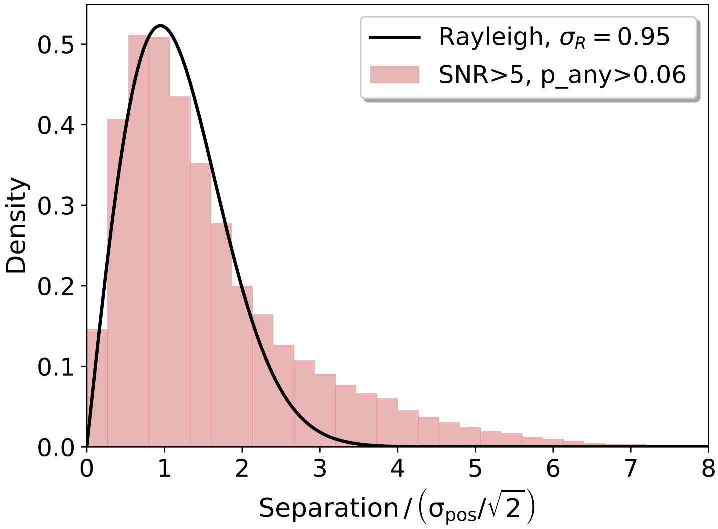


Figure 2.10: Rayleigh curve fit to the histogram of the ratio of counterpart separation to the 1D positional error for the sample of 22,754 LOFAR-LS9 sources. Visually, a good agreement can be seen with $\sigma_R \sim 1$, as expected, although there is a remnant tail (see text for discussion).

For positional (Gaussian) matching experiments, it is a common validation test to plot the distribution of the separation between counterparts divided by $\sigma_{\mathrm{pos}}/\sqrt{2}$ (Salvato et al. 2022a; Pineau et al. 2017), for which the best-fit should follow a Rayleigh distribution with $\sigma_R = 1$ (note: the division by $\sqrt{2}$ is to plot the one-dimensional positional error). Fig. 2.10 shows the radio-to-optical distribution for $S/N > 5$, `p_any` > 0.06 sources in light red. Even though the best fit Rayleigh distribution (black curve) is at $\sigma_R = 0.95$ and there is an excess of sources in the tail of the distribution, this is normal for ‘real’ distributions (Pineau et al. 2017). In particular, radio sources often do not have symmetric and Gaussian RA/Dec errors, which is one of the assumptions made in having a Rayleigh distribution with $\sigma_R = 1$, and the radio centring of complex morphology radio sources is not always trivial.

As a final step, after having determined the set of robust optical (LS9) counterparts, the LOFAR detections with LS9 counterparts are matched to the GAMA09 galaxies using a simple $2''$ positional match, similarly to the X-ray catalogue. In addition to this, six large radio galaxies with a GAMA09 counterpart, are appended to the sample after a further visual inspection procedure (see Ch. 2.3.3). With this addition, there are in total 22,759 LOFAR detections with an LS9 optical counterpart and 2,619 radio sources among the GAMA09 galaxies (see Table 2.1).

Of this overall total of 2,619 radio sources with counterparts in GAMA09, 1,901 have compact morphology, 718 complex. In Chapter 2.3.4, we further characterise these LOFAR-LS9-GAMA09 radio sources in terms of the origin of their radio emission and the properties of their host galaxies.

2.3.3 Visual inspection of radio sources

Visual inspection allows for the statistically identified counterparts by NWAY to be verified and the radio morphology to be further categorised. As this process is time-intensive and requires human resources, only the 'mass-complete' sources to be used in the final incidence analysis are visually inspected (see Ch. 2.3.5 for details on the mass and luminosity completeness considerations). These are marked in the LOFAR-eFEDS catalogue by the flag `vis_inspected=True`. We also investigate FRI and FRII sources (Fanaroff & Riley 1974b), which are powerful jetted AGN with core- and lobe-dominated emission, respectively, as well as giant radio galaxies (GRGs), which have a largest linear size of > 0.7 Mpc (e.g. Saripalli et al. 2005). This enhances the scientific impact of the LOFAR-eFEDS catalogue as such subsets of radio sources with complex morphologies are still yet to be fully characterised in terms of their radio, host galaxy and environmental properties. We investigate this further in Ch. 3.3.3 where we measure the fraction of FRII sources in different stellar mass bins and comment on how this impacts the shape of the incidence distribution in Ch. 4.2.

Three cutouts per source were created to be visually inspected:

1. ($1' \times 1'$) LS9 one-band image centred on the optical coordinates, with radio contours overlaid, as in Fig. 2.8.
2. ($1.8' \times 1.8'$) radio intensity image.
3. ($10' \times 10'$) radio intensity image (adjusted for visualising large dynamic ranges) to see the surroundings of the radio source in case of larger scale associated emission.

Figure 2.11 shows an example of these three cutouts for an example LOFAR-eFEDS source. As a result of the visual inspection, two sources, which had catalogued $S/N > 5$, were removed as they appeared to be associated with noise in the radio images, due to calibration errors in the spokes around bright sources (LOFAR Source id: 10347, 27051). Then four sources had to be rematched to different optical counterparts as the asymmetric or complex jetted morphologies resulted in the radio centring being wrong, meaning that the radio coordinates did not align with the visually identified location of the radio core (LOFAR Source id: 7310, 10975, 25001, 29295; for example id 29295 is a head-tail radio galaxy where the two jets are 'bent' behind the core, usually due to the galaxy moving through a dense ICM). There was one apparently compact case (LOFAR Source id: 22763) where the wider area radio intensity cutout revealed four additional associated radio components (diffuse lobes and compact hot spots) in a large-scale FRII source. In this case, the emission from the five components was summed to give the total radio flux (similarly, the two components associated with LOFAR Source id: 29295 were also summed). Fig. 2.11 shows the three cutouts created for this source, as described above, underlining the need to look at the larger scale environment. Likewise, a head-tail radio source (LOFAR Source id: 26644), catalogued as 'compact' was moved to the complex sample.

Moreover, during visual inspection, three inspectors would assign values of 1, 0.5, or 0 to each source, corresponding respectively to: secure FRII morphology (clear edge-brightened

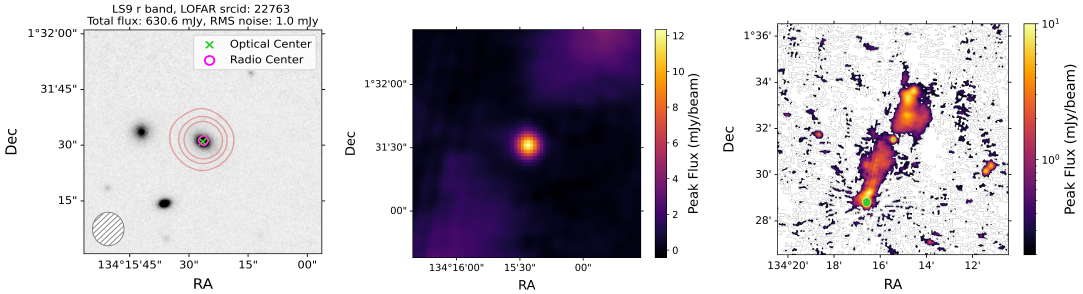


Figure 2.11: The three radio and optical cutouts of LOFAR Source id 22763, used for visual inspection and for classifying this source as an FRII (see text for details). We note that the green colour on the right-most panel indicates regions with peak flux greater than 10 mJy/beam, chosen to help better visualise the full dynamic range in the radio image.

double lobes); likely FRII (potentially unclear if edge- or core- brightened; and/or complicated asymmetric morphology and/or faint radio flux preventing secure classification, but source definitely has some extended structures) ; or unlikely FRII (no edge-brightened double-lobed structure). If the average of the three verdicts was $> 0.5, = 0.5$ or < 0.5, the source was classed as a 'secure', 'likely' or 'unlikely' FRII, respectively, and flagged accordingly in the VAC (`FRII_flag`). We note that only those sources which can visually be resolved into edge-brightened, double lobed components at the $8'' \times 9''$ resolution of LOFAR are considered FRIIs here. We find 28 FRIIs this way, which are combined with six further larger FRIIs (in addition to LOFAR Source id 22763 shown in Fig. 2.11), as described below.

The optical host identification method adopted in this paper is not tailored for finding large and powerful radio galaxies or sources with lobe components catalogued as separate IDs, thus these objects may be missed. To ascertain the completeness of the complex radio AGN sample, a test was made to visually inspect (with $10' \times 10'$ cutouts) all LOFAR-eFEDS sources in the GAMA09 area with F_{Tot} >10 mJy and `Maj`>19.1$''$ and `LOFAR_scodeS_flag`=`False`, matching to at least one other nearest neighbour within $2'$ (528 catalogue entries in total). A total of 78/528 catalogue entries were flagged as containing possible large, disconnected radio components. Each entry was then matched to GAMA09 within $5'$ to visually search for host galaxy counterparts. Nine sources were identified to match to a host galaxy detected in GAMA09, three of which were already present in the 'G9 radio AGN' sample (defined in the next section). This brings the total number of identified secure FRIIs in this investigation to 34 sources. Three out of these 34 sources are classified as giant radio galaxies, marked in the VAC with `GRG_flag`=`True`. The LOFAR components are summed to get the total flux and the largest linear projected size is used to calculate the extent of the source in kpc. The large-scale FRII source shown in Figure 2.11 is in fact an example of a GRG in the LOFAR-eFEDS field, previously discovered by Prescott et al. (2016), with jets stretching ∼1.4 Mpc in largest linear size.

2.3.4 Radio AGN versus star-forming galaxies

Radio emission can have a variety of origins, including star formation, AGN radio jets, AGN wind interactions and coronal emission (Panessa et al. 2019) and so it is vital for studies of radio AGN to be able to distinguish among these.

Different methods to separate star-forming galaxies from radio AGN have been widely discussed in the literature and have been refined significantly over the years with the advent of large surveys. Examples include: (i) radio SEDs and correlations with infrared parameters (Calistro Rivera et al. 2017; Gürkan et al. 2018; Yun et al. 2001; Delvecchio et al. 2021); (ii) brightness temperature (Morabito et al. 2022); (iii) using correlations between SFR (or proxies thereof, e.g. Hα) and radio emission to identify excess emission (Smith et al. 2021; Best et al. 2005; Kauffmann et al. 2008); (iv) emission line diagnostics, BPT diagrams (Baldwin et al. 1981; Kewley et al. 2006), (v) or combinations of the above and other methods, as discussed in Best & Heckman (2012); Sabater et al. (2019); Hardcastle et al. (2019), and references therein.

The method used in this work takes advantage of the highly reliable FUV to FIR SED fitting of the GAMA sources to calculate the SFR of all GAMA09 galaxies, as well as the tight correlation between SFR and radio luminosity for star forming galaxies (Condon 1992; Smith et al. 2021; Best et al. 2023a; Heesen et al. 2023). This relation is able to trace recent star formation effectively via synchrotron radiation emitted from massive stars ending their short lifetimes in supernovae explosions. Radio AGN can then be identified by measuring an excess with respect to the predicted SFR-related radio emission ('radio-excess AGN').

Fig. 2.12 shows the SFR (in units of $M_\odot$ yr^{-1}) plotted against the radio luminosity of the 1,901 compact (left panel) and 718 complex (right panel) radio sources detected by LOFAR and associated with a GAMA09 galaxy. Radio luminosity is calculated the standard way:

$$L_{144\mathrm{MHz}}\ [\mathrm{W\ Hz^{-1}}] = L_\mathrm{R} = 4\pi\ d_L^2\ F_\mathrm{Tot}\ 10^{-30}\ (1+z)^{\alpha-1}, \tag{2.5}$$

where d_L is the luminosity distance in cm, F_Tot is the total integrated flux[9] in units of Jansky (Jy) and $(1+z)^{\alpha-1}$ is the K-correction, with radio spectral index $\alpha = 0.7$ (Condon 1992).

The black solid line is the best fit derived by Best et al. (2023a) using the LoTSS Deep Fields (accounting for non-detections such that the relation is not biased by radio imaging depth):

$$\log_{10}(L_\mathrm{R}/[\mathrm{W\,Hz^{-1}}]) = 22.24 + 1.08\ \log_{10}(\mathrm{SFR}/[M_\odot\ \mathrm{yr}^{-1}]). \tag{2.6}$$

This relation is fully consistent within 0.1 dex with other recent relations (e.g. Smith et al. 2021) and tracks well the star-forming cloud of objects shown in grey in Fig. 2.12. As the overall population of sources above and below the best fit line are asymmetric (see Fig. 8 from Best et al. 2023a), the one-sided Gaussian spread of the distribution of radio luminosities below Eq. 2.6, which has $\sigma = 0.22$, is used to determine the cut for a source

---

[9] The following convention for the radio flux density as a function of frequency, S_ν, is used here: $S_\nu \propto \nu^{-\alpha}$.

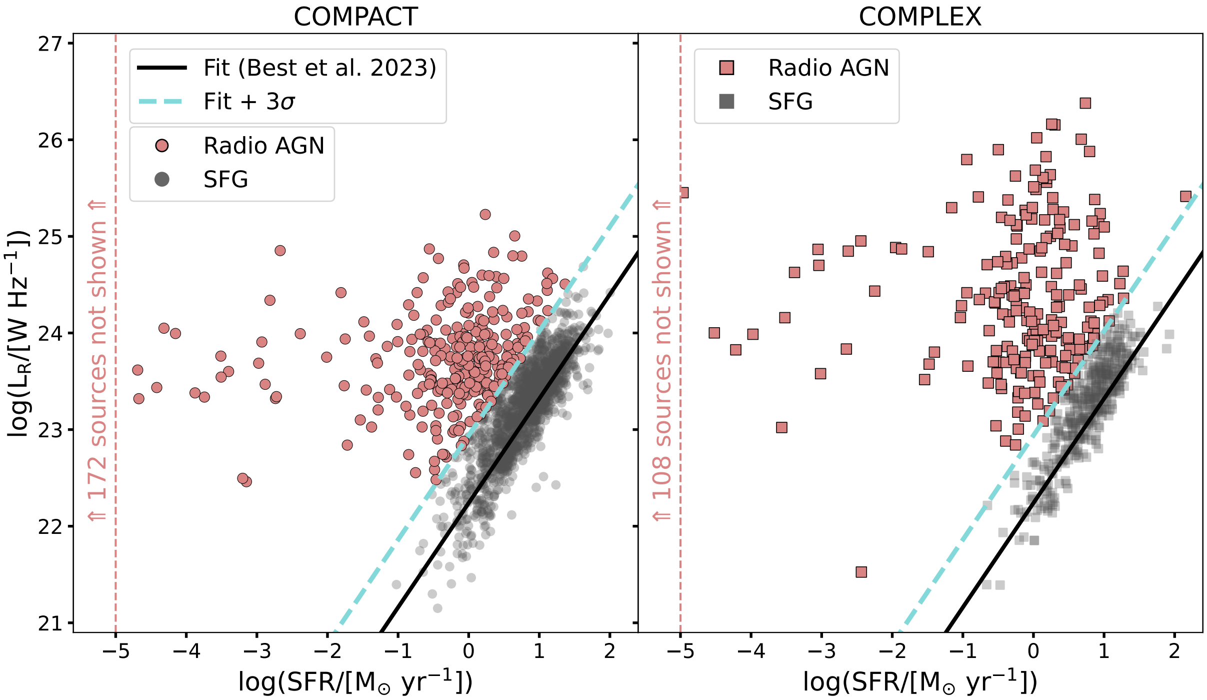


Figure 2.12: Radio luminosity versus SFR for radio sources within the GAMA09 galaxy sample. The black solid line, taken from Best et al. (2023a), describes the relation between SFR and $L_{\rm R}$ for star-forming galaxies hosting compact (left panel, grey circles) and complex (right panel, grey squares) radio sources. Sources lying 3σ above this relation (cyan, dashed line) form the sample of compact and complex radio AGN (light red circles, squares, respectively).

to be considered a radio AGN. All sources to the left of this cut, corresponding to 3σ (0.7 dex) above the relation (dashed cyan line), are defined as radio AGN, as they have radio luminosities in excess of what is expected from pure star formation. The 172 compact and 108 complex sources with very low SFR ($\log(\mathrm{SFR}/[M_\odot\ \mathrm{yr}^{-1}]) < -5$) are marked with text on the left hand side of each panel in Fig. 2.12 and also classify as radio AGN.

In our subsequent analysis, we only consider the radio AGN sample (light red points; marked in the catalogue by `G9_radioAGN=True`). From the total G9 radio sources, 445/1901 and 319/718 are compact and complex AGN, respectively. This sample has already been used in Popesso et al. (2023a) to study the incidences of radio AGN in brightest cluster galaxies (BCGs).

2.3.5 Stellar mass and radio luminosity complete radio AGN samples

Figure 2.13 (left) shows the stellar mass versus redshift distribution of the GAMA09 parent sample (grey points) and of the radio AGN among them (defined above), where the compact and complex radio emitters are marked with light circles and dark squares, respectively. The same stellar mass completeness curves, calculated as described in Ch. 2.2 above, are shown in black, since the completeness is dictated by the underlying GAMA09 galaxy mass distribution. White-filled markers are the radio AGN which are excluded as a result of this cut. Overall, there are 682 radio AGN in the 'mass-complete' (to 70%) sample, of which 404 are compact and 278 are complex. This corresponds to a total fraction of GAMA09 galaxies hosting a radio AGN detected by LOFAR of about 3% (682/21462).

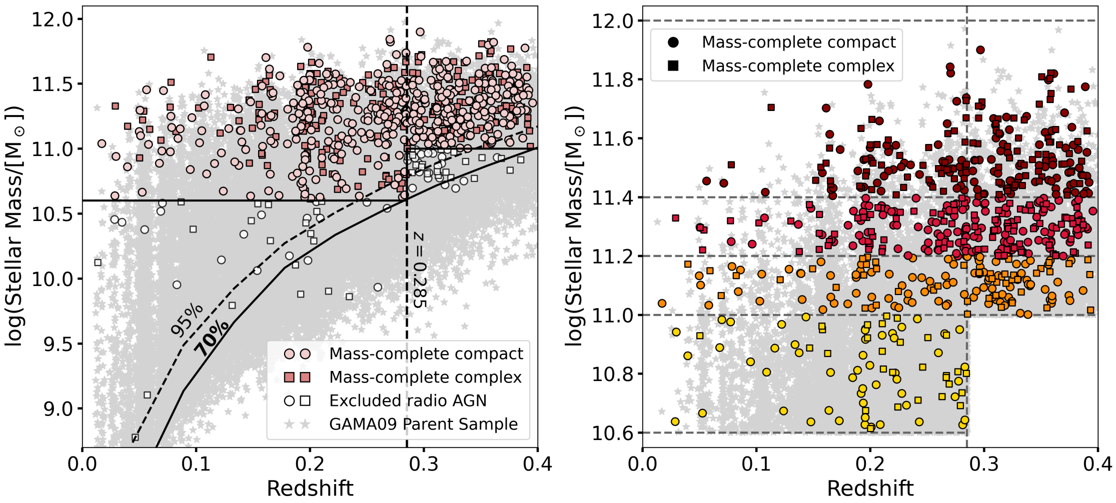


Figure 2.13: Stellar mass versus redshift distribution for the GAMA09 galaxies (grey points) and for the compact (light filled circles) and complex (dark filled squares) radio AGN. A vertical line divides the sample into two redshift bins. Completeness curves (70%, 95% with solid, dashed black lines, respectively) and horizontal thresholds are used to exclude sources incomplete in stellar mass (unfilled markers). A zoom-in of the mass-complete sample, split into four stellar mass bins (yellow, orange, red, crimson), is presented in the right panel.

Figure 2.14 shows the radio luminosity distribution with respect to redshift of the mass-complete G9 radio AGN. The colours and symbols are as above. Black dashed and dotted lines show the 80% and 95% radio luminosity completeness thresholds, respectively. The generation of the completeness curves is similar to that described in Shimwell et al. 2019 (their Fig. 14 and Section 3.6). A residual image of the entire LOFAR-eFEDS field is generated using PyBDSF. Then, 45,000 sources with flux densities ranging from 0.1 mJy to 10 Jy are injected into the residual image (in the image, not $u - v$, plane). The injected sources are searched and counted with PyBDSF. The injection procedure is done 50 times to improve the injection/detection statistic. Fig. 2.15 shows that the point-source completeness depends on the integrated flux density of the injected sources. For instance, 50% of the injected sources with flux densities above 0.34 mJy are detected. The completeness is 80% for sources brighter than 0.85 mJy, and it increases to 95% for sources

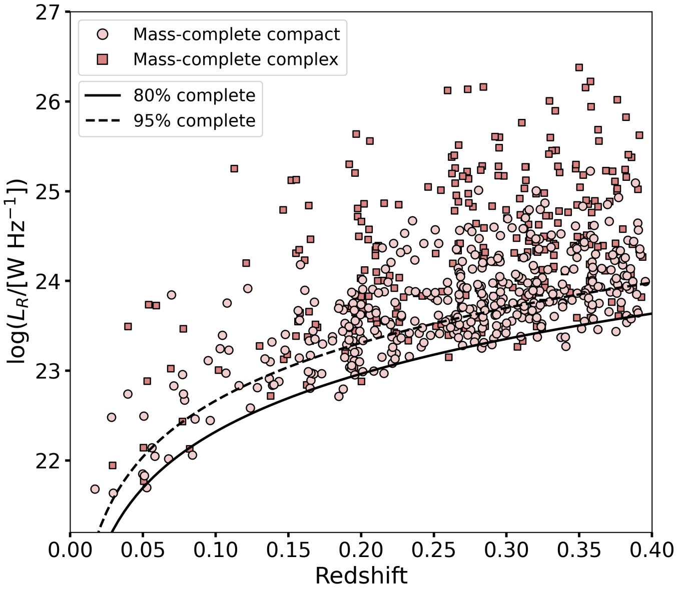


Figure 2.14: Radio luminosity versus redshift distribution for the compact and complex G9 radio AGN in the two redshift bins (colours and symbols are as above). Black dashed and dotted curves show the 80% and 95% radio luminosity completeness limits, respectively.

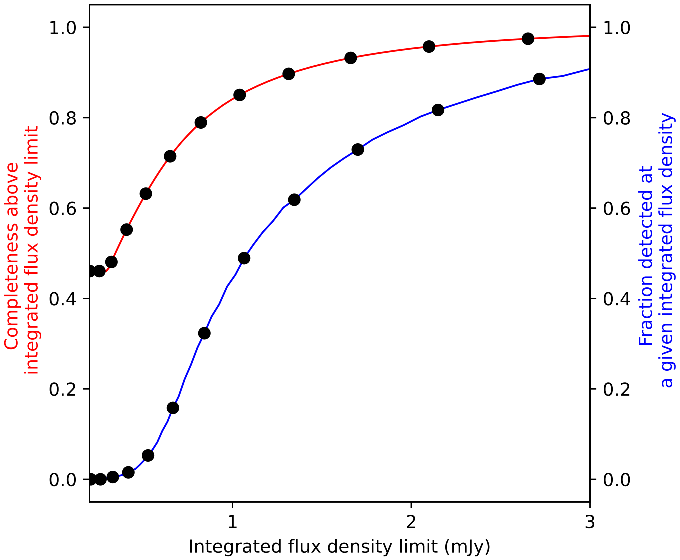


Figure 2.15: Point-source completeness functions for the LOFAR-eFEDS field, calculated by injecting simulated sources with a range of radio intensities onto the field's residual image. The red and blue lines show the cumulative completeness above and the fraction of detected sources at a given integrated flux density, respectively. The former is used to derive the luminosity completeness curves shown in orange on Fig. 2.14.

with flux densities above 1.88 mJy. At the same completeness level, this LOFAR-eFEDS data requires sources to be a factor of $2-3$ brighter than those in the LoTSS-DR1 images of Shimwell et al. (2019) to be detected, mainly due to the higher noise level of this low declination field. In a similar way to the X-ray AGN, a weighting per bin is applied to account for radio luminosity dependent incompleteness (see details in Ch. 3.1).

The radio physical size is also calculated via Eq. 2.7, to better classify the complex sample and comment on potential surface brightness limitations (see Ch. 3.5).

$$R_{\mathrm{kpc}} = \theta * d_L/(1+z)^2, \tag{2.7}$$

where θ is major axis in radians and d_L is the luminosity distance in kpc.

2.3.6 Host galaxy properties of radio AGN

In analogy to our analysis of the X-ray AGN sample, Fig. 2.16 shows SFR versus stellar mass for the GAMA09 parent sample (purple hexbins) and the LOFAR radio AGN (yellow, orange; compact: circles, complex: squares). The MS relation is plotted in black for the two redshift bins, along with a dashed line marking SFRs $3\sigma_{\mathrm{intr}}$ below the MS. The final numbers of sources are listed in Table 2.1. All sources with $\log(\mathrm{SFR}/[\mathrm{M}_\odot\ \mathrm{yr}^{-1}]) \leq -5$ are plotted below the horizontal black dotted line with arbitrary y-axis shift to show their distribution in stellar mass. In the low (high) redshift bin, there are a total of 164 (116) and 5453 (2121) radio AGN and GAMA09 galaxies, respectively which have $\log(\mathrm{SFR}/[\mathrm{M}_\odot\ \mathrm{yr}^{-1}]) \leq -5$.

2.4 Combined X-ray and radio AGN sample characterisation

The LOFAR-LS9-GAMA09 (radio) and eFEDS-LS9-GAMA09 (X-ray) catalogues can be combined by matching the GAMA09 source IDs (or coordinates). This results in 121 sources emitting at both X-ray and radio wavelengths (marked by `G9_radioXray_sources =True`), of which 74 and 32 are mass-complete X-ray and radio sources, respectively. However, only 24 of these mass-complete sources classify as both radio and X-ray AGN, as set by the criteria defined in the previous sections.

For X-ray detected AGN, following Aird et al. (2012), we adopt the following definition of the specific black hole radiative power or specific accretion rate:

$$\lambda_{\mathrm{Edd}} = \frac{L_{\mathrm{Bol}}}{L_{\mathrm{Edd}}} = \frac{k_{\mathrm{bol}}\ L_{\mathrm{X,Hard}}}{1.26 \times 10^{38}\ \mathrm{erg\ s^{-1}\ M_{BH}/M_\odot}}, \tag{2.8}$$

where a simple bolometric correction factor of $k_{\mathrm{bol}} = 25$ is chosen to convert from hard ($2-10$ keV) X-ray luminosity to bolometric luminosity[10] (L_{Bol}). The Eddington luminosity

---

[10] We note that although bolometric correction factors depend on AGN luminosity, the chosen value of 25 agrees well with the range of 15-30 found in past studies on large samples for the luminosity range probed in this study (Vasudevan & Fabian 2009).

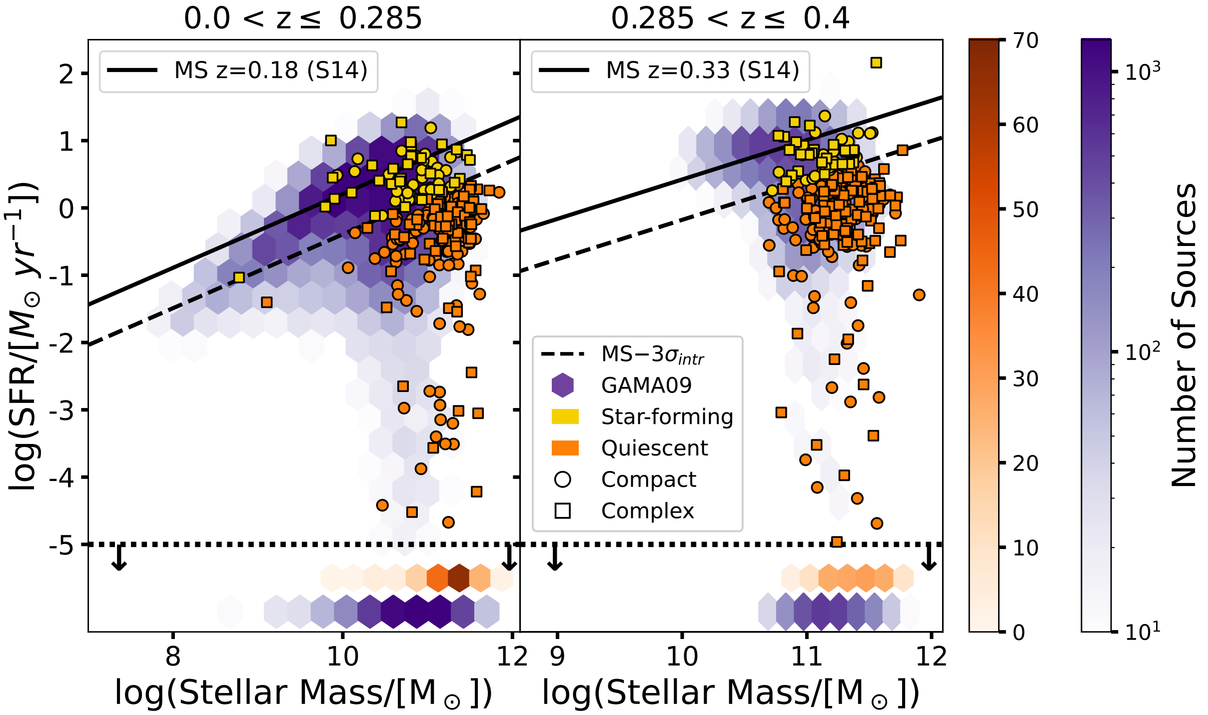


Figure 2.16: SFR versus stellar mass for the GAMA09 galaxy sample (purple hexbins) and for the compact/complex (circles/squares) radio AGN. The solid black line in each redshift panel marks the star-forming galaxy main sequence from (S14; Speagle et al. 2014). Sources 3σ below this line (black, dashed) are considered to be quiescent galaxies (orange), otherwise they are classed as star-forming (yellow). Quiescent sources below $\log(\mathrm{SFR}/[\mathrm{M}_\odot\ \mathrm{yr}^{-1}]) = -5$ are marked as upper limits and their distributions are shown in the bottom of each panel.

(L_{Edd}) is calculated using an estimate of the black hole mass as $M_{BH} \sim 0.002\ M_*$, assuming the mass of the bulge is equal to M_* (Marconi & Hunt 2003). The goal of λ_{Edd} is to serve as a mass-scaled scaled power indicator, not necessarily as the true 'Eddington ratio' of the AGN, which is inherently difficult to constrain given the ~ 0.4 dex systematic uncertainties on black hole mass measurements.

For radio AGN, the derivation of a mass-scaled jet power is more complicated. Two common methods to estimate jet power (Q) are either to calculate the work done by jets to inflate cavities in nearby cluster AGN using X-ray observations and combine with an estimate of the source age, or to infer it from correlations between narrow emission line luminosity and radio emission (Willott et al. 1999; Hardcastle et al. 2007; Merloni & Heinz 2008; Cavagnolo et al. 2010; Daly et al. 2012; Godfrey & Shabala 2013; Heckman & Best 2014; Ineson et al. 2017; Hardcastle 2018). Although there are several caveats that come with using such scaling relations (see Ch. 3.5), here we adopt the empirical relation from Heckman & Best (2014) (their Eq. 2) to define jet power, as it is based on a combination of the approaches mentioned above:

$$Q = 2.8 \times 10^{37} \left(\frac{L_{1.4\mathrm{GHz}}}{10^{25}\ \mathrm{W\ Hz}^{-1}} \right)^{0.68} \mathrm{W}, \tag{2.9}$$

where the LOFAR 144MHz radio luminosity is converted to 1.4 GHz radio luminosity assuming a spectral index of $\alpha = 0.7$. As can be seen, the equation is non-linear. Normalising by L_{Edd}, we then define the specific black hole kinetic power:

$$\lambda_{\mathrm{Jet}} = \frac{Q}{L_{\mathrm{Edd}}}. \tag{2.10}$$

Figure 2.17 compares λ_{Jet} with the well-known measure of 'radio-loudness' (R), a measure of the dominance of the radio emission over mid-IR emission, with 6μm luminosity ($L_{6\mu m}$), assumed to be a direct tracer of the reprocessed primary emission from accretion processes. $L_{6\mu m}$ is calculated via a log-linear interpolation (or extrapolation) of WISE fluxes and the threshold for a source to be considered 'radio-loud' is $R > -4.2$ (Klindt et al. 2019). Only mass-complete sources and those passing the 80% radio luminosity completeness curve are plotted on Fig. 2.17 (552/682 radio AGN; 21/24 radio and X-ray AGN). Sources lacking good (any) WISE data are marked as lower limits in their radio-loudness and their $L_{6\mu m}$ upper limits are calculated from the WISE 5σ PSF depths (Wright et al. 2010). Star-forming galaxies are marked as upper limits in Q/L_{Edd} as their possible AGN emission is indistinguishable from their star-formation emission. One can clearly see the relatively tight correlation, as expected for higher kinetic power objects to be more 'radio-loud'. Partial correlation analysis reveals a strong positive correlation (Pearson coefficient, r, of 0.911) when controlling for stellar mass as a covariate, although the correlation becomes weaker ($r = 0.428$) when controlling for radio luminosity, which makes sense as it is a common variable in both axes. Overall, 60% (47%, 77%) of the total (compact, complex) radio AGN sample are radio-loud (light red circles and squares), compared to only 2% for the star-forming galaxies (grey upper limits). Around 76% (61%, 100%) of the total (compact, complex) radio AGN also detected in X-rays (green circles and squares) are radio loud. FRII-like morphologies and giant radio galaxies, marked with black crosses and stars, respectively, tend to populate the high radio-loudness regime, in line with their expected powerful jets.

Figure 2.18 plots the radiative versus kinetic power for all the radio and X-ray AGN detected among GAMA galaxies. Different samples of sources are shown, namely radio and X-ray AGN (green), X-ray AGN in star-forming galaxies (black upper limits), X-ray AGN with no radio detections (blue), and radio AGN with no X-ray detection (light red). All samples are complete for stellar mass at the 70% level. For the radio-undetected sources, the 99% flux limit (3.2 mJy) from Fig. 2.15 is adopted. For the X-ray-undetected sources the eFEDS survey-average $0.5 - 2$ keV flux 80% completeness limit equal to 6.5×10^{-15} erg cm^{-2} from Brunner et al. (2022) is converted to a $2 - 10$ keV luminosity using a $\Gamma = 2$ power law at the redshift of the source. A sample representative error of 0.4 dex in both $x - y$ variables is plotted in the bottom right corner. The grey solid

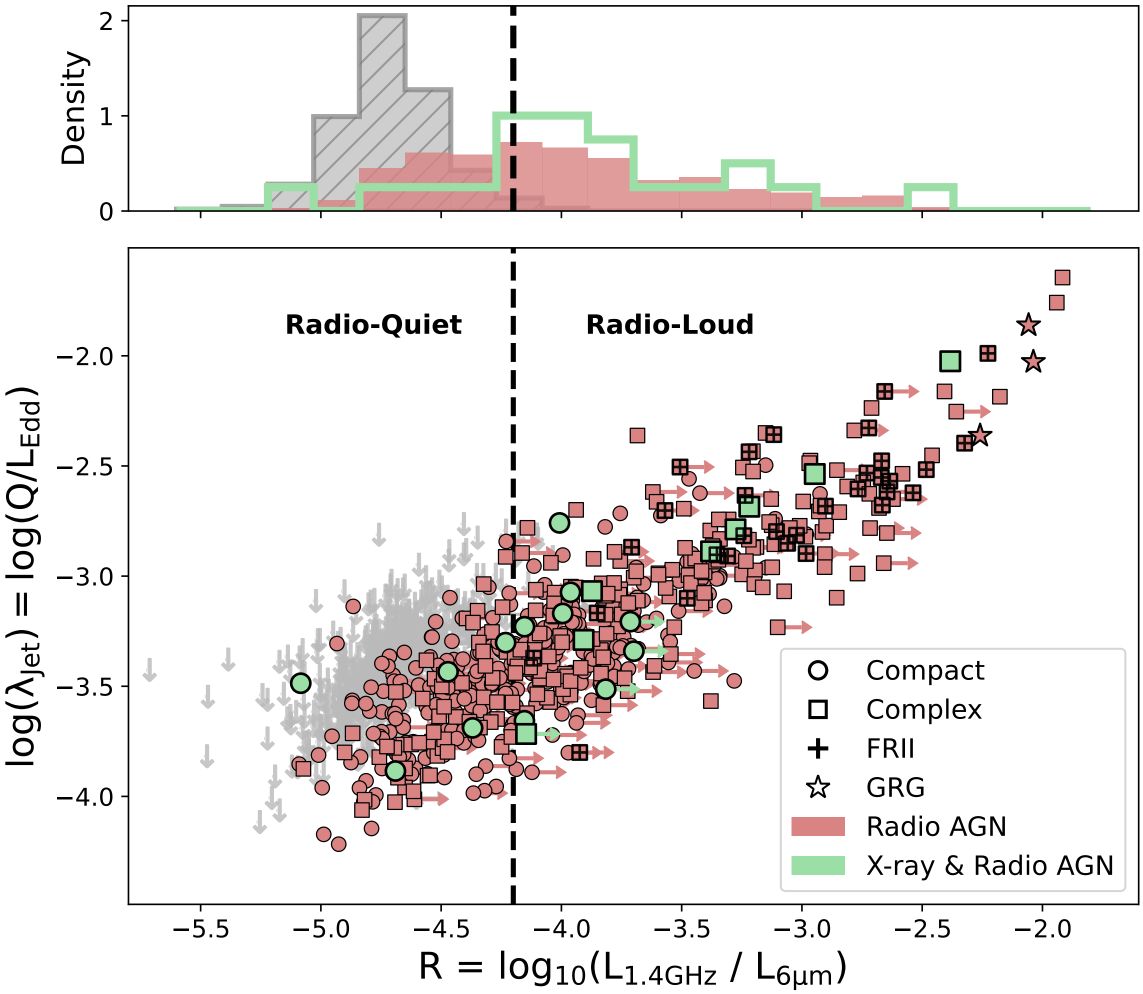


Figure 2.17: λ_{Jet} plotted against radio-loudness, using 6μm luminosity as a proxy for accretion luminosity, for the mass-complete compact (circles) and complex (squares) radio AGN (light red), star-forming galaxies (grey upper limits), those radio AGN also detected in X-rays (green), and those that have secure FRII morphologies (black crosses) or are giant radio galaxies (stars). It can be seen that sources with $\log(\lambda_{\text{Jet}}) \gtrsim -3.0$ are almost exclusively 'radio-loud'.

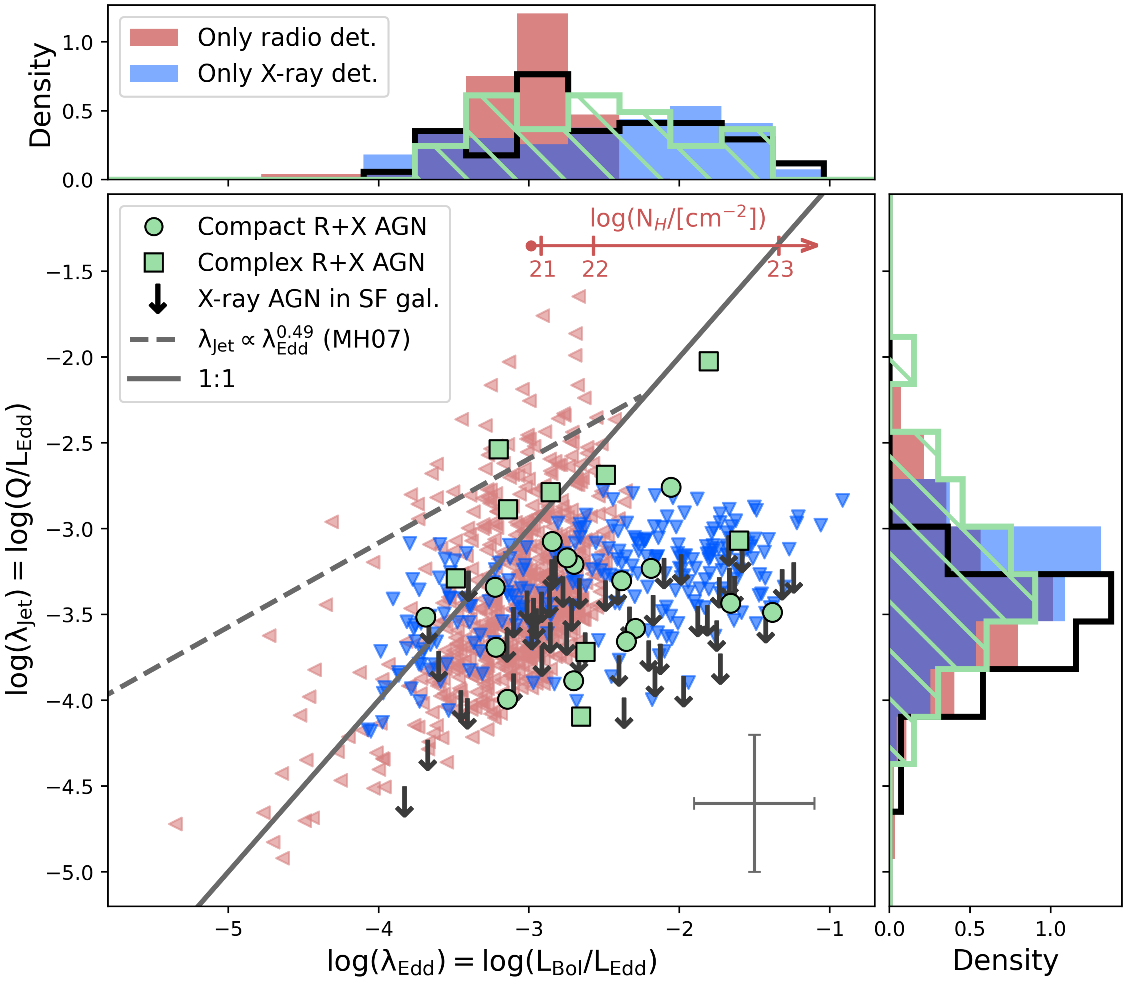


Figure 2.18: Mass-complete distribution of $\lambda_{\rm Jet}$ versus $\lambda_{\rm Edd}$ for the radio and X-ray AGN (green), X-ray AGN in star-forming galaxies (black), only radio-detected AGN (light red) and only X-ray-detected AGN (blue). Solid grey and dashed grey lines mark the 1:1 and $\lambda_{\rm Jet} \propto \lambda_{\rm Edd}^{0.49}$, respectively. Horizontal arrow marks the effect of obscuration on only radio-detected AGN (see text).

line marks the 1:1 relation, whereas the grey dashed line is adapted from Eq. 3 of Merloni & Heinz (2007) and describes the radiatively inefficient ADAF mode where $\lambda_{\rm Jet} \propto \lambda_{\rm Edd}^{0.49}$ (with intrinsic scatter of 0.39), as observed for local radio galaxies in groups and clusters. It can be seen that the different samples populate the region where the radiatively inefficient mode starts branching away from the 1:1 relation (which effectively traces the radiatively efficient mode). This means that the determination of the accretion mode for a given object is unreliable (considering also the relatively large systematic errors in the determination of $\lambda_{\rm Jet}$ and $\lambda_{\rm Edd}$). Therefore, it is unfeasible to make statements about the effect of different accretion modes on jet power using this sample. In fact, the bulk of the sources, especially the ones detected only in X-rays, do not populate the $\lambda_{\rm Jet} \propto \lambda_{\rm Edd}^{0.49}$ (radiatively inefficient, kinetically dominated) branch, where the fundamental plane is supposed to be valid. Overall, the eFEDS observations are not sensitive enough to probe the low-power population, where an accretion mode transition would be more obvious; at the same time, the survey volume is too small to detect many high-power sources.

In general, the radio-detected population scatters around $\lambda_{\rm Jet} \sim \lambda_{\rm Edd}$ line, whereas the X-ray detected one has $\lambda_{\rm Jet} \ll \lambda_{\rm Edd}$, as expected from kinetically and radiatively dominant accretion modes, respectively. However, there is a hint that X-ray AGN that are radio-loud and radiatively efficient (upper right region of Fig. 2.18) may appear as a distinct population from their radio-quiet counterparts at the same $\lambda_{\rm Edd}$ (see also Ichikawa et al. (2023) for a discussion of the balance of power in higher redshift radio and X-ray detected AGN). Additionally, we note that in the parameter space populated by the bulk of the X-ray detected AGN, coined as the 'high radiative' (HR; radio-quiet) branch (Merloni & Heinz 2008), even the origin of the radio emission is highly debated (e.g. wind, jet, or coronal origin; see Panessa et al. 2019). This is important as it would require different physical models than the one of a scale-invariant, freely expanding jet model presented in Heinz & Sunyaev (2003).

The G9 radio AGN sample also contains some radio-detected sources with $\lambda_{\rm Jet} \gg \lambda_{\rm Edd}$ at high intrinsic jet kinetic power. However, the location of these sources could be affected by X-ray obscuration. We test the effect that different levels of $\log(N_H/[\rm cm^{-2}]) = 21, 22, 23$ have on the measurement of $\lambda_{\rm Edd}$ estimate for a source with observed flux equal to the eFEDS 80% limit, an average redshift of 0.24 and an average $\log(M_*/M_\odot) = 11$. A horizontal red arrow in the upper right corner of Fig. 2.18 shows the extent to which the intrinsic flux, and thereby $\lambda_{\rm Edd}$, increases as a function of increasing obscuration (for a constant observed flux), resulting in a shift of radio-detected sources toward the 1:1 line. Future work could investigate the effect of obscuration in more detail using X-ray spectral analysis of the detected sources or by conducting X-ray stacking experiments in two different X-ray bands for the undetected sources. Overall, it is very difficult to get a reliable measurement of the fundamental accretion mode without having simultaneous measurements of accretion and jet-driven power for many of our objects, with the majority of the AGN being detected either in radio or in X-rays, but not in both.

2.5 General sample statistics

The general AGN sample statistics show that 3% and 1.5% of GAMA09 galaxies are detected as radio and X-ray AGN, respectively (see Fig. 3.1 later for more details on the fraction of AGN in different stellar mass and redshift bins). Taking only the mass-complete samples, 7% of X-ray AGN are also radio AGN, in line with the commonly expected population of 'radio-loud' QSOs. Yet only 4% of LOFAR-detected radio AGN are X-ray detected.

At face value, Figures 2.5 and 2.16 show that radio AGN tend to lie mostly (87%) in quiescent galaxies, in contrast to X-ray AGN which are found in star-forming galaxies 62% of the time (percentages computed from mass-complete samples only, see Table 2.1). However, in Ch. 3.3 we show, albeit with limited statistics, that quiescent and star-forming radio AGN have in fact similar incidence as a function of mass-normalised jet power, once mass- and luminosity- completeness has been accounted for.

2.6 Details of LOFAR-eFEDS Value-Added Catalogue

Table 2.2 gives a description of all columns in the VAC. A boolean value of 1 (0) is equivalent to `True` (`False`) in the columns containing flags, from which one can select various subsamples. We note that only the subset of mass-complete G9 radio AGN were visually inspected in this work, as described in Ch. 2.3.3; the other sources, especially those with complex radio morphologies, may need further validation to ascertain the correctness of their optical counterparts as radio centring may not be trivial for such objects (as discussed in Ch. 2.3). This also means that the `FRII_flag` entries are only valid for the subset which was visually inspected.

Twelve LOFAR sources with different handling in their optical counterpart identification and/or radio property calculation are marked with `special_flag`=1. Full details are given in Ch. 2.3.3 but a summary is presented here. Four sources (LOFAR Source id: 7310, 10975, 25001, 29295) have been manually rematched to their correct LS9 optical counterpart during the visual inspection process. Two sources (LOFAR Source id: 22763, 26644), with correctly identified optical counterparts, were moved from the compact to the complex radio morphology sample and had their radio properties updated manually, including summing associated component emission. Six sources (LOFAR Source id: 8153, 14599, 23634, 29440, 29781, 32863), from the additional visual inspection process to find large FRIIs, also had their optical counterparts and radio properties updated manually. For sources with manually assigned counterparts, the `p_any` and `p_i` entries are NULL. For sources with summed components, the `LOFAR_Total_flux` and `LOFAR_E_Total_flux` are summed/propagated accordingly from the individual component entries; the `LOFAR_Peak_flux` and `LOFAR_E_Peak_flux` entries are NULL; the `LOFAR_Maj` entry corresponds to the projected largest linear size in degrees; and `CTP_Separation` is taken as the distance from the given LOFAR Source id and the matched LS9 counterpart.

To summarise, out of the total 36,631 LOFAR sources in the eFEDS region where the

X-ray exposure time exceeds 500s, 33,769 matched to an LS9 galaxy. After certain quality cuts to do with `p_any` and SNR thresholds, 22,754 matches remain. The addition of 5/6 large FRIIs from the visual-inspection brings that total up to 22,759 LOFAR detections with an LS9 optical counterpart (the remaining large FRII source with LOFAR Source id 8153 was already present among the matched sources, albeit the optical CTP had to be corrected, see Table 2.1). Matching via optical coordinates to the GAMA09 catalogue, using a 2″ match radius, results in 2,619 radio sources among the GAMA09 galaxies, of which 764 (or 682 mass-complete) are identified as radio AGN. This is the science sample of G9 radio AGN that is the subject of the next Chapter.

Table 2.2: Column descriptions for the LOFAR-eFEDS value-added catalogue.

Column Name	**Description**
LOFAR	
`LOFAR_Source_id`	Unique number that identifies the source.
`LOFAR_RA`	Right ascension of the source (for the equinox of the image), in degrees.
`LOFAR_DEC`	Declination of the source (for the equinox of the image), in degrees.
`LOFAR_Total_Flux`	Total, integrated Stokes I flux density of the source at the reference frequency, in Jy.
`LOFAR_E_Total_Flux`	1σ error on the total flux density of the source, in Jy.
`LOFAR_Peak_Flux`	Peak Stokes I flux density per beam of the source, in Jy/beam.
`LOFAR_E_Peak_Flux`	1σ error on the peak flux density per beam of the source, in Jy/beam.
`LOFAR_Maj`	FWHM of the major axis of the source, in degrees.
`LOFAR_pos_err`	Positional error calculated using Eq. 2.4.
LOFAR: Morphology Flags	
`LOFAR_scodeS_flag`	`True` if `LOFAR_S_code=S`, i.e. fit with only a single Gaussian by PyBDSF.
`LOFAR_fluxratio_flag`	`True` if `LOFAR_Total_Flux/LOFAR_Peak_Flux` < 3.6.
`LOFAR_maj_flag`	`True` if `LOFAR_Maj` $< 19.1''$.
`LOFAR_isolated_flag`	`True` if no nearest neighbours within $45''$.
`LOFAR_compact_flag`	`True`: compact source; all four flags above are `True`. `False`: complex source.
Legacy Survey DR9	
`LS9_UNIQUE_OBJID`	Unique source identifier: `BRICKID_OBJID` (`RELEASE`=9010 for sources present here).
`LS9_TYPE`	Morphological model.

Column Name	**Description**
`LS9_RA`	Right ascension at equinox J2000.
`LS9_DEC`	Declination at equinox J2000.
`LS9_pos_err`	Positional error.
`LS9_mag_g_dered`	Dereddened g-band magnitude.
`LS9_mag_r_dered`	Dereddened r-band magnitude.
`LS9_mag_z_dered`	Dereddened z-band magnitude.
`LS9_mag_W1_dered`	Dereddened W1-band magnitude.
`LS9_mag_W2_dered`	Dereddened W2-band magnitude.
`LS9_mag_W3_dered`	Dereddened W3-band magnitude.
`LS9_mag_W4_dered`	Dereddened W4-band magnitude.
NWAY Match	
`CTP_Separation`	Separation, in arcsec, between LOFAR source and best-match LS9 counterpart.
`p_i`	Probability for the counterpart to be the correct one.
`p_any`	Probability for a source to have any counterpart in the search region. Optimal: `p_any` > 0.06.
GAMA09	
`uberID`	Unique GAMA ID of object.
`CATAID`	Unique numeric GAMA object identifier.
`RAcen`	Right Ascension of flux-weighted centre (ICRS).
`Deccen`	Declination of flux-weighted centre (ICRS).
`Z`	Spectroscopic redshift.
`StellarMass_50`	Median stellar mass from MCMC chain.
`StellarMass_16`	16th percentile stellar mass from MCMC chain.
`StellarMass_84`	84th percentile stellar mass from MCMC chain.
`SFR_50`	Median SFR from MCMC chain.
`SFR_16`	16th percentile SFR from MCMC chain.
`SFR_84`	84th percentile SFR from MCMC chain.
`SC`	Science sample class. `SC` ≥ 6 is used here.
Additional Flags	
`radioAGN_flag`	`True` if LOFAR source (with S/N>5, `p_any`>0.06) fulfils radio-excess criterion defined by Eq. 2.6.
`vis_inspected`	`True` if source has been visually inspected (see Appendix 2.3.2).
`G9_radioAGN`	Final sample of mass-complete G9 radio AGN, with visual inspection results applied (see Table 2.1).
`FRII_flag`	`FRII_flag` $= 1, 0.5, 0$: secure, likely, unlikely FRII-morphology, respectively.
`GRG_flag`	Giant radio galaxy flag (largest linear size > 0.7 Mpc).

Column Name	Description
`G9_radioXray_sources`	`True` if the LOFAR source has an X-ray match in eROSITA eFEDS (see Table 2.1).
`special_flag`	`True` if source required special cross-matching and/or property estimation (see Appendix A).

2.7 Summary

In this Chapter, we present the work done to define and characterise a new sample of radio and X-ray AGN, and their optical counterparts in LS9, in the LOFAR-eFEDS field. We additionally exploit the well-characterised, spectroscopic sample of galaxies in the GAMA09 field to build a volume-limited parent galaxy sample, complete in stellar mass, from which we can compute the incidence of our X-ray or radio luminosity-complete target sample of X-ray or radio AGN (see Chapter 3). Using purely the radio information, we automatise the classification of radio sources into compact or complex morphologies. We use the robust SFR and stellar mass estimates from the GAMA09 survey to define samples of radio and X-ray AGN, which we further split into those AGN hosted in quiescent or star-forming galaxies. We find that radio AGN predominantly populate quiescent galaxies, whereas X-ray AGN tend to be found in star-forming galaxies. Yet, it still remains to be seen how radio morphology and host galaxy properties, including stellar mass and SFR, affect the powering of jets in radio AGN. This will be the main subject of the following Chapter. Lastly, we see hints of different accretion modes present in our sample of our radio and X-ray sources, given their location on the $\lambda_{\mathrm{Edd}} - \lambda_{\mathrm{Jet}}$ plot (Fig. 2.18). However, in order to better understand the true interplay between the corona and the jet, we would need both more sensitive *and* larger volume X-ray and radio surveys to detect the numerous faint, radiatively-inefficient accretion mode and rarer bright, radiatively-efficient AGN.

Chapter 3

The incidence of radio and X-ray AGN and the disk-jet coupling

In this Chapter, we build on the X-ray and radio AGN samples from the LOFAR-eFEDS survey to calculate incidence distributions as a function of mass-scaled power indicators. In particular, we investigate whether or not jet powering shows the same mass-invariance as is suggested by the X-ray AGN incidence, and we explore the constraints that can be placed on the disk–jet connection by the measured incidences of radio and X-ray AGN.

The key findings are the following. Despite the soft-X-ray eROSITA-selected sample, the incidence of X-ray AGN as a function of λ_{Edd} shows the same mass-invariance and power law slope (-0.65) as that found in Aird et al. (2012) once corrected for completeness. Across the M_* range probed, the incidence of compact radio AGN as a function of λ_{Jet} is described by a power law with constant slope, showing that it is not only high mass galaxies hosting high power jets and vice versa. This slope is steeper than that of the X-ray incidence, which has a value of around -1.5. Furthermore, higher-mass galaxies are more likely to host radio AGN across the λ_{Jet} range, indicating some residual mass dependence of jet powering. Upon adding complex radio morphologies, including 34 FRIIs, three of which are giant radio galaxies, the incidence not only shows a larger mass dependence but also a jet power dependence, being clearly boosted at high λ_{Jet} values. Importantly, the latter effect cannot be explained by such radio AGN residing in more dense environments (or more massive dark matter haloes). Moreover, the similarity in the incidence of quiescent and star-forming radio AGN reveals that radio AGN are not only found in ‘red and dead’ galaxies.

The striking difference in the compact versus complex radio AGN incidence as a function of λ_{Jet}, discovered in this work, highlights a varied mass and morphology dependence of radio jet powering in AGN. This will be investigated in more detail in Chapter 4 to understand how this disparity translates into the ways in which these radio AGN feed back energy to their host galaxies and halos.

The contents of this Chapter are published in Sections 4 – 6 and Appendix D of Igo et al. (2024).

3.1 Calculating the incidence of AGN among GAMA09 galaxies

In this section, we present the methodology adopted to calculate AGN incidence as a function of stellar mass and specific black hole kinetic (from radio) or radiative (from X-rays) power, along with extra corrections accounting for completeness.

3.1.1 Stellar mass–redshift binning

Firstly, as shown on Fig. 2.13 (left), the radio data is split into two redshift bins, to limit evolutionary or redshift dependent completeness effects on the analysis. They are determined by dividing the sample of sources equally in two: (i) $0 < z \leq 0.285$; (ii) $0.285 < z \leq 0.4$. The same two redshift bins are then adopted for the X-ray analysis as shown on Fig. 2.3 (left), for consistency, and are represented, where relevant, with light grey and black colours throughout this Chapter.

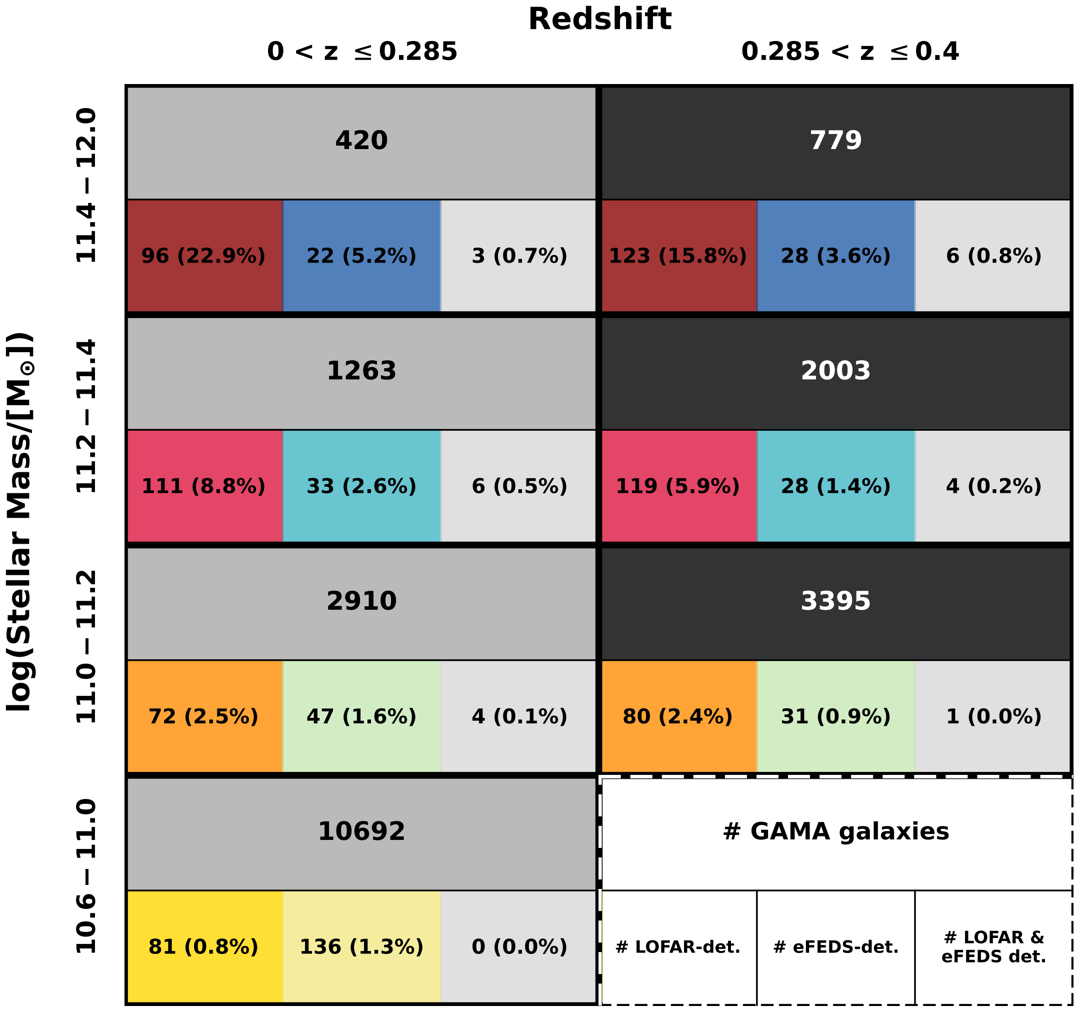
Secondly, four stellar mass bins are introduced, as shown on the right panels of Figs. 2.3 and 2.13, in ranges of $\log(M_*/M_\odot)$: *(i)* $10.6 - 11.0$; *(ii)* $11.0 - 11.2$; *(iii)* $11.2 - 11.4$; *(iv)* $11.4 - 12.0$. These are chosen to achieve an optimal splitting of the parameter space whilst keeping the bin sizes larger than the average error on stellar masses calculated by GAMA (around 0.1 dex).

These redshift and stellar mass bins are used throughout to combine both the mass-complete G9 radio and X-ray sources, as well as the GAMA09 galaxies themselves (serving as the parent sample). A summary infographic is shown in Fig. 3.1, presenting the numbers and incidences of X-ray AGN, radio AGN and AGN detected at both wavelengths, in the different $M_* - z$ bins.

3.1.2 Measuring AGN incidences as a function of mass-scaled power indicators

As introduced in Chapter 2.4, the observable X-ray and radio luminosities, along with stellar mass estimates from SED fitting can be combined to obtain proxies of the fundamental dimensionless power rates: the specific black hole radiative power (λ_{Edd}) and the specific black hole kinetic power (λ_{Jet}). The analysis of the incidence of radio AGN as a function of mass-scaled jet power indicators is presented here for the first time.

Normalising the luminosity by stellar mass is important as it unmasks correlations with respect to the underlying radiative and kinetic power output distribution. For example, Aird et al. (2012); Bongiorno et al. (2012); Georgakakis et al. (2017); Birchall et al. (2022) and others show that the increasing fraction of X-ray detected AGN with stellar mass is just a selection effect of magnitude-limited surveys being able to detect objects down to lower accretion rates at higher mass for the same luminosity. In other words, looking at the incidence of AGN as a function of just luminosity is degenerate to the high accretion rate, small mass black holes and low accretion rate, large mass black holes.



log(Stellar Mass/[M$_\odot$])	Redshift	# GAMA galaxies	# LOFAR-det.	# eFEDS-det.	# LOFAR & eFEDS det.
11.4 – 12.0	0 < z ≤ 0.285	420	96 (22.9%)	22 (5.2%)	3 (0.7%)
11.4 – 12.0	0.285 < z ≤ 0.4	779	123 (15.8%)	28 (3.6%)	6 (0.8%)
11.2 – 11.4	0 < z ≤ 0.285	1263	111 (8.8%)	33 (2.6%)	6 (0.5%)
11.2 – 11.4	0.285 < z ≤ 0.4	2003	119 (5.9%)	28 (1.4%)	4 (0.2%)
11.0 – 11.2	0 < z ≤ 0.285	2910	72 (2.5%)	47 (1.6%)	4 (0.1%)
11.0 – 11.2	0.285 < z ≤ 0.4	3395	80 (2.4%)	31 (0.9%)	1 (0.0%)
10.6 – 11.0	0 < z ≤ 0.285	10692	81 (0.8%)	136 (1.3%)	0 (0.0%)

Figure 3.1: Infographic showing the incidence of radio AGN, X-ray AGN and AGN detected at both wavelengths, in different stellar mass–redshift bins (colours are as defined above). The legend is in the bottom right (e.g. there are 10692 GAMA09 galaxies in the lowest stellar mass bin, of which 81, 0.8%, are LOFAR detected, etc.). We note that the AGN detected in both radio and X-ray are a subset of the individual pure radio and pure X-ray detected numbers.

The method to calculate the incidence of AGN (valid for all target samples: radio-only, X-ray only, or radio and X-ray samples) is to estimate the confidence intervals on binomial population proportions using Beta distributions (Cameron 2011). In essence, this returns a measure of the fraction of target objects compared to GAMA09 parent galaxies in each given bin. The errors on these values are denoted by the 16th and 84th percentiles (1σ) of the distribution. This method is favoured over others, including for example that of Gehrels (1986), as the Poisson error on population proportions is systematically underestimated for small samples or large samples with extreme population proportions (either very low or very high detection fractions). As seen in Table 2.1 and Fig. 3.1, the GAMA09 sample is large, yet the X-ray and radio detections, especially when split into different bins, are sometimes orders of magnitude less, necessitating the use of Cameron (2011) confidence intervals. Moreover, the effects that the SF properties of the host galaxy can have on the incidences are examined by splitting up the sample (see Ch. 3.3), thus further reducing the statistics. This is also done for the compact and complex radio morphologies.

Once the fraction of galaxies hosting the given target sample of AGN as a function of the different (mass-scaled) parameters have been calculated, a power law is fit to the data in the form of $y = A \times (10^{x-x_0})^B$, using UltraNest[1] (Buchner 2021c). A power law slope (B) and normalisation (A) at a given (log) x-axis value (x_0) can then be obtained and compared across the different stellar mass and redshift bins in order to extract trends.

3.1.3 Accounting for radio and X-ray luminosity incompleteness

Using the information regarding the flux sensitivity of the eROSITA and LOFAR instruments observing the GAMA09 field (see Figs. 2.4 and 2.14), it is possible to apply a correction to the incidence in the bins which are not fully complete in luminosity. We note that a weighting per bin, instead of per source, is the appropriate method here as the incompleteness is a result of the survey limitations, not of the sources themselves.

For example, for a given $L_\mathrm{R} - z$ bin, one can calculate the median luminosity, convert it back to an observed radio total integrated flux and interpolate to find the survey sensitivity at that flux level. The incidence in that bin would then be weighted by a factor of 1/sensitivity (for any sensitivity greater than 50%, otherwise it is considered incomplete and removed). Similarly, for every $L_X - z$ bin, the sensitivity can be directly interpolated from Fig. 2.4, given the median luminosity and redshift. We note that these corrections mainly affect the low $\lambda_\mathrm{Jet,Edd}$ and the low M_* bins at higher redshift.

We also account for the potential missed radio AGN in highly star-forming galaxies, resulting from the 3σ cut in Fig. 2.12 preferentially removing higher mass galaxies. This is because higher stellar mass corresponds to a higher SFR, a consequence of the main sequence of star-forming galaxies (e.g. Speagle et al. 2014, see Ch. 2.3.6). Thus, as one moves to higher SFRs on Fig. 2.12, the sources have to have higher and higher radio luminosities, meaning stronger jets, to systematically dominate over the stellar emission. However, as indicated by the radio luminosity function for radio AGN (Smolčić et al. 2008;

---

[1] `https://johannesbuchner.github.io/UltraNest/example-line.html`

Sabater et al. 2019; Kondapally et al. 2022), higher power jetted systems are less common in the universe and therefore, it gets harder to pass the cut for higher SFRs.

To mitigate this, the incompleteness is accounted for by determining the fraction of sources which could be missed using a combination of the main sequence presented in Speagle et al. (2014) and the radio AGN cut from Best et al. (2023b). Firstly, the $\log M_*$ and $\log Q/L_{\rm Edd}$ bin limits are used to calculate a parallelogram in $\log M_* - \log L_{\rm R}$ space covered by those limits (e.g. top right corner of this parallelogram would have an $\log L_{\rm R}$ value calculated by combining the maximum bin limits of $\log M_*$ and $\log Q/L_{\rm Edd}$). Secondly, a linear relation in $\log M_* - \log L_{\rm R}$ space is computed, combining the radio AGN 3σ cut (relating $\log L_{\rm R}$ and $\log {\rm SFR}$) and the MS equation (relating $\log {\rm SFR}$ to $\log M_*$, within a given redshift bin). Anything with a radio luminosity below this linear relation is incomplete as it bridges into the star-forming MS and means that radio AGN below that line are missed. Therefore, the fraction of sources missed within a given $\log M_*$ and $\log Q/L_{\rm Edd}$ bin is simply the geometric ratio of the parallelogram area below vs. above the linear relation. The weighting, applied by multiplying the incidence in a given bin, is then the inverse of this missed fraction. We note that this correction only affects the lowest $L_{\rm R}/M_*$ or $\lambda_{\rm Jet}$ bins (crosses on Figures 3.7 and 3.6), in some cases being <50% complete and therefore removed.

Faint X-ray AGN in highly star-forming galaxies may equally be missed (recall the selection in Fig. 2.2). However, as pointed out in Merloni (2016), the X-ray emission for a typical $10^8\ M_\odot$ AGN in $10^{10.5}\ M_\odot$ main sequence star-forming host dominates over star-formation for $\lambda_{\rm Edd} > 10^{-5}$. The G9 X-ray AGN do not extend to such low $\lambda_{\rm Edd}$ and so any incompleteness from this effect would be negligible in the context of this analysis.

3.2 Incidence of eFEDS X-ray AGN

Figure 3.2 shows the fraction of GAMA09 galaxies hosting eROSITA-eFEDS detected X-ray AGN as a function of $\lambda_{\rm Edd}$ in different stellar mass and redshift bins. We note that the stringent X-ray luminosity completeness limits leave too few sources to split the X-ray incidences into AGN residing in quiescent versus star-forming galaxies, and therefore, they are combined.

As seen in past studies (e.g. Aird et al. 2012; Bongiorno et al. 2012), the X-ray incidences across the wide range of $\log(M_*/M_\odot)$, from 10.6 to 12.0, are remarkably similar. Regardless of the stellar mass, the X-ray incidence depends on the value of $\lambda_{\rm Edd}$, with higher accretion rate AGN having a lower incidence (rarer) than those at lower accretion rates. Specifically, around 0.1% and 1% of galaxies host an X-ray accreting at $\lambda_{\rm Edd} \sim 0.1$ and ~ 0.01, respectively. This has been attributed to a universal, stellar mass-invariant (and thereby black hole mass-invariant?) fuelling and triggering mechanism present in X-ray detected AGN. It is associated with a universal underlying $\lambda_{\rm Edd}$ distribution with power law slope -0.65, independent of host galaxy stellar mass, that evolves to higher normalisations with increasing redshift, as shown in Fig. 3.2.

Our results shown in Fig. 3.2 serve as a validation of the methods described in this

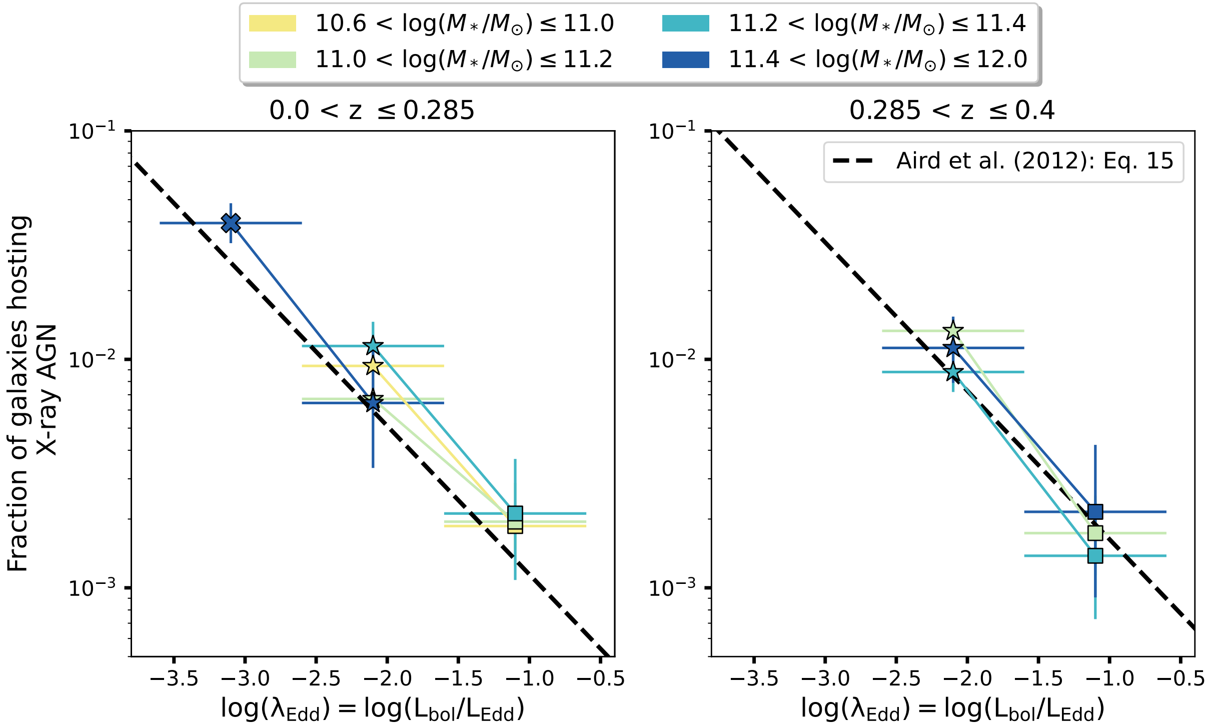


Figure 3.2: Fraction of GAMA09 galaxies hosting eROSITA-eFEDS-detected X-ray AGN as a function of λ_{Edd} in different stellar mass (yellow, green, teal, blue) and redshift (two panels) bins. The results agree well with those of Aird et al. (2012) (black dashed lines), corroborating the idea that there is a mass-invariant triggering and fuelling mechanism at play in X-ray AGN.

Chapter and a proof of concept that the soft response of eROSITA (with thorough consideration for completeness) is able to recover past results obtained mainly with harder X-ray instruments, less susceptible to absorption (e.g. *XMM-Newton, Chandra*), at least for the low-redshift samples probed here.

3.2.1 Unobscured X-ray AGN Incidences

We note that relying solely on an unobscured AGN selection introduces absorption-related incompleteness, which becomes significant at the lowest values of λ_{Edd}. Fig. 3.3 shows the unobscured (i.e. assuming $\log N_H < 21$ cm^{-2}) sensitivity corrections (orange curves), over-plotted on the G9 X-ray AGN sample. It is seen that many more X-ray AGN are now $>$ 90% complete, as would be expected from eROSITA's soft selection. However, when considering the fraction of galaxies hosting such X-ray AGN as a function of λ_{Edd} (Fig. 3.4) there is a clear lack of detections at the lower accretion rate end. This is due to the effects of obscuration affecting the lower luminosity population more, making them drop out of the sample. Fig. 3.4 is interesting to show the levels of λ_{Edd} which start to become significantly affected by obscuration.

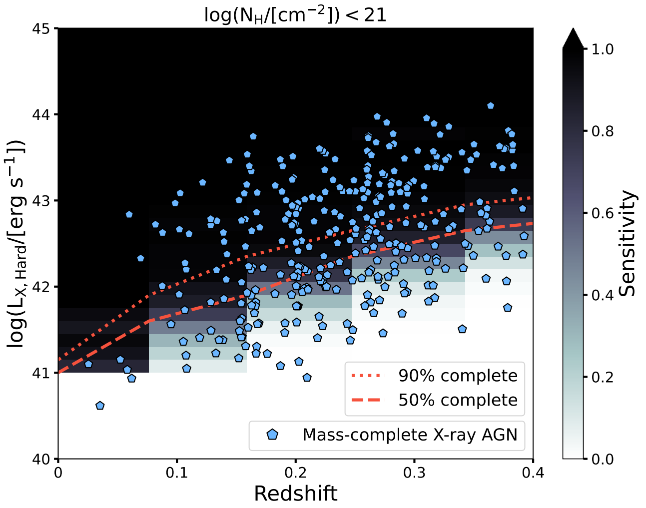


Figure 3.3: Intrinsic hard (2 – 10 keV) X-ray luminosity versus redshift of the G9 X-ray AGN (pentagons). A sensitivity grid for only unobscured ($\log N_H < 21$ cm^{-2}) sources is plotted in the background from simulations done by Liu et al. (2022e) and used to compute the 50% and 90% X-ray luminosity completeness limits, respectively.

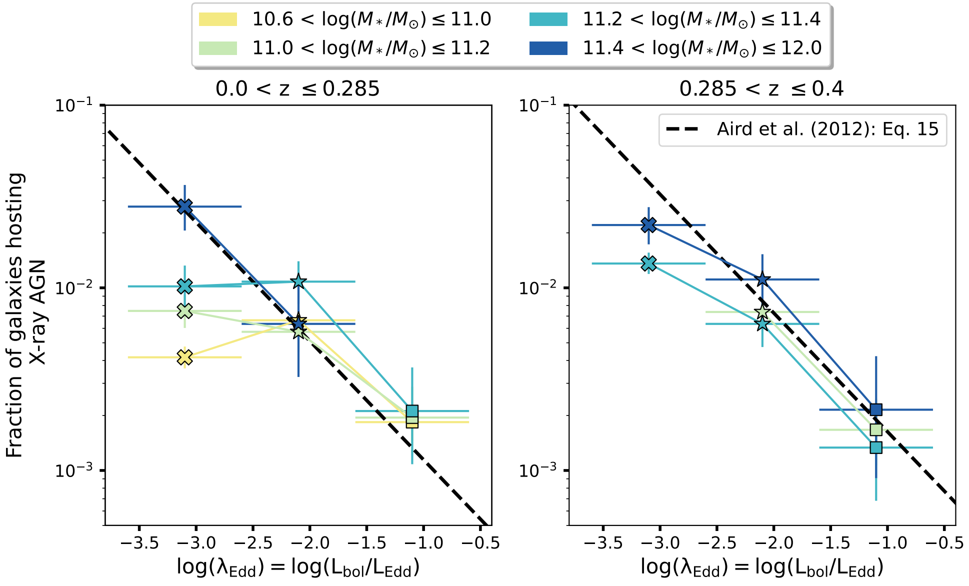


Figure 3.4: Fraction of GAMA09 galaxies hosting X-ray AGN as a function of λ_{Edd} in different stellar mass and redshift bins, using only a soft selection (meaning that the X-ray luminosity is corrected to be complete for $\log N_H < 21$ cm^{-2}). As expected, the lowest λ_{Edd} bins are incomplete, compared to the black dashed line, due to obscuration.

3.3 Incidence of radio AGN

3.3.1 Incidence of radio AGN as a function of stellar mass

Figure 3.5 shows that the fraction of GAMA09 galaxies hosting (detectable) radio AGN is a strongly increasing function of stellar mass, in different redshift and cumulative luminosity bins. For $\log(L_R/[\mathrm{W\ Hz}^{-1}]) \geq 24$, around 10% of galaxies host radio AGN at the highest masses $\log(M_*/M_\odot) > 11.5$, whilst at $\log(M_*/M_\odot) = 11.1$ the prevalence is only around 1% (up to $z < 0.4$). These results agree well with Sabater et al. (2019), as shown by the orange shaded regions over-plotted onto Fig. 3.5, taken from their Fig. 5 (left panel). Of course, as for the case of X-ray selected AGN, this strongly increasing radio AGN incidence as a function of stellar mass is again a selection effect resulting from the underlying λ_{Jet} distribution and our survey flux limits, which is why is it essential to probe quantities normalised by stellar mass (see below).

Unfortunately, it is difficult with present data to comment on the redshift evolution of radio AGN incidence as a function of radio luminosity (e.g. Smolčić et al. 2017), yet a weak increasing trend in normalisation with redshift is apparent, as expected from past studies.

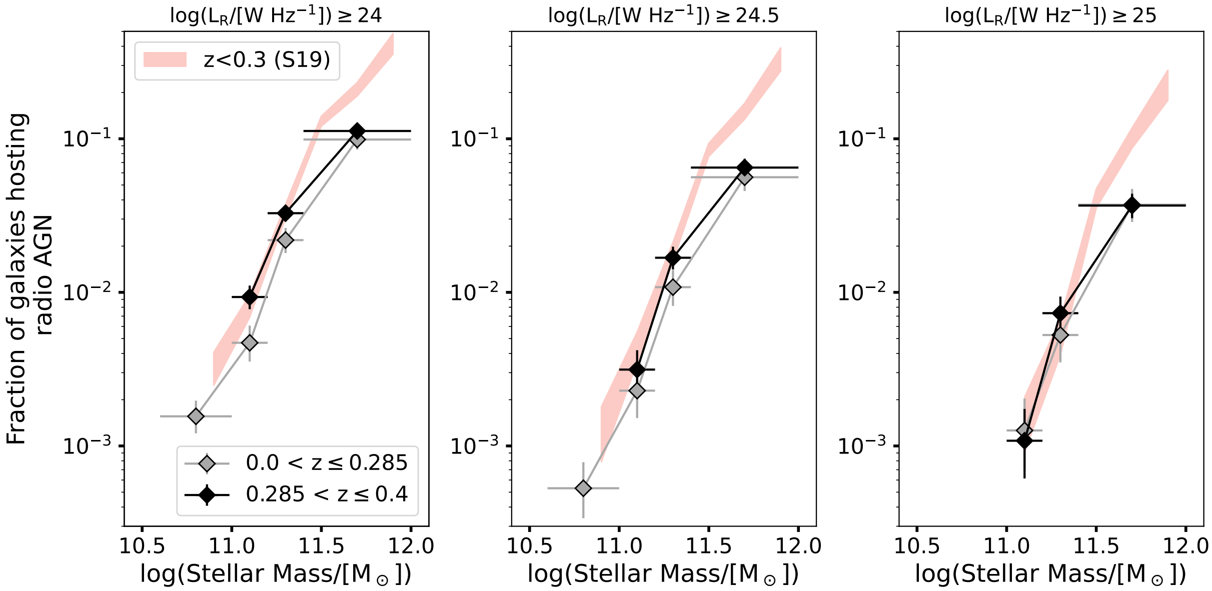


Figure 3.5: Fraction of GAMA09 galaxies hosting both complex and compact radio AGN as a function of stellar mass, in different redshift (light grey and black) and luminosity (panels) bins. A strong increase in the fraction of detected radio AGN with increasing stellar mass is observed. Orange shaded curves mark the results from Fig. 5 of Sabater et al. (2019) in the redshift range $0 < z < 0.3$.

3.3.2 Incidence of radio AGN as a function of λ_{Jet} for compact radio morphologies

Mass-scaled power indicators in the radio regime are not as straightforward as the X-ray regime where $\lambda_{\mathrm{Edd}} \propto L_{\mathrm{X}}/M_*$. This is because the radio luminosity does not scale linearly with jet power (as given by observations, recall Ch. 2.4). Nevertheless, it is still informative to show the radio AGN incidence as a function of the simple observable L_{R}/M_*, as done in Fig. 3.6. The average power law slopes are constant around -0.8. A mass dependence is clearly seen by the increasing power law normalisations with stellar mass. Quantitatively, comparing the power laws at $\log(L_{\mathrm{R}}/M_*) = 13$, the highest mass galaxies are 13.5 and 4.2 times more likely to host radio AGN than the lowest mass galaxies in the low and high redshift bins, respectively. However, one must note that the observable L_{R}/M_* parameter is an indirect (and complex) tracer of the underlying jet power.

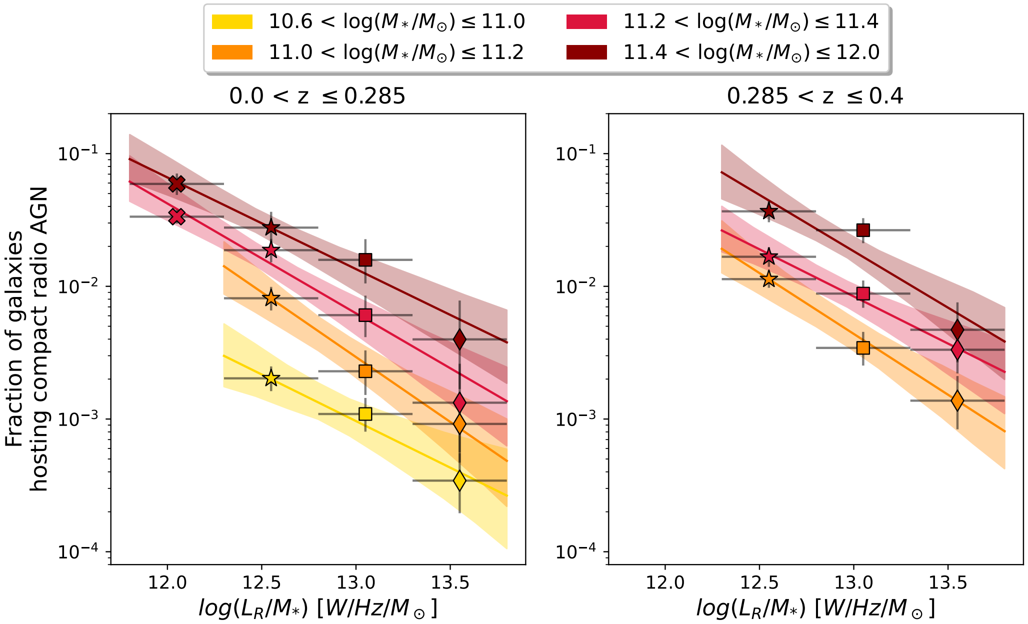


Figure 3.6: Fraction of GAMA09 galaxies hosting compact radio AGN as a function of radio luminosity normalised by stellar mass (L_{R}/M_*), in different redshift (panels) and stellar mass bins (colours). Power laws, and associated errors, are fit to each stellar mass bin and are plotted with the corresponding colour.

To examine the physical nature of jet powering, Fig. 3.7 shows the fraction of GAMA09 galaxies hosting compact radio AGN as a function of the specific black hole kinetic power, λ_{Jet}, in different redshift and stellar mass bins. Power law slopes and normalisation of the fit to the data points are shown in Fig. 3.8 and summarised in Table 3.1. A linear parameterisation of the power law normalisation as a function of stellar mass can be given by $A = 10^{-3}\ [m \times (\log M_* - 11.4) + c]$, where (m, c)=$(12.8^{+4.2}_{-4.7}, 7.63^{+1.3}_{-1.4})$ and $(13.9^{+8.2}_{-6.8}, 10.7^{+1.9}_{-2.0})$ for the low and high redshift bins, respectively.

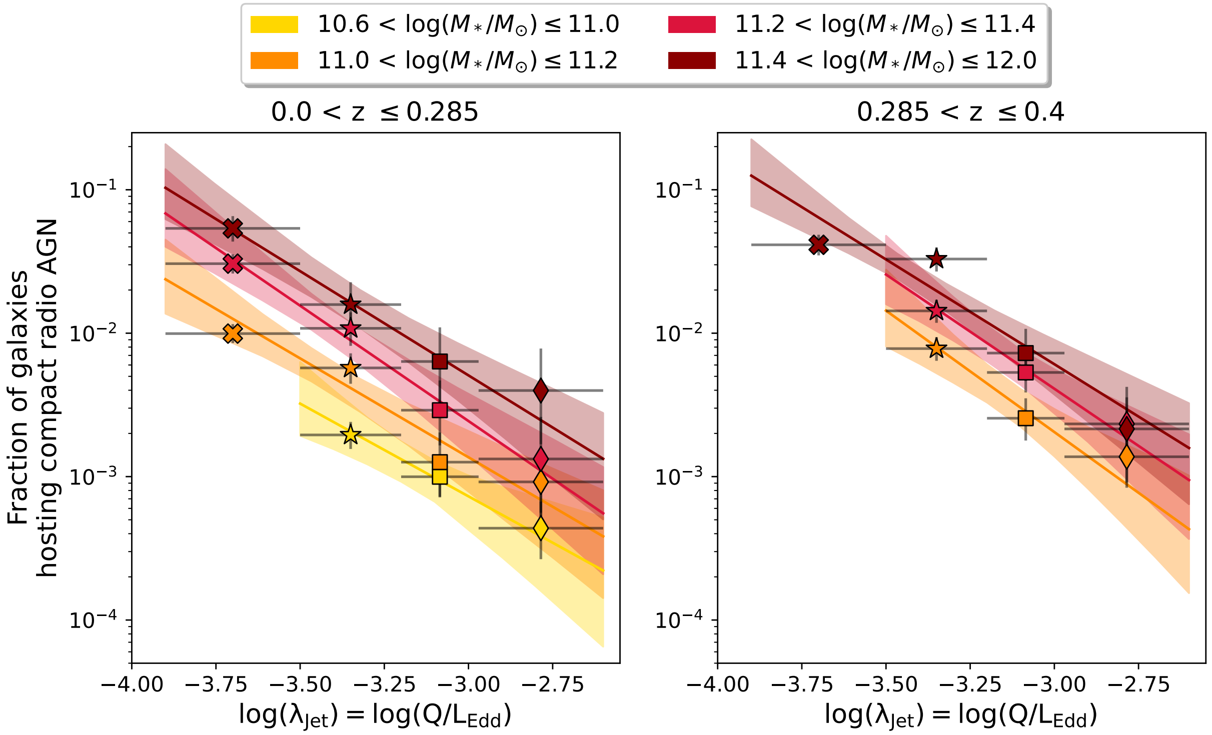


Figure 3.7: Fraction of GAMA09 galaxies hosting compact radio AGN as a function of the specific black hole kinetic power, $\lambda_{\rm Jet}$, in different stellar mass and redshift bins. The power law fit slope and normalisation values are shown in Fig. 3.8 and Table 3.1.

Similarly to the X-ray AGN incidence, the radio AGN incidence decreases as $\lambda_{\rm Jet}$ increases, because higher radio power objects become less common at all masses in the sample. On the other hand, there is a non-zero mass dependence, shown by the increasing power law normalisations with stellar mass, that is not present in Fig. 3.2. In fact, at $\log \lambda_{\rm Jet} = -3.25$, the highest mass galaxies are 7.8 and 2.6 times more likely to host radio AGN compared to the lowest mass bins in the low and high redshift bin, respectively. Possible reasons why the incidence of radio AGN shows this mass dependence, along with the caveats in the calculation of Q are discussed in Ch. 3.5.

An important takeaway from Fig. 3.7 is that the slopes of the observed power law distributions for all stellar mass ranges probed are the same, with a value equal to about -1.5. This shows clearly that it is not only the massive galaxies that host powerful jetted AGN, nor do only the low mass galaxies host low-power jets.

There is also a slight tendency for increased detection fractions with increasing redshift (see increasing intercept values in Fig. 3.8), possibly relating to an increased characteristic $\lambda_{\rm Jet}$ distribution at different epochs. Nevertheless, a larger redshift range would be needed to probe any redshift dependence further.

We also study the incidence of radio AGN in quiescent versus star-forming galaxies. Indeed, as mentioned in Ch. 1.6, the differences between quiescent and star-forming hosts,

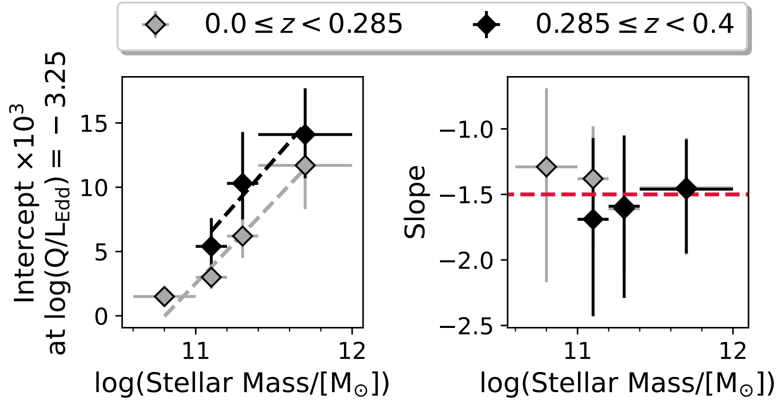


Figure 3.8: Results of a power law fit, $y = A \times (10^{x-x_0})^B$, to all the different mass, redshift bins present in Fig. 3.7. The left panel plots the normalisation ($A * 10^3$) as the y-intercept at $x = x_0 = -3.25$; the right panel plots the slope (B). The slope is consistently around -1.5 for all M_* values (red dashed line). The normalisations show a slight mass dependence of the incidence, with some redshift evolution. Light grey and black dashed lines show the result from a linear fit (parameters listed in Table 3.1).

such as temperature and fraction of gas, could have direct effects on the powering of jets (see e.g. Kondapally et al. 2022). At each λ_{Jet} value, we sample 1000 points in the range of the 1σ uncertainty on the incidence of quiescent and star-forming radio AGN separately. We then find the average ratio between the two, with the standard deviation on the mean giving the 1σ error. Fig. 3.9 shows the ratio of the measured incidence of compact[2] radio AGN in star-forming versus quiescent galaxies in the same redshift, stellar mass and λ_{Jet} bins as above. It can be seen that the fraction of quiescent galaxies hosting radio AGN is similar to that of star-forming galaxies. In general, there is no evidence of a suppressed radio AGN incidence in star-forming galaxies (with the exception of the lowest λ_{Jet} sources in the low redshift bin). Importantly, this indicates that, contrary to older findings (e.g. Matthews et al. 1964; Dunlop et al. 2003; Best et al. 2005; Hickox et al. 2009), radio AGN are not predominantly hosted by ‘red and dead’ giant elliptical galaxies, when the incidences are properly computed from complete samples. Indeed, the LOFAR survey and availability of ample multi-wavelength data is finally enabling the field of radio astronomy to probe radio AGN in even the most star forming galaxies, by allowing a better understanding of the origin of the radio emission.

However, due to the still limited sample size, it is not possible within the scope of this investigation to further probe the differences in jet powering resulting from the host galaxy properties (see e.g. Kondapally et al. 2022; Aird et al. 2019; Birchall et al. 2023, for work on this topic in the radio and X-ray regimes). Therefore, the two samples are combined, as already done for Fig. 3.7, in order to increase sample statistics.

---

[2] We note that the sample of complex morphology star-forming hosts is too small to robustly compare to its quiescent equal.

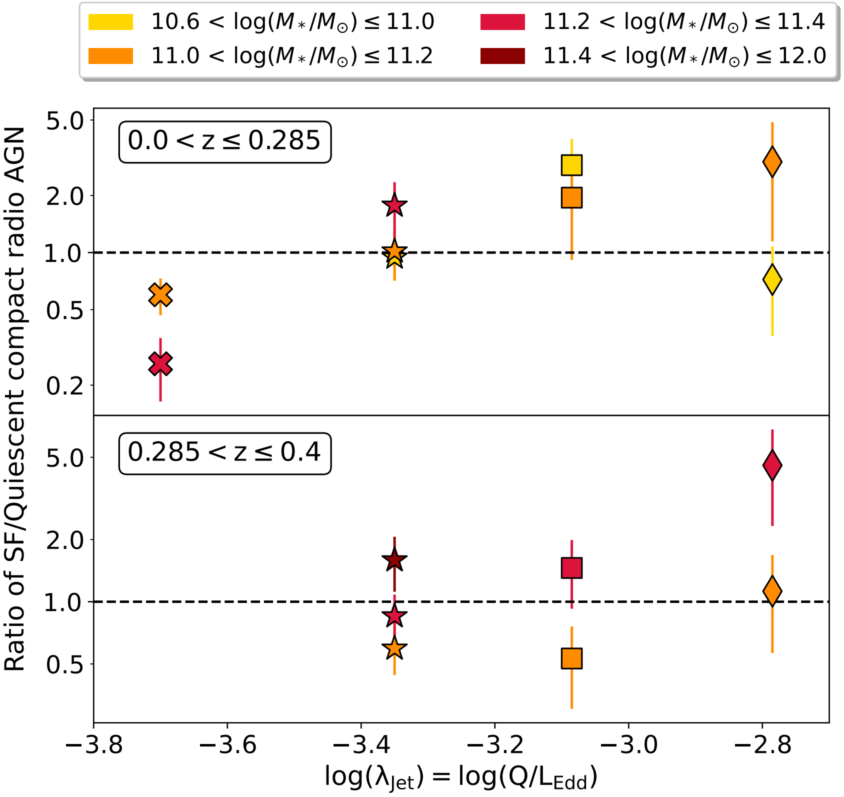


Figure 3.9: Ratio of the measured incidence of radio AGN in star-forming to quiescent galaxies (colours and symbols as above), showing that the fraction of quiescent galaxies hosting radio AGN is similar to that of star-forming ones.

3.3.3 Incidence of radio AGN as a function of $\lambda_{\rm Jet}$ for both compact and complex radio morphologies

Until now we have only studied the incidence of compact radio sources, as they provide the largest statistics. Fig. 3.10 shows instead the incidence of radio AGN with *both* compact and complex radio morphologies, as a function of $\lambda_{\rm Jet}$. An additional high $\lambda_{\rm Jet}$ bin (pentagon) is added as complex sources reach higher radio luminosities than the compact sample (recall Fig. 2.12). Focusing first on the plotted markers (all compact and complex), the main difference, compared to Fig. 3.7, is that the incidences are boosted at high $\lambda_{\rm Jet}$ values (diamonds), especially for the higher stellar mass bins where our sample contains a larger number of complex radio AGN. There is also an increased mass dependence of the incidence compared to the compact-only case. This implies that, although both compact and complex radio AGN are in general more frequently detected in more massive galaxies, it is the complex sample preferentially driving this mass dependence.

Interestingly, when plotting the subset of compact and complex sources satisfying the condition that their radio emission is modelled only by a single Gaussian (`S_code`=S), the

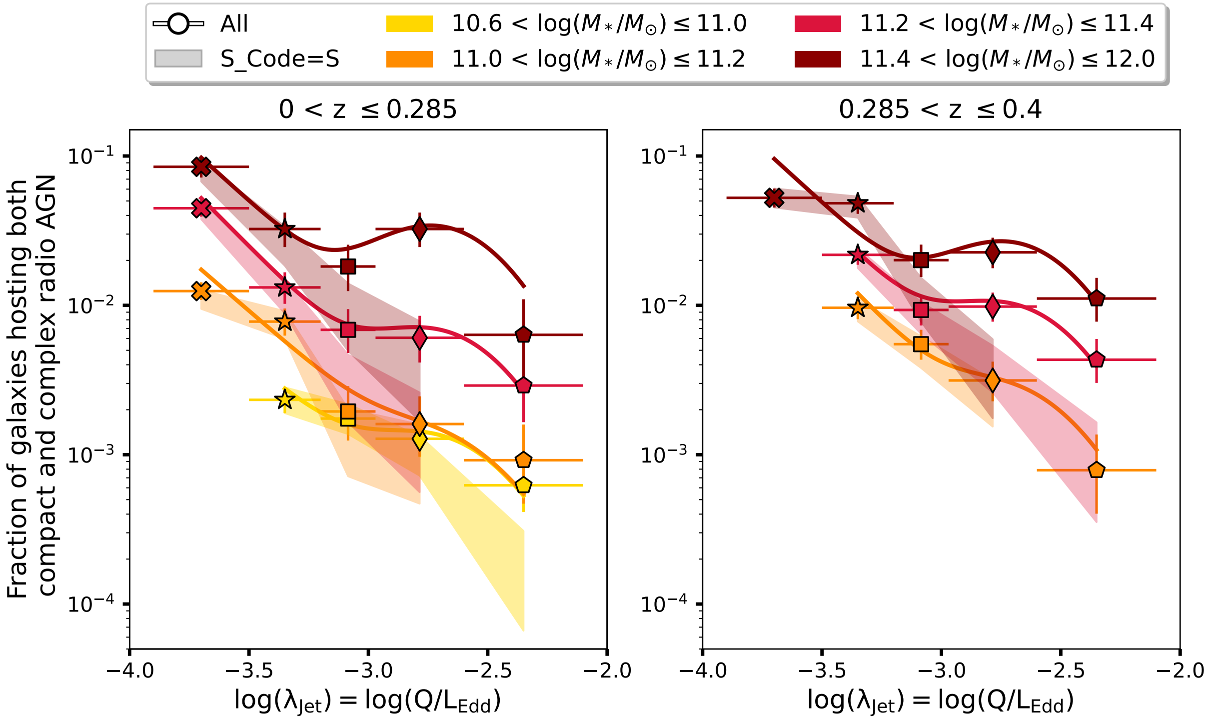


Figure 3.10: Fraction of GAMA09 galaxies hosting both compact and complex radio AGN as a function of the specific black hole kinetic power, λ_{Jet}, in different stellar mass and redshift bins. The markers plot all the sources, while the shaded region shows the incidences of just the `S_code`=S sources. In contrast to the compact-only case, a larger mass dependence is present, as well as a boosted incidence at high λ_{Jet} (diamonds) values, indicating a jet power dependence. A power law plus Gaussian fit is used to approximate the data (solid curves), for which fit parameters are presented in Table 3.1.

upturn at high jet powers disappears, resembling the simple power law distributions seen in compact radio AGN (Fig. 3.7). We note that a significant fraction, 52% (144/278), of complex (mass-complete) radio AGN have a single Gaussian component. These are radio sources which are morphologically simple, but too large for being classified as compact.

We approximate the compact and complex radio AGN incidence with a power law plus a Gaussian, marked by solid curves on Fig. 3.10. The power law slopes are fixed to the individual best-fit values from the compact-only incidence (Fig. 3.8) and the Gaussian is centred at $\mu = \log \lambda_{\mathrm{Jet}} = -2.71$ with fixed width $\sigma = 0.19$. Both power law and Gaussian normalisations are left free in the fit. All fit parameters are presented in Table 3.1.

Fig. 3.11 shows the histogram of the different subsets of compact (light blue filled), complex with `S_code`=S (black, unfilled), complex with `S_code`=M (magenta, hatched) and complex `S_code`=M sources with physical sizes $R_{\mathrm{kpc}} > 60$ kpc (dark blue filled). Complex (i.e. non-compact) sources show a clear bimodal distribution in both radio luminosity (top right) and λ_{Jet} (bottom right). Therefore, we conclude that the sources modelled with

	Compact-only				Compact and Complex			
	Pow norm.		Pow slope		Pow norm.		Gauss norm.	
$\log(M_*/M_\odot)$	z1	z2	z1	z2	z1	z2	z1	z2
10.6 – 11.0	$1.5^{+0.5}_{-0.6}$	–	$-1.29^{+0.88}_{-0.6}$	–	$2.0^{+0.5}_{-0.6}$	–	$1.0^{+0.6}_{-0.7}$	–
11.0 – 11.2	$3.0^{+0.9}_{-0.9}$	$5.4^{+1.5}_{-2.2}$	$-1.38^{+0.52}_{-0.4}$	$-1.69^{+0.74}_{-0.62}$	$4.2^{+1.1}_{-0.9}$	$8.1^{+2.3}_{-2.4}$	$0.8^{+1.2}_{-1.5}$	$2.0^{+1.7}_{-2.0}$
11.2 – 11.4	$6.2^{+1.7}_{-2.1}$	$10.3^{+2.8}_{-4.0}$	$-1.61^{+0.51}_{-0.39}$	$-1.59^{+0.7}_{-0.54}$	$9.8^{+2.9}_{-3.2}$	$15.4^{+4.3}_{-5.4}$	$5.6^{+3.3}_{-4.0}$	$8.2^{+4.2}_{-5.0}$
11.4 – 12.0	$11.7^{+3.4}_{-4.0}$	$14.1^{+3.4}_{-3.6}$	$-1.45^{+0.51}_{-0.38}$	$-1.46^{+0.49}_{-0.38}$	$21.7^{+5.8}_{-6.0}$	$21.1^{+5.2}_{-4.5}$	$30.5^{+10.9}_{-19.6}$	$23.2^{+10.7}_{-16.4}$

Table 3.1: Fit results for the power law and power law plus Gaussian trends, fit to the incidence of radio AGN in different stellar mass and redshift bins ($z1$ and $z2$ for the low and high redshift bins, respectively). The power law norm (pow norm.) is the value of the intercept $\times 10^3$ at $\log \lambda_{\mathrm{Jet}} = -3.25$. Gauss norm. represents the normalisation of the Gaussian function, with an additional multiplicative factor of 1000. These values are used to plot the trends in Figs. 3.7 and 3.10.

multiple Gaussians and sources with large physical sizes drive the upturn in the incidence seen at higher jet powers. Possible explanations for this are discussed in Sect. 3.5.2.

Lastly, since all mass-complete sources were visually inspected, it is possible to also give an estimate of the prominence of FRIIs in the sample (flagged as `FRII_flag`=1). Table 3.2 shows the percentage of total (compact and complex) radio AGN with FRII-like morphologies in different stellar mass and redshift bins. Percentages are only shown for the incidence of the filled diamond λ_{Jet} sources (♦ on Fig. 3.10 where the boosted incidence is present) and for the extended range of $-3 < \log \lambda_{\mathrm{Jet}} \leq -1.5$, which includes all of the most powerful FRIIs of the sample. It can be seen that FRII objects are in fact not so ‘rare’ and make up around 10 – 30% of the high jet power sources at $z < 0.4$. There does not seem to be a distinct trend with host galaxy stellar mass, although a larger sample with higher resolution would be needed to test this.

	$0 < z \leq 0.285$		$0.285 < z \leq 0.4$	
$\log(M_*/M_\odot)$	♦	◇	♦	◇
10.6 – 11.0	$11.7^{+10.0}_{-6.6}$	$21.9^{+9.7}_{-8.0}$	-	-
11.0 – 11.2	$31.4^{+21.0}_{-16.7}$	$16.2^{+13.2}_{-9.0}$	$23.6^{+13.7}_{-10.7}$	$21.2^{+10.7}_{-8.5}$
11.2 – 11.4	$32.1^{+16.9}_{-14.2}$	$32.6^{+12.9}_{-11.3}$	$13.1^{+8.5}_{-6.2}$	$22.7^{+8.3}_{-7.0}$
11.4 – 12.0	$25.6^{+12.4}_{-10.2}$	$30.9^{+11.3}_{-9.9}$	$25.5^{+10.9}_{-9.1}$	$26.1^{+8.6}_{-7.5}$

Table 3.2: Percentage of total (compact and complex) radio AGN that have FRII-like morphologies. The statistics are only shown for the incidence of the filled diamond λ_{Jet} sources (♦ on Fig. 3.10 where the boosted incidence is present) and for the range $-3 < \log \lambda_{\mathrm{Jet}} \leq -1.5$ (unfilled diamond, ◇) in each stellar mass and redshift bin.

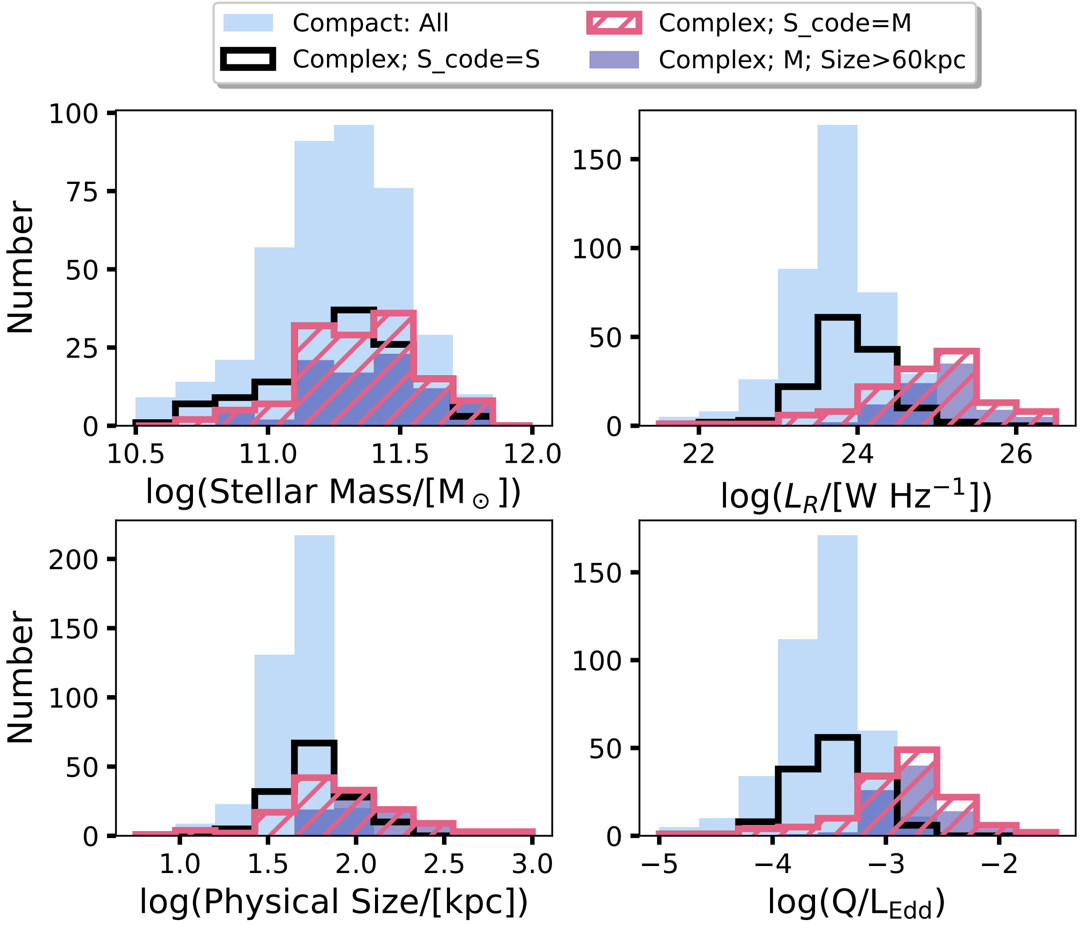


Figure 3.11: Histograms showing the stellar mass, radio luminosity, physical size and $\lambda_{\rm Jet}$ distributions of different subsets of radio AGN: compact (light blue filled), complex `S_Code`=S (black, unfilled), complex `S_Code`=M (magenta, hatched) and complex `S_Code`=M sources with physical sizes $>$ 60 kpc (dark blue filled). This shows that the sources responsible for the boosted radio AGN incidences at high jet powers are the ones best modelled by multiple Gaussian components and large physical sizes.

3.4 Incidences of both X-ray and radio AGN

In general, as host galaxy stellar mass increases, it gets increasingly likely to host both radio and X-ray AGN (as found in, e.g. Best et al. 2005; Smolčić et al. 2009; Brusa et al. 2009). However, at $11.4 < \log(M_*/[M_\odot]) \leq 12$, there is a factor $\sim$4 higher probability to host a radio AGN compared to an X-ray AGN, in both the low and high redshift bins.

Moreover, we can combine the limited number of radio and X-ray detected AGN (same as Fig. 3.2) to calculate the incidence of both radio and X-ray AGN as a function of $\lambda_{\rm Edd}$, as shown in Fig. 3.12. The mass bins are now combined into one, keeping all mass completeness requirements fulfilled. Only the 18/24 X-ray and radio AGN which have radio luminosity in excess of the 95% completeness threshold (dashed line on Fig. 2.14) are included.

Again it is seen that the pure X-ray detected sources (light blue pentagons) follow well the Aird et al. (2012) power law. The incidence of radio and X-ray AGN (light green diamonds) is around 10% less than that of the pure X-ray AGN, in line with the generally accepted fraction of 'radio-loud' AGN in X-ray surveys (see e.g. Kellermann et al. 2016).

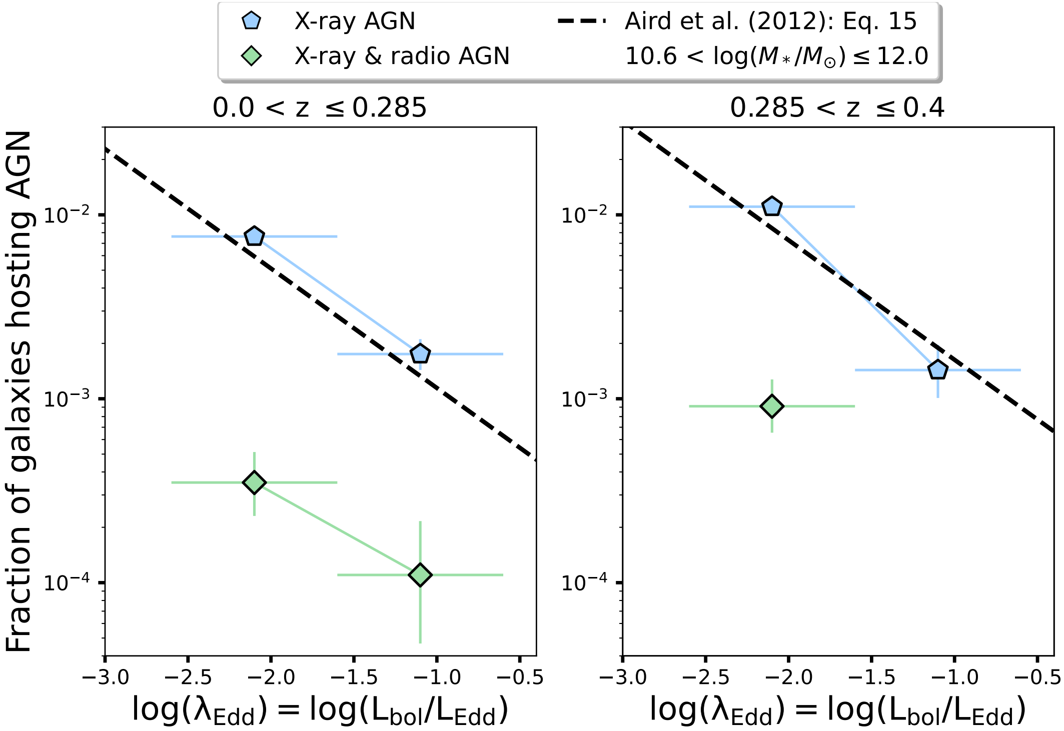


Figure 3.12: Fraction of GAMA09 galaxies hosting purely X-ray-detected AGN (blue) and both X-ray and radio AGN (green) as a function of $\lambda_{\rm Edd}$. The stellar mass, radio luminosity and X-ray luminosity are ensured for completeness. The pure X-ray fractions follow well the Aird et al. (2012) trend, whilst the radio and X-ray AGN lie an order of magnitude below it.

3.5 Discussion

Overall, we show that a universal AGN triggering and fuelling mechanism is present for X-ray detected AGN (as also found by several previous studies), yet this phenomenon does not seem to translate into a fully mass-invariant jet power mechanism, as seen by the incidence of radio AGN as a function of $\lambda_{\rm Jet}$ in Fig. 3.7. It even shows an additional jet power dependence when including complex radio emitters (Fig. 3.10). This section discusses the possible reasons for these differences, given the caveats and limitations of the surveys, and asks whether there even should be a universal jet powering mechanism present in AGN. Firstly though, the GAMA survey completeness to AGN and its reliability to estimate stellar masses is vetted.

3.5.1 Incompleteness due to the lack of bright quasars in GAMA

GAMA is inherently a galaxy survey and bright Type I QSOs may have been missed (or de-selected; see Section 3.1 of Baldry et al. 2010), therefore it is important to consider the implications this might have on our measures of AGN incidences. Comparat et al. (2022) estimate the fraction of missed eROSITA eFEDS X-ray extra-galactic point sources (AGN) in GAMA09, in the redshift range $0.05 < z < 0.3$, by firstly taking their Legacy Survey DR8 (LS8) counterparts, with spectroscopic or good quality photometric redshifts (Salvato et al. 2022a), within the GAMA09 footprint. Then, after matching to the GAMA09 galaxy catalogue, they find that 88.8% of the X-ray sources with an LS8 r-band magnitude < 19.8 have a GAMA09 counterpart, meaning that around 10% of X-ray bright QSOs may be missed. This is in line with the known 10% incompleteness for GAMA galaxies at $r < 19.77$ mag and is accounted for in Ch. 2.2.3. A much smaller fraction of missed radio AGN is expected, considering their dominant kinetic, rather than radiative emission.

Meanwhile, for the brighter objects that are present in the GAMA sample, it could be questioned whether the stellar mass estimates are robust and not overestimated due to the AGN contamination (there is no AGN component in the SED fitting for GAMA galaxies). However, we show in Fig. 3.13 that this is not the case for our samples. We compare the GAMA derived stellar masses of the G9 X-ray AGN to the ones from Hyper Supreme-Cam (HSC) obtained for a subset of the eFEDS X-ray sources (which have better AGN and host galaxy decomposition, Aihara et al. 2018; Li et al. 2023). Using the ∼300 matched objects, the stellar masses agree well, as shown in Fig. 3.13 (bottom), with a slight systematic offset of around 0.1 dex (constant with mass) for the GAMA derived measurements, well within the stellar mass bins used in this work (note that both HSC and GAMA use the Chabrier 2003 IMF). Thus, this is not a dominant uncertainty, especially not for the radio results as those objects are not expected to be hosted by bright quasars (corroborated by none of the G9 radio AGN lying within the ‘WISE’ wedge of luminous AGN; Mateos et al. 2013).

Thorne et al. (2022) confirm that the addition of an AGN component to the SED fitting has no significant impact on the GAMA derived stellar masses. Fig. 3.13 (top) shows the excellent agreement between M_* measurements from Thorne et al. (2022) and the ones used in this work for the G9 radio AGN. However, Thorne et al. (2022) find that stellar properties such as SFRs may be overestimated in cases where an AGN component is significant but not considered. This could potentially lead to an underestimate of radio-excess AGN, although it is not deemed a large impact for the same reasons as outlined above.

3.5.2 Caveats in the inference of jet power

Jet power is a notoriously difficult quantity to estimate accurately. Even though the most widely used relation is applied in this Chapter, there are numerous caveats that must be

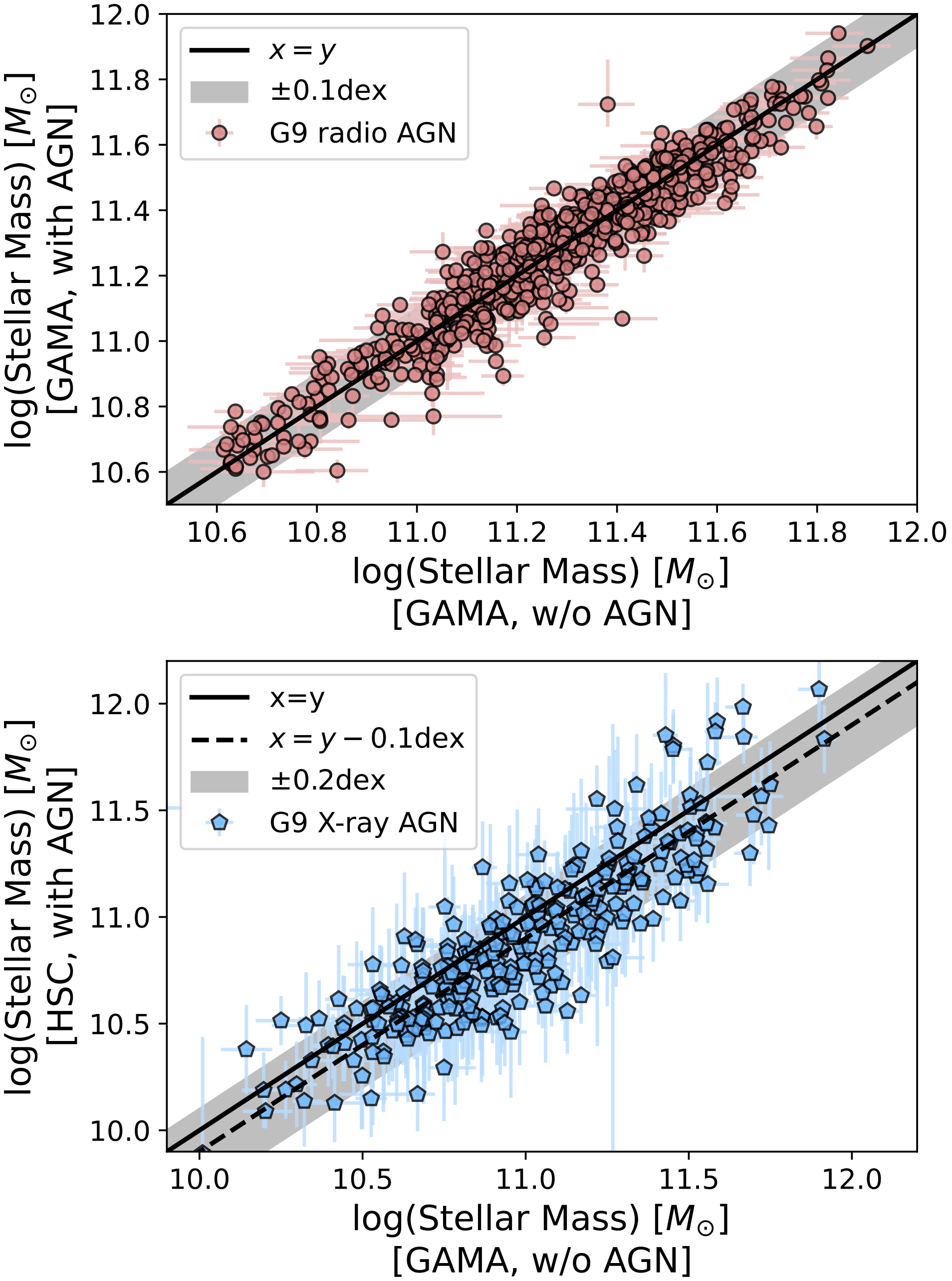


Figure 3.13: Top: Comparison of GAMA M_* measurements with and without AGN component in the SED fitting for the G9 radio AGN sample (with 1σ errorbars). Bottom: Comparison of M_* derived from GAMA without AGN component versus those derived from HSC with AGN-host decomposition for a subset of G9 X-ray AGN.

discussed. However, in Chapter 4 we re-examine in detail the incidence of radio AGN as a function of λ_{Jet} for compact and complex sources, separately, and find that there are still striking differences, which cannot simply be explained by potentially biased jet power measurements.

The $Q - L_R$ relation

Firstly, as stated in Hardcastle et al. (2019), a single conversion from radio luminosity to jet power is inherently flawed as, although jet power could be constant in time, radio luminosity by default cannot be: it must start from zero (and is predicted to decay at late times). Use of such relations requires an assumption to be made on the type of radio source it has been calibrated to, usually large sources in relatively rich environments.

Secondly, there exist a large number of radio luminosity to jet power conversions in the literature, not only with (slightly) different normalisations but also with vastly varying power law slopes, ranging from 0.4 (Bîrzan et al. 2004; Best et al. 2006, 2007) all the way to 1 (Hardcastle et al. 2019, see their Fig. 12 for a visual comparison with respect to the simulated jet powers). The slope of the $Q - L_R$ relation has a direct impact on the mass dependence of the incidence, whereby the more sublinear the slope is, the more mass-invariant the incidence becomes.

Thirdly, some works predict jet power conversions to depend on additional parameters such as radio spectral index (Godfrey & Shabala 2013), proton content (Croston et al. 2018) or spin (see below). Heinz & Sunyaev (2003) and Merloni et al. (2003) derive, for a self-similar freely expanding jet coupled to a radiatively inefficient accretion flow, $L_{\text{Kin}} \propto L_R^{12/(17+8\alpha)}$, where L_{Kin} is the kinetic luminosity, equivalent to Q here. Thus, jet power scaling depends on the observed radio spectral index and this acts to flatten the $Q - L_R$ relation for steeper α values. This may be important as a constant $\alpha = 0.7$ is used here, which better describes larger, more diffuse sources (typically in the 'complex sample). Meanwhile, compact sources with synchrotron self-absorbed cores have much flatter $\alpha \sim 0$, causing a potentially underestimated jet power. However, not all the G9 radio AGN sources in our sample can be associated with a jet coupled to a low accretion rate, radiatively inefficient mode of accretion (see Fig. 2.18), so the applicability of the Heinz & Sunyaev (2003) scaling for the full sample is questionable.

Moreover, the scaling we adopt from Eq. 2.9 (Heckman & Best 2014) is a cavity-based relation, making use of the minimum energy condition to find the energy needed to inflate the radio lobes to their observed volume, calibrated with pressure of the surrounding medium estimated from X-ray observations. These are observations which have been conducted mainly on the relatively few nearby, large cluster AGN with deep X-ray and radio coverage. Therefore, they may not be accurate in describing the jet power of the large majority of low-luminosity ($\log(L_R/[\text{W Hz}^{-1}]) < 25$), compact sources. In fact, Eq. 2.9 may underestimate the jet power for these sources if there are diffuse, large-scale lobes that are too faint to detect. Such surface brightness limitations in LOFAR are a known observational bias (see e.g. Hardcastle et al. 2019) but we do not attempt to correct for this here. Croston et al. (2019) also report that FRI cavity scaling relations may overestimate

FRII jet powers by up to an order of magnitude.

The jet composition (e.g. proton content, electron energy distribution) and collimation/interaction with environment could further impact the observed relation. For example, it has been found that low-luminosity, larger physical size radio AGN jets have a higher proton content, consistent with entrainment arguments (Croston & Hardcastle 2014; Hardcastle 2018; Hardcastle et al. 2019; Croston et al. 2018, Croston et al., in prep.). Although this effect is taken into account in the pressure calculation within the cavity-based method, there may be a systematic underestimation for low-power sources with respect to Eq. 2.9 (i.e. the assumption of a universal proton content for all types of radio AGN is invalid).

These arguments could equally explain the upturn in the incidence of high jet power sources (diamond points) shown in Fig. 3.10. This is because the systematically underestimated jet power for lower luminosity radio-detected objects, once corrected for, would shift these sources to higher jet powers, thus populating a power law distribution at the level of the current diamond points.

The mass dependence of the incidence could also inherently be due to an unaccounted for correlation of jet power with stellar mass itself. However, Turner & Shabala (2015) find no such correlation using their modelled FRI and FRII AGN ($z < 0.1$) in realistic galaxy environments. We note that if future black hole mass estimates become available for such large samples, it could be educational to look for direct correlations of the radio incidence with black hole mass, instead of using stellar mass as their proxy.

Lastly, spin (a) and magnetic flux density/configuration are also important parameters in the context of driving jets in AGN (Blandford & Znajek 1977; Sikora & Begelman 2013). Recall that the Blandford-Znajek mechanism (Blandford & Znajek 1977) predicts the spin to directly impact jet power as $Q \propto a^2$ (Meier 2002; Amarantidis et al. 2019). However, this is only one specific model and there are too few objects (with high selection bias) for which spin measurements exist (Reynolds 2021) to be able to qualitatively test the effect of spin on jet power, and its dependence on stellar mass or accretion rate.

Observed radio morphology

The striking difference between the purely compact radio AGN incidence as a function of Q/L_{Edd} and the one including the compact and complex sources, opens up the discussion about whether the jet powering mechanisms may vary with observed (LOFAR) radio morphology (see also e.g. Mingo et al. 2022).

Firstly, a given source radio morphology determination may be plagued with orientation effects or be dependent on the observing frequency. For example, the uncertainty in observing pure-core or core and lobe emission in projection, may hinder the ability to draw conclusions on the fundamental nature of jet powering, due to contaminating factors such as environment. Also, the increase in sensitivity towards older electron populations at lower frequencies (Condon 1992), could be problematic for jet power estimates of sources no longer possessing an active core. This is complicated by the fact that jet power is a lifetime average quantity, in contrast to the X-ray emission that can respond relatively fast to accretion events. Therefore, measuring ‘instantaneous’ jet power is almost impossible.

To try to understand why there is such a strong increase in the fraction of galaxies hosting radio AGN at high jet powers, the stellar mass, radio luminosity, physical size and $\lambda_{\rm Jet}$ distributions of radio AGN exhibiting different morphological properties are shown in Fig. 3.11. The stellar mass distributions of all subsets are similar, but a dichotomy is present in the radio luminosity between the compact and the complex single component sources versus the complex multiple component and also large sources. Fig. 2.12 showed hints of the relatively numerous complex radio AGN population at $\log(L_R/[\text{W Hz}^{-1}]) > 25$. This bi-modality is then translated to the $\lambda_{\rm Jet}$ distribution, resulting in the upturn in the incidence being caused by these large, complex (multiple components) radio AGN, more specifically the subset with physical sizes > 60kpc. We note that the lack of upturn in the lowest stellar mass bins in Fig. 3.10, is due to the lack of large complex sources in those mass bins as shown in Fig. 3.11.

However, there are conflicting results in the literature about how morphology affects jet power estimates. For example, Godfrey & Shabala (2013) find good agreement between the jet power relations derived for both FRIs versus FRIIs and low versus high power sources. They find that the supposed higher $L_{\rm R}/Q$ ratios in FRII sources, due to lower fraction of the energy in non-radiating particles, are counteracted by the effects of lower density environments, spectral ageing and strong shocks, bringing FRII $Q - L_{\rm R}$ relations into agreement with that of FRIs. On the other hand, Turner & Shabala (2015) find that larger physical size sources tend have higher jet powers compared to smaller sources with equivalent morphology and luminosity, and FRII-like objects have a factor of two higher jet powers than FRI-like sources (see their Fig. 9).

Radio AGN environment

It is known that radio emission can be boosted by denser (e.g. group- or cluster-like) environments, due to the reduced adiabatic expansion losses and more synchrotron shocks as the multiple component, large jets interact with the surrounding interstellar and galactic medium (Barthel & Arnaud 1996; McNamara et al. 2006; Hardcastle & Krause 2013). Therefore, a straightforward hypothesis for the increased incidence at high $\lambda_{\rm Jet}$ compared to the compact-only case, would be that complex sources are more often located in dense environments. This is because the incidence of complex sources, which become increasingly dominant at increasing $\lambda_{\rm Jet}$, would then be counted in preferentially higher $\lambda_{\rm Jet}$ bins due to the boost in radio luminosity.

To test this, we match the G9 radio AGN to the publicly released GAMA DR2 Geometric Environment catalogue[3] (Eardley et al. 2015), in which galaxies have been rigorously sorted into four Cosmic Web environments: knots, filaments, sheets and voids (in decreasing order of density). This is only valid for $0.04 < z < 0.263$, so all samples and completeness limits have been adjusted accordingly. In total, there are 27 radio AGN in low-density voids and 228 in sheets, filaments and knots.

In addition, the radio AGN are also matched to the eFEDS X-ray cluster catalogue

---

[3] Available at: `http://www.gama-survey.org/dr4/schema/dmu.php?id=96`

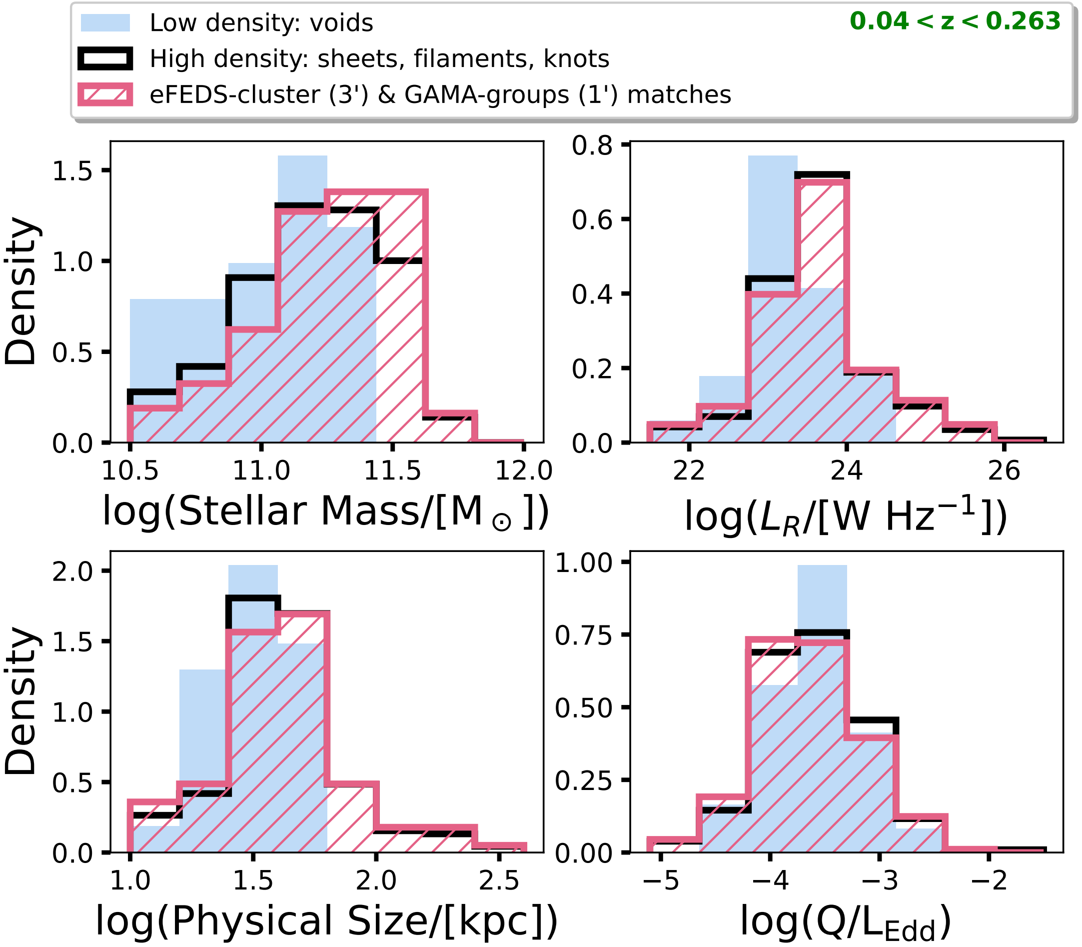


Figure 3.14: Histograms showing the parameter distributions of the G9 radio AGN split into different environments. Radio AGN in dense environments (black unfilled, magenta hatched) are not the cause for the upturn in the measured incidence seen in Fig. 3.10, as they have the same λ_{Jet} distribution as the ones in low density (light blue filled) environments. We note that the GAMA cosmic web environment catalogue is limited to $0.04 < z < 0.263$, so this is applied to all datasets shown here, along with the appropriate mass completeness cuts.

(Liu et al. 2022a) using a 3 arcmin match radius and to the GAMA09 groups catalogue[4] (Robotham et al. 2011) within 1 arcmin, as a further probe of dense environments. Out of the 264 mass-complete radio AGN in $0.04 < z < 0.263$, 171 and 26 match to the GAMA groups and eFEDS cluster catalogue, respectively.

Figure 3.14 shows that, as expected, the radio AGN in higher density environments reside preferentially in larger stellar mass galaxies, and do have increased radio emission, as well as larger physical sizes. However, they are not preferentially hosting radio AGN

---

[4] Available at: `http://www.gama-survey.org/dr4/schema/dmu.php?id=115`

with higher $\lambda_{\rm Jet}$ compared to those in low density voids.

This is interesting as numerous past work finds environment to be an important factor in the driving and propagation of jets (e.g. Hardcastle & Krause 2013; English et al. 2016; Croston et al. 2019; Morris et al. 2022; Mingo et al. 2022). The incidence of radio AGN is also reported to be higher in dense, cluster environments (Best 2004; Best et al. 2005, 2007; Sabater et al. 2013; Best & Heckman 2012; Heckman & Best 2014). Yet, not many studies have probed differences in environment using also mass-scaled parameters and samples well controlled for mass and luminosity completeness.

We therefore conclude that environment is *not* the primary cause of the observed increased incidence of powerful, complex, large-scale radio AGN at high stellar masses.

Radio size and lifetimes

An alternative explanation for the boosted incidence of high jet power sources in the complex radio morphology sample would be the idea put forward by Hardcastle (2018); Hardcastle et al. (2019). They state that large, powerful radio sources are just the older, longer-lived population of normal radio AGN, rather than a marker of different host galaxy or larger scale environmental properties. Hardcastle et al. (2019) show that simulated samples of small, low power radio AGN (possibly akin to the ‘compact’ sample here) are better modelled with log-uniformly distributed lifetimes (more short lived than long lived), whereas large, high power AGN (similar to the ‘complex’ sample), fit better to uniformly distributed lifetimes.

A consequence of the longer-lived nature of the complex radio AGN could manifest through the boosted detection fractions of this population, compared to the compact sources, potentially explaining the elevated incidence measured at high $\lambda_{\rm Jet}$ (Fig. 3.10). Recall here that there is a $\sim 10-30\%$ FRII fraction (see Table 3.2), objects known to have large physical sizes and powerful jets, among the high $\lambda_{\rm Jet}$ sources at all masses. This would also explain the similar stellar mass distributions between the different subsets in Fig. 3.11, but clear dichotomy in the radio luminosity and therefore $\lambda_{\rm Jet}$ distribution. Furthermore, this could be the reason why the mass dependence of the incidence increases with respect to the compact-only case: at higher jet powers for all masses, there will be more larger, longer lived radio AGN; in fact this is shown by the diverging incidences between the single (shaded region) and multi-component (markers) sources on Fig. 3.10. Just considering the single component sources, the incidence as a function of stellar mass is approximately similar to the compact only case.

3.5.3 Towards probing disk–jet coupling through radio and X-ray incidences

The incidence analysis presented here is a powerful probe of the bulk behaviour of a population, in this case the disk–jet coupling in radio and X-ray AGN. Above we show that there is a non-zero mass dependence of the incidence of compact radio AGN as a function of $\lambda_{\rm Jet}$ and that an additional jet power dependence becomes apparent when we include complex radio AGN. However, the origin of these differences and whether they can be related to intrinsic differences in accretion modes is unclear. Drawing a firm connection between empirical incidences and accretion models may be hampered by the fact that the radio and X-ray AGN in our sample populate a mixture of different accretion modes (see Fig. 2.18 and associated discussion). Therefore, there may be no dominant population driving clear trends in the total incidence of radio AGN.

Past studies of (bright) radio AGN have made attempts to separate the radiatively efficient and inefficient accreting sources, by classifying them into 'high' and 'low' excitation radio galaxies (HERGs/LERGs), depending on whether strong optical emission lines are present or not (Best & Heckman 2012). Interestingly, Smolčić et al. (2009) find that LERGs preferentially live in higher mass galaxies compared to HERGs, in line with their expected lower accretion rates, but both populations can span similar ranges in radio luminosity, although HERGs tend to be skewed towards the more luminous regime (Best & Heckman 2012). These properties put HERGs exactly in the region of mass and intrinsic jet power where we observed a marked increase in the incidence for 'complex' radio sources (Fig. 3.10). Indeed, Fig. 2.17 shows that the majority (96%) of sources above $\log(\lambda_{\rm Jet}) \gtrsim -3.0$, the value above which there is a boosted incidence, are complex. This means that the majority of the sources lying above the $\lambda_{\rm Jet} \propto \lambda_{\rm Edd}$ line in Fig. 2.18 (i.e. 'jet dominated' AGN) are also complex in their observed (LOFAR) morphology. Considering that the incidences of compact versus complex radio AGN showed such striking differences in their mass and jet power dependencies, and that it is the complex morphologies which dominate at these high jet (and radiative) powers, it could be postulated that morphology (and/or radio physical sizes, as shown on Fig. 3.11), could be a signpost of jet domination, and a tracer of a specific accretion mode in this sample. This may be important in qualitatively understanding the observed radio AGN incidence results, which we elaborate upon below.

Let us assume the intrinsic distribution of accretion rate, $P_{\dot{m}} = d\,P(\dot{m}, M)/d\log\dot{m}$, as a function of dimensionless accretion rate ($\dot{m} = \eta\dot{M}c^2/L_{\rm Edd}$, with η the accretion efficiency) and black hole mass ($M = M_{\rm BH}/M_{\odot}$), follows a power law distribution described by $P_{\dot{m}} \propto \dot{m}^{-\alpha}\, M^{\beta}$. Observationally (Fig. 3.2), we find that the incidence follows a power law $P_{\lambda_{\rm Edd}} = {\rm d}\,P(\lambda_{\rm Edd}, M)/{\rm d}\log\lambda_{\rm Edd} \propto \lambda_{\rm Edd}^{-\gamma}{\rm M}^{\delta}$, where $\gamma \sim 0.65$ and $\delta \sim 0$. Using a scaling of $\lambda_{\rm Edd}$ with $\dot{m}$ in different accretion modes, it is possible to determine the value of α (assuming $\delta = 0$). Adopting a two-modes relation (Merloni & Heinz 2008):

$$\lambda_{\rm Edd} \propto \begin{cases} \dot{m}, & \dot{m} > \dot{m}_{\rm crit}. \\ \dot{m}^2, & \dot{m} < \dot{m}_{\rm crit}, \end{cases} \tag{3.1}$$

where $\dot{m}_{\rm crit}$ is the critical dimensionless accretion rate at which a putative state transition takes place, one finds $\alpha = \gamma$ for $\dot{m} > \dot{m}_{\rm crit}$ and $\alpha = 2\gamma$ for $\dot{m} < \dot{m}_{\rm crit}$.

Therefore, in this scenario the incidence of X-ray AGN as a function of $\lambda_{\rm Edd}$ is expected to flatten at low accretion rates below the critical value. This flattening, and possible turnover, has recently been found observationally by Aird et al. (2017); Torbaniuk et al. (2024), who probe the X-ray AGN incidence at $\log(\lambda_{\rm Edd}) \ll -3$ using larger samples and deeper X-ray data, as well as by Volonteri et al. (2016) using cosmological hydrodynamical simulations. On the other hand, for high accretion rates, the incidence of X-ray AGN has been shown to drop off dramatically as the sources reached their Eddington limits (Bongiorno et al. 2012; Georgakakis et al. 2017).

The same parameterisation can be made for the observed radio AGN incidence, $P_{\lambda_{\rm Jet}} = d\,P(\lambda_{\rm Jet}, M)/d\log\lambda_{\rm Jet} \propto \lambda_{\rm Jet}^{-x}\ M^{y}$, where $x \sim 1.5$ (Fig. 3.8) and y is poorly constrained, but likely > 0. Next, we make a first assumption that the same intrinsic mass accretion distribution is responsible for the energy provided to power the jet, meaning that the jet power at the base is governed by the internal energy of the system (Heinz & Sunyaev 2003) and so $P_{\lambda_{\rm Jet}} \propto P_{\dot{m}}$. Although the mass dependence is found to be small but non-zero, we start by examining the simplest $y = 0$ case. Given a generalised formulation of Eq. 3.1 of $\lambda_{\rm Edd} \propto \dot{m}^{A}$, we can write $\lambda_{\rm Jet}^{-x} \propto \dot{m}^{-A*\alpha}$. Then, $\lambda_{\rm Jet}$ scales linearly with $\dot{m}$ in the LK mode, where $A = 2$ (Heinz & Sunyaev 2003; Merloni & Heinz 2007). Therefore, we obtain $x = 2\alpha = 1.3$, which is consistent, within errors, with the steep compact radio AGN incidence slope of ~ -1.5, and with the prediction above that compact sources trace the relatively radiatively inefficient branch.

For the complex morphologies, we cannot make concrete statements as the observed incidence does not follow a power law distribution and the assumption of $y = 0$ is not appropriate. A larger sample of high-power sources, possibly exploring larger volumes than what is accessible in the GAMA09 field will be needed to properly characterise the functional form of their incidence, and we defer this analysis to future work.

3.6 Summary

In this Chapter, the incidence of radio and X-ray AGN as functions of several mass-normalised power indicators are presented, with our aim being to test whether or not the mass-invariant AGN triggering and fuelling mechanisms seen in the X-ray selected AGN samples also translate to a mass-invariant jet powering mechanism for radio AGN.

Firstly, even with the soft response of eROSITA, we are able to recover the mass-invariance of the X-ray AGN incidence as a function of $\lambda_{\rm Edd}$ found with X-ray AGN samples selected from *Chandra* and *XMM-Newton* (Fig. 3.2). This is possible after carefully controlling for incompleteness due to the loss of obscured AGN, and is enabled by the good spectroscopic capabilities of the eROSITA CCDs (Meidinger et al. 2020).

The novelty of this work is the consideration of the incidence of radio AGN as a function of specific black hole kinetic power, $\lambda_{\rm Jet}$. We use the new fully characterised sample of radio AGN, as defined in Chapter 2, with host galaxies identfied in the GAMA09 spectroscopic galaxy survey. These GAMA09 galaxies provide a well-controlled parent sample of known completeness as a function of stellar mass, within which we then compute the incidence of radio AGN. For the GAMA09 galaxies, we also compile a complete sample of 34 sources hosting radio AGN with secure FRII morphologies, three of which are classified as giant radio galaxies.

The main results of this Chapter are summarised below:

- The strongly increasing incidence of radio AGN detected above the LOFAR flux limit as a function of stellar mass is recovered (Fig. 3.5), as in past studies (e.g. Sabater et al. 2019), and hints of increasing normalisation with increasing redshift are present.

- The fraction of quiescent versus star-forming GAMA09 galaxies hosting compact radio AGN are similar (Fig. 3.9), suggesting that radio AGN are *not* only found in 'red and dead' galaxies.
- The fraction of GAMA09 galaxies hosting compact radio AGN as a function of $\lambda_{\rm Jet}$ shows approximately constant power law slope of -1.5, but increasing normalisation with increasing stellar mass, and redshift (Figs. 3.7 and 3.8). The strong observed increase of the incidence as a function of stellar mass (see first point), is a *selection bias* favouring radio AGN with lower specific radio luminosity to be detected in higher mass galaxies as a result of flux limited surveys.
- The constant slope of the incidence for different stellar masses highlights that it is not only the most massive radio AGN that host the most powerful radio AGN and the low-mass radio AGN that host the low-power radio AGN. This slope is also steeper than the power law describing the X-ray AGN incidence as a function of $\lambda_{\rm Edd}$, ~ -0.65, which may be understood in the context of reprocessing of accretion energy into coronal X-ray and radio (jetted) emission and suggests that compact radio AGN trace a relatively radiatively inefficient mode of accretion. Incidence analysis is thus useful to gain insight into the average properties of a given accretion mode.
- Including complex radio morphology sources reveals a striking boosted incidence at high $\lambda_{\rm Jet}$, due to large physical size (> 60 kpc) and multiple component radio AGN, in the (high) stellar mass bins where such sources are dominant (Fig. 3.10). We find that this enhanced incidence is *not* due to environmental effects (i.e. to powerful complex radio AGN residing in more massive haloes), as probed by the GAMA sample. A different mass dependence of the incidence in comparison to the pure compact case is also clear at all $\lambda_{\rm Edd}$.

We discuss the numerous caveats associated with calculating jet power for different radio luminosities, sizes, and morphologies that complicate the interpretation of the physical mechanisms driving jets. Finally, we explore the disk–jet connection in various accretion modes by postulating different accretion rate and black hole mass scaling of the radio and X-ray AGN incidence.

Overall, larger sample sizes probing larger volumes and depths, along with improved jet power estimation, taking into account the numerous influential factors, are the key to better understand the disk–jet connection through AGN incidences. Future machine-learning-based algorithms will also help in finding and characterising X-ray and radio AGN (Alegre et al. 2022; Mostert et al. 2022; Barkus et al. 2022), such that incidence analysis can be completed for the large current/upcoming wide and deep surveys, for example WEAVE/LOFAR (Smith et al. 2016), eRASS (Merloni et al. 2024), ASKAP (McConnell et al. 2020) and VLASS (Lacy et al. 2020).

Chapter 4

Global Energetics of radio AGN kinetic feedback in the local Universe

In this Chapter, we further explore the difference in the compact versus complex radio AGN incidence distributions. We use the detailed knowledge we have now gained on this sample of 682 radio AGN, including the duty cycle of their activity (via the incidence), the stellar mass of their hosts (via the GAMA09 survey) and their jet powers (via the LOFAR information), to make statements about the global energetics of radio AGN kinetic feedback in the local universe ($z < 0.4$). We derive the average jet power of massive galaxies and its variation as a function of stellar mass (M_*), halo mass (M_h) and radio morphology. We compare the incidence distributions of compact and complex radio AGN as a function of specific black hole kinetic power, $\lambda_{\rm Jet}$, and synthesise, for the first time, the radio luminosity function (RLF) by M_* and radio morphology. Our RLF and derived total radio AGN kinetic luminosity density, $\log \Omega_{\rm kin}/[{\rm W\ Mpc^{-3}}] = 32.15^{+0.18}_{-0.34}$, align with previous work.

Our key findings are the following. Kinetic feedback from radio AGN dominates over any plausible inventory of radiatively driven feedback for galaxies with $\log M_*/M_\odot > 10.6$. More specifically, it is the compact radio AGN which dominate this global kinetic energy budget for all but the most massive galaxies ($10.6 < \log M_*/M_\odot < 11.5$). Subsequently, we compare the average injected jet energy ($\overline{E_{\rm Jet}}$) against the galaxy and halo binding energy ($U_{\rm bin}$), and against the total thermal energy of the host gas ($E_{\rm th}$) within halos. We find that both compact and complex radio AGN lack the energy to fully unbind their host galaxies (when considering the appropriate scales for such kinetic feedback to take place) and host dark matter halos; but they can contribute significantly to locally disrupting the kinematics of the gas distribution within their hosts. On the other hand, $\overline{E_{\rm Jet}}$ greatly surpasses the global $E_{\rm th}$ for groups, thereby providing a crucial input to the gas and thermodynamical balance in these systems. Finally, we show that AGN jets can also significantly impact the *local* thermodynamical balance in the cores of large groups and massive clusters. Overall, our findings provide important insights on jet powering, accretion processes and black hole-galaxy coevolution via AGN feedback, as well as a clear observational benchmark to calibrate AGN feedback simulations.

The contents of this Chapter are published in Igo & Merloni (2025).

4.1 Gaining empirical insight on AGN feedback

As motivated in Chapter 1, the importance of AGN feedback in galaxies, groups and clusters has been empirically elucidated over the past two decades (e.g. Fabian 2012; King & Pounds 2015; Eckert et al. 2021; Hlavacek-Larrondo et al. 2022; Bahar et al. 2024), and this phenomenon is now a required ingredient in our theoretical understanding of galaxy formation, as demonstrated by several cosmological simulations (e.g. Croton et al. 2006; Bower et al. 2006; Hopkins et al. 2006; Sijacki et al. 2007; Schaye et al. 2015; Croton et al. 2016). However, we still lack a quantitative understanding of the different modes of feedback, of the dependencies on AGN and environmental parameters over different scales and cosmic times, and how these should guide fine-tuned inputs to simulations (e.g. see recent reviews by Hardcastle & Croston 2020b; Harrison & Ramos Almeida 2024).

Both the radiative and kinetic feedback modes can have positive and/or negative impact on the black hole growth itself, on the structure of the host galaxy or even on the larger-scale host halo. For example, suppression or enhancement of star formation, depletion or replenishing of gas reservoirs, halting or fuelling the growth of the supermassive black hole (SMBH) itself (e.g. Ferrarese & Merritt 2000; Bower et al. 2012; Kormendy & Ho 2013; Harrison 2017) are all secular phenomena in the evolution of galaxies linked to AGN feedback phenomenology, and are known to act on varying timescales. These aspects make it challenging to accurately disentangle causality from effect, to quantify the energetics of a given mode of feedback and to understand on what scales each of these modes has the greatest impact, not to mention how this varies with cosmic time.

Nevertheless, one aspect that is now observationally clear, and relevant in the context of this Chapter, is that the kinetic power attributed to particle jets can efficiently offset cooling flows in galaxy groups and clusters (e.g. Fabian 1994; Peterson et al. 2004; McNamara et al. 2005; McNamara & Nulsen 2012), thereby preventing extreme star formation in the massive central galaxies of those halos (note that gas does not necessarily have to be expelled from the host galaxy or halo to be considered 'negative feedback'). In fact, bubbles inflated by jets have been found to be coincident with X-ray cavities in clusters, showing direct evidence of AGN kinetic feedback and allowing estimates of the jet power to be computed (e.g. Boehringer et al. 1993; Carilli et al. 1994; Bîrzan et al. 2004; Dunn & Fabian 2006; Cavagnolo et al. 2010; Timmerman et al. 2022). However, the efficiency with which jets of varying power, orientation and morphology couple to the gas in the surrounding inter- or circumgalactic medium (ISM, CGM) is unclear (e.g. Wagner et al. 2012; Mukherjee et al. 2018; Perucho 2019; Meenakshi et al. 2022; Bourne & Yang 2023).

One approach to gain empirical insight on AGN feedback is to use large, well-characterised samples to trace the global properties of a given AGN population, such as their incidence and power distributions. The study of incidences as a function of mass-scaled kinetic power, due to its statistical completeness, provides important information about how AGN with different jet powers and radio morphologies, in galaxies of different masses, release energy into their host galaxies (or halos). For example, the incidence of radio AGN as a function of specific black hole kinetic power, λ_{Jet}, can inform us about the average jet power of specific subsets of the galaxy population. The same measure of incidence, when convolved with the galaxy stellar mass function (SMF; Driver et al. 2022; Bernardi et al. 2018), should recover the radio AGN luminosity function. For example, Aird et al. (2013) use the X-ray AGN incidence previously found in Aird et al. (2012) to compute an X-ray luminosity function (XLF) in agreement with observations (Aird et al. 2010).

In this Chapter, building on the results from Chapters 2 and 3, we use the same complete sample of low-redshift ($z < 0.4$) radio AGN to provide a quantitative measure of the average jet kinetic feedback as a function of galaxy stellar mass (M_*), distinguishing between compact and complex radio morphologies. Additionally, we aim to quantify the ability of jet kinetic energy to drive out gas from the host galaxy or halo, as well as its contribution to the heating and cooling balance in the population of massive galaxies (above our completeness limit for the parent sample of host galaxies, $\log(M_*/M_\odot) > 10.6$). This will allow us to estimate, in all generality, the impact of kinetic feedback from radio AGN on different subsets of the galaxy population, ranging from the galaxy ($12 < \log(M_h/M_\odot) < 13$), to the group ($13 < \log(M_h/M_\odot) < 14$) and cluster ($14 < \log(M_h/M_\odot) < 15$) regimes. These estimates are vital for furthering our understanding of black hole accretion processes, including jet powering, as well as black hole-galaxy coevolution and growth.

4.2 Radio AGN incidence

In Chapter 3, we presented for the first time an accurate measure of the radio AGN incidence, that is, the fraction of (massive) galaxies hosting radio AGN as a function of the specific black hole kinetic power for compact (see Fig. 3.7) and both compact and complex radio morphologies (see Fig. 3.10). However, since one of the aims of this Chapter is to quantitatively understand how radio morphology impacts the power, efficacy and scale of the kinetic feedback, we re-compute the incidence individually for compact and complex radio AGN. It is important to highlight here that the split in radio morphology is dependent on the resolution of the survey, in this case the $8'' \times 9''$ resolution of the LOFAR-eFEDS field, so small-scale sources may become resolved at higher resolution. Interestingly, LOFAR Very Long Baseline (LOFAR VLBI) studies find that small-scale sources may become resolved complex sources at higher resolution ($0.3''$), although only 40% of unresolved sources detected at the Dutch station resolution of $6''$ have a high resolution counterpart, and of this sample 89% remain compact (Morabito et al. 2022; Sweijen et al. 2022). The radio AGN incidence is computed in bins of $\Delta \log \lambda_{\rm Jet} = 0.3$, where the incidence is defined as the probability density per logarithmic $\lambda_{\rm Jet}$ interval (units of $[\log \lambda_{\rm Jet}]^{-1}$).

Figure 4.1 (top) shows the compact radio AGN incidence, fit by a simple power law of the form

$$f(x) = A \left(\frac{x}{x_0}\right)^B , \tag{4.1}$$

where A is the normalisation, $\log(x_0) = -3.25$, and B is the slope. Bayesian fitting, using

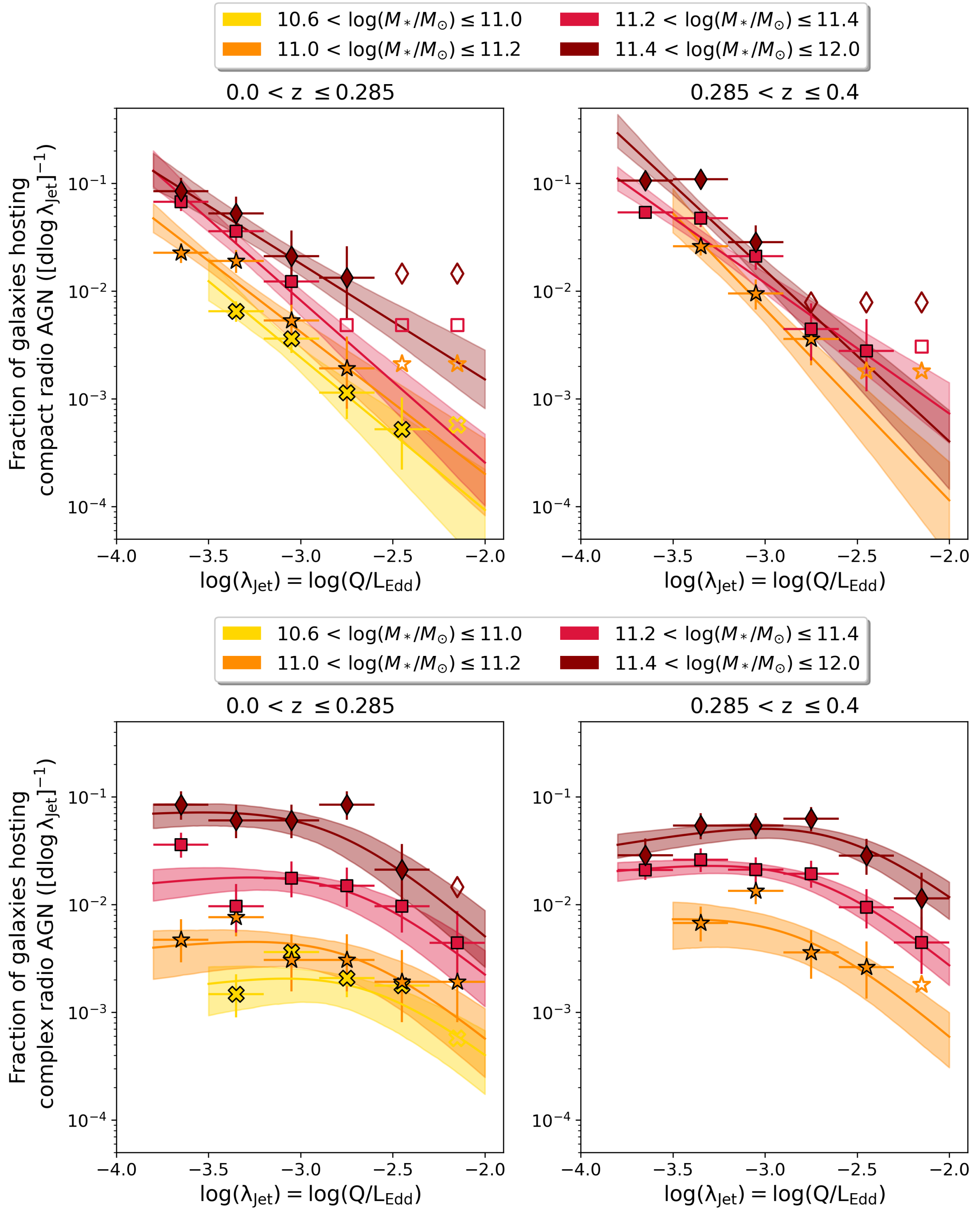


Figure 4.1: Fraction of GAMA09 galaxies hosting compact (top two panels) and complex (bottom two panels) radio AGN as a function of λ_{Jet}, in different stellar mass and redshift bins, modelled by a single and double power law, respectively (solid line with 1σ error margin). The 3σ upper limits are shown with unfilled markers.

UltraNest[1] (Buchner 2021c) is performed by sampling an asymmetric point cloud defined by the 1σ uncertainties on the values (Cameron 2011, upper limits are sampled with a bounded box at the 3σ value). Results of the power law fitting are shown in Fig. 4.2, where a clear increase in normalisation, i.e. a mass-dependence, is seen in both low (green) and high (purple) redshift bins (left panel), whilst the slope remains relatively constant around -1.4 ± 0.2 (right panel), as already discussed in Chapter 3. The average M_* within each bin is used in the fitting instead of the geometric bin centre to reflect the true distribution.

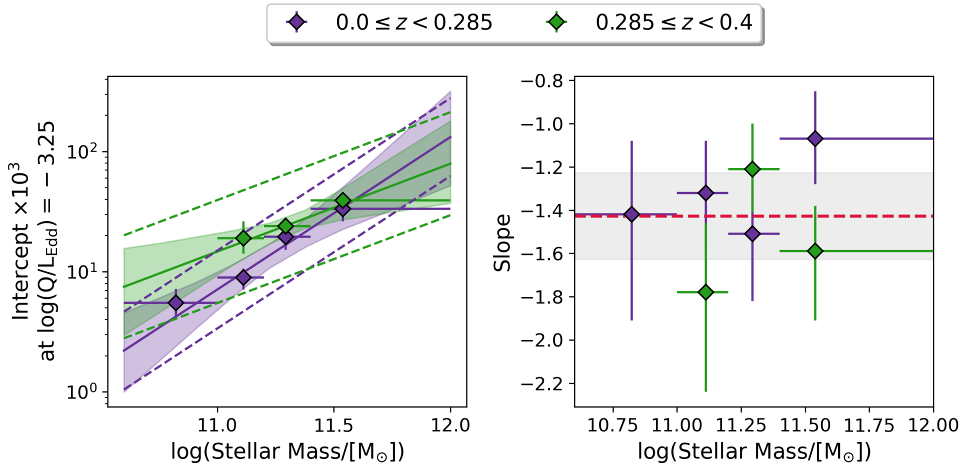


Figure 4.2: Left: Increase in the normalisation of the compact radio AGN incidence found as a function of stellar mass for the low (purple) and high (green) redshift bins. Solid lines with 1σ shaded error regions mark the best-fit and dashed lines mark the intrinsic scatter. Right: Approximately constant steep slope of -1.4 (red dashed line) with standard error of 0.2 (grey shaded region) found as a function of stellar mass and redshift.

Figure 4.1 (bottom) instead shows the complex radio AGN incidence, fit by a double power law described by Eq. 4.2,

$$f(x) = \frac{k}{\left(\frac{x}{x_b}\right)^{-\alpha_1} + \left(\frac{x}{x_b}\right)^{-\alpha_2}}, \tag{4.2}$$

where k is the normalisation, x_b is the break and α_1, α_2 are the faint-end and bright-end slopes, respectively. The power law indices are well-fit on average by a shallow faint-end, $\alpha_1 \sim 0.2$, and a steep bright-end, $\alpha_2 \sim -1.4$, slope. Similarly to the compact-only fits, other than the normalisation of the incidence (k), the other fit parameters remain constant within the standard error deviation, as shown in Fig. 4.3. We note that the double power law fit provides an adequate parameterisation of the incidence distribution for the purpose of this work, but that future larger samples could highlight more complexity in the shape of the distribution, especially when looking into the redshift-evolution in more detail.

---

[1] `https://johannesbuchner.github.io/UltraNest/example-line.html`

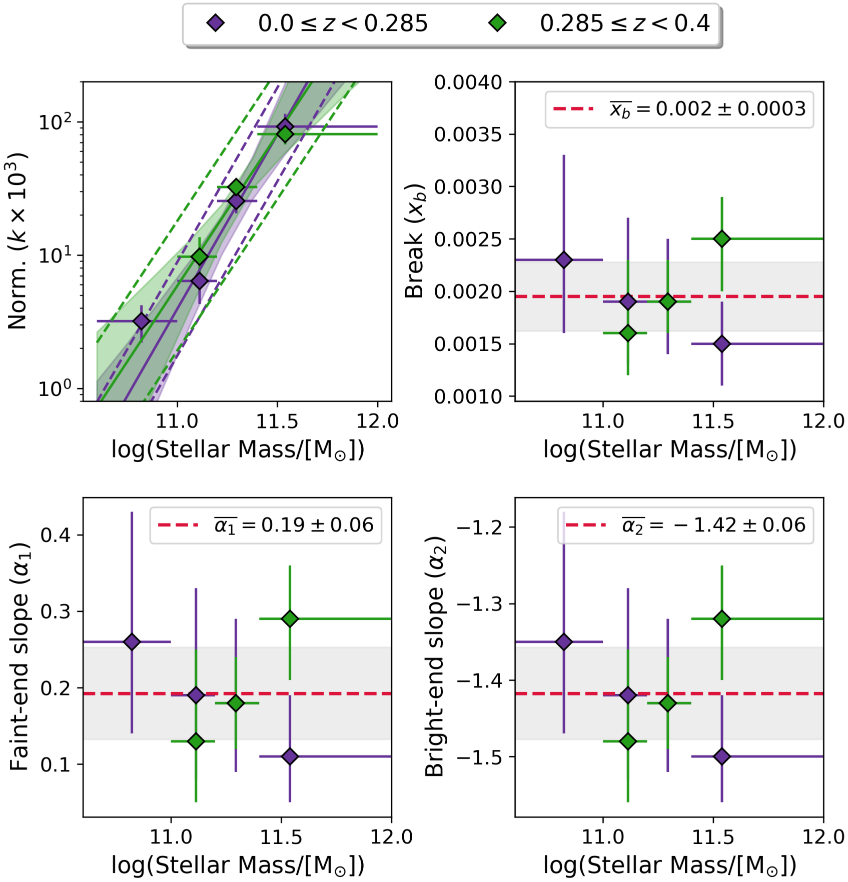


Figure 4.3: Top left panel: increase in the normalisation of the complex radio AGN incidence found as a function of stellar mass for the low (purple) and high (green) redshift bins. Solid lines with 1σ shaded regions mark the best-fit and dashed lines mark the intrinsic scatter. Top right, and bottom panels: approximately constant break, faint- and bright-end slopes (red dashed lines with grey shaded standard error) are found as a function of stellar mass and redshift.

Fig. 4.1 highlights explicitly the stark differences present between the $\lambda_{\rm Jet}$ distributions of compact versus complex sources, already found in Chapter 3. Although a different functional form (power law plus Gaussian) was used to fit the incidence curves in Fig. 3.10, the results are fully consisted with those presented in Fig. 4.1, now represented in units of $[\log \lambda_{\rm Jet}]^{-1}$, and are independent of the choice of binning. Fig. 4.4 shows this in more detail by plotting the incidence of both compact and complex radio AGN, as well as showing the sum of the power law (Eq. 4.1; Figure 4.2) and double power law (Eq. 4.2; Figure 4.3) fits to the total compact and complex incidence of radio AGN (solid curves with shading) as a function of $\lambda_{\rm Jet}$ from this Chapter. The dot-dashed and dashed curves here show the compact and complex contributions, respectively, to the total incidence curves; these are

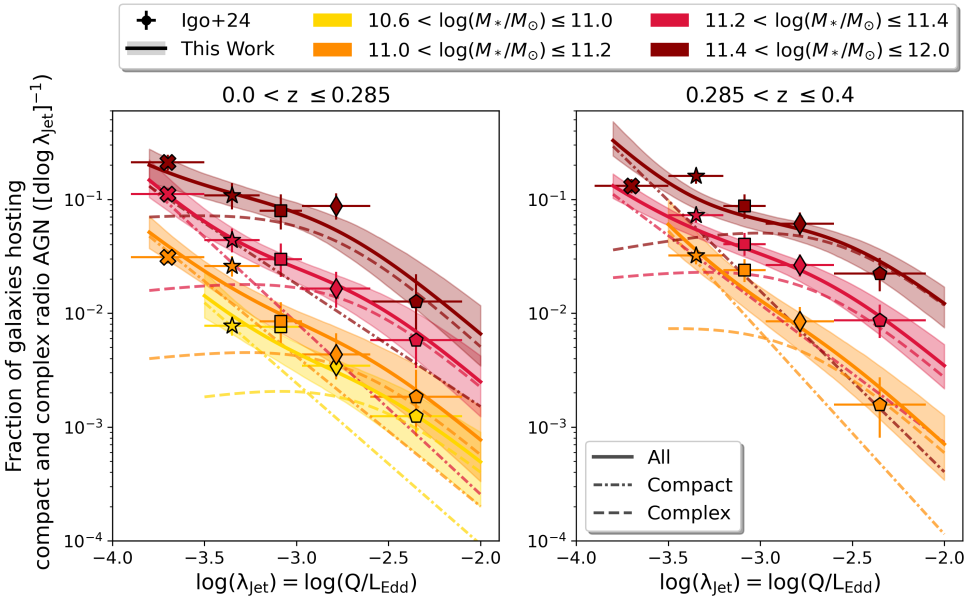


Figure 4.4: Comparison of the incidence of radio AGN as a function of $\lambda_{\rm Jet}$ from Chapter 3 (markers; Igo et al. 2024) and this Chapter (best-fit model and error margin), in units of $[\log \lambda_{\rm Jet}]^{-1}$, showing consistency.

shown in detail in Fig. 4.1.

Overall, the steep power law-like incidence of compact radio AGN (Fig. 4.1, top) indicates that compact (i.e. unresolved by LOFAR) radio AGN dominate the lower jet powers and drop out rapidly at higher jet powers. In contrast, the flatter incidence distribution of complex radio AGN, and their stronger mass scaling (Fig. 4.1, bottom) indicates that complex radio AGN progressively dominate at high M_* and $\lambda_{\rm Jet}$. Importantly, the strong dependence of the incidence distributions on radio morphology cannot simply be explained by caveats in jet power determination discussed in Ch. 3.5, thereby pointing to a true physical origin.

Interestingly, the break in the jet power distribution occurs around the point where the fraction of FRII double-lobed radio AGN out of the total compact and complex radio AGN sample increases rapidly. The average secure FRII fraction, defined as those classified unanimously by visual inspection to have a double-lobed, edge-brightened structure (see Chapter 2), between $-3.8 < \log \lambda_{\rm Jet} \leq -2.9$ is 4.6%, whereas between $-2.9 < \log \lambda_{\rm Jet} < -2.0$ it reaches 37.3% (averaged across all mass and redshift bins). A larger sample would be needed to understand the true nature of this break, as it could be an indication of a change of accretion mode (see Ch. 3.5.3), inadequacy of a universal $Q - L_R$ relation (see Ch. 3.5.2) or something else entirely.

In the following sections we will elaborate on the implications of such stark differences in the radio AGN incidence and jet power distributions for the synthesis of the radio AGN luminosity function, and for our estimates of AGN kinetic feedback for galaxies (and halos) of different masses.

4.3 Synthesis of the radio luminosity function

Under the assumption that our samples are complete (or completeness-corrected), convolving the incidence distributions — the functions describing how many galaxies host radio AGN in a given luminosity range — with the galaxy SMF — a measure of how many galaxies there are in a given stellar mass range per comoving volume — is an alternative way to recover the AGN radio luminosity function (RLF). The RLF describes how many radio AGN there are within a given luminosity range per comoving volume.

This is analogous to previous works, for example Aird et al. (2013), where it was shown to be possible to recover the XLF by convolving the X-ray AGN incidence, as computed in Aird et al. (2012), with the SMF from Moustakas et al. (2013). Based on our work in the eFEDS/GAMA09 field, we could also reproduce the results of Aird et al. (2013): the eFEDS X-ray AGN incidence results presented in Chapter 3 are indeed in good agreement with Aird et al. (2012); therefore, we do not discuss XLFs further in this Chapter. Instead, we focus here on computing the RLF and its dependence on radio morphology and host stellar mass.

Following the method of Aird et al. (2013), the RLF of AGN can be expressed as:

$$\begin{aligned}\psi(L_{\mathrm{R}}, z) &= \phi(M_*, z) * p(\lambda_{\mathrm{Jet}} \mid M_*, z) \\ &= \int \phi(M_*, z)\, p\Big(\lambda_{\mathrm{Jet}} \mid M_*, z\Big) \frac{\mathrm{d}\log(\lambda_{\mathrm{Jet}})}{\mathrm{d}\log(L_R)}\, \mathrm{d}\log M_* \end{aligned} \tag{4.3}$$

where $\phi(M_*, z)$ is the galaxy SMF and $p(\lambda_{\mathrm{Jet}} \mid M_*, z)$ describes the probability in $[\log \lambda_{\mathrm{Jet}}]^{-1}$ units, for a galaxy of given M_* and z to host an AGN with a specific kinetic power, λ_{Jet}.

Given the use of GAMA09 data in this work, the appropriate SMF would be the one using GAMA galaxies from Driver et al. (2022). However, as seen in their Figure 12 and as pointed out by the authors, the GAMA SMF is unable to accurately capture the high-mass end due to the small survey volumes. For our study on radio AGN, which is complete only for massive galaxies ($\log(M_*/M_\odot) > 10.6$), we therefore adopt the SMF from Bernardi et al. (2018) who use the larger volume Sloan Digital Sky Survey (SDSS) and find consistent results with Driver et al. (2022) up to $\log(M_*/M_\odot) \sim 11$.

Figure 4.5 shows our derived RLF (black), in the redshift range $0 < z < 0.285$, broken down, for the first time, into the contribution from compact (light blue) and complex (dark blue) radio AGN and their respective contributions in different host stellar mass bins (same colourbar as Figure 4.1). Note that we only plot the range of massive galaxies with $10.5 < \log(M_*/M_\odot) < 12.0$, meaning that at the lowest luminosities, where lower mass galaxies dominate, the RLF shown in Fig. 4.5 should be regarded as a lower-limit.

The RLF is plotted with a solid line in the range $\log(L_R/[\mathrm{W\ Hz^{-1}}]) = 23.65 - 26$, where the lower bound corresponds to the 95% radio luminosity completeness level at $z = 0.285$ (see Fig. 2.15) and the upper bound is driven by the LOFAR-eFEDS survey's volume-limit. To obtain the RLF in this range, only the radio AGN incidence distributions in the observed range of $-4 < \log \lambda_{\mathrm{Jet}} < -2$ (see Fig. 4.1) are convolved with the SMF as described in Eq. 4.3.

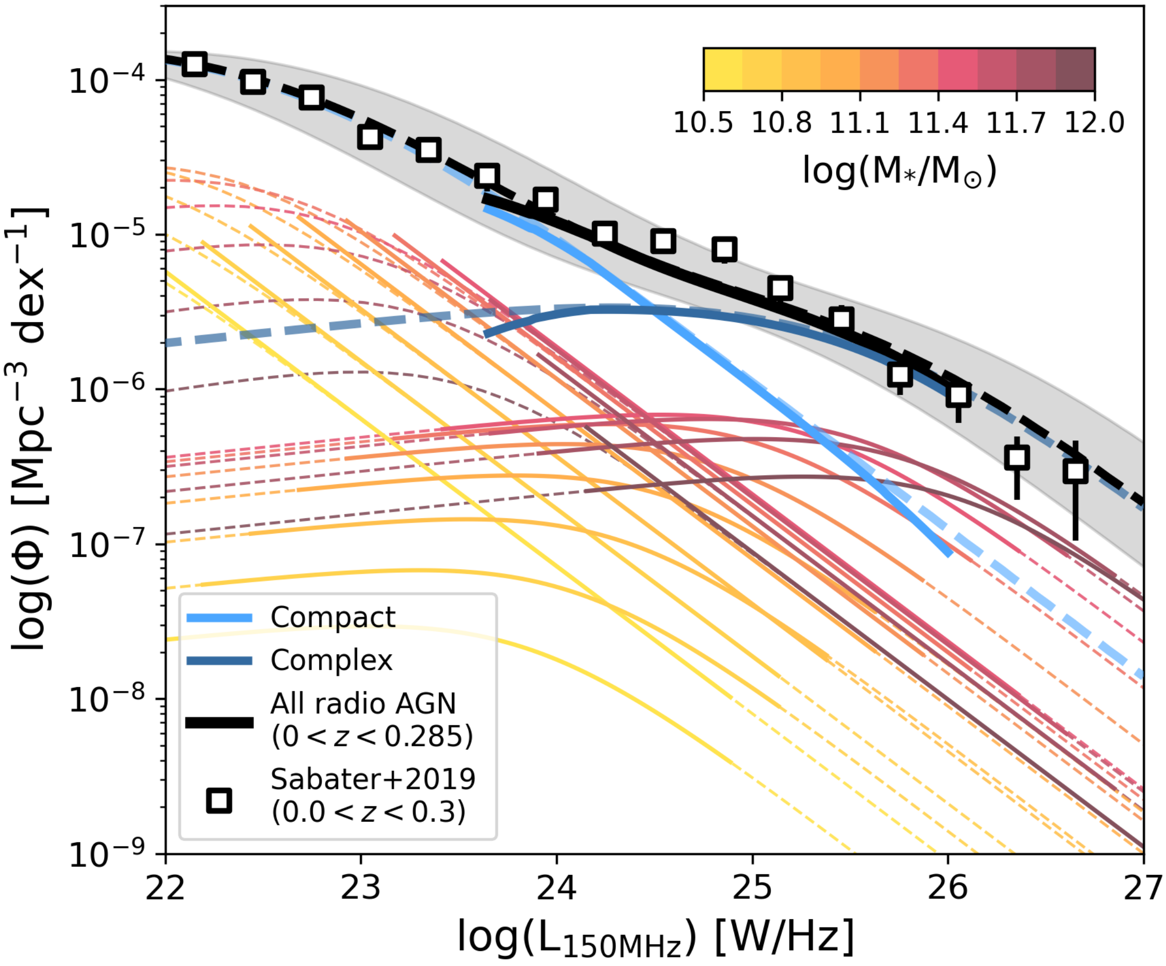


Figure 4.5: The radio luminosity function (RLF; black solid line), derived by convolving the radio AGN incidence with the stellar mass function, decomposed into the contribution from radio AGN of varying radio morphology (compact – light blue; complex – dark blue) and host galaxy stellar mass (same colour map as Fig. 4.1). The RLF is in good agreement with Sabater et al. (2019) (squares; error bars often too small to be seen). Dashed lines indicate extrapolation of the incidence distribution, and thus the RLF, and the grey-shaded region represents the uncertainty on the RLF (see text for more details).

The dashed lines in Fig. 4.5 are extrapolations to the incidence distributions, given the assumptions that follow, in order to sample the faint and bright ends of the RLF. Firstly, for the incidence of complex radio AGN, we simply extrapolate the low- and high-ends of the double power law in the range $-8 < \log \lambda_{\rm Jet} < 0$ and perform the same convolution procedure as described in Eq. 4.3. We do a similar extrapolation for the incidence of compact radio AGN, except for the low-power end we assume a mass-independent turnover at $\log \lambda_{\rm Jet} = -4.4$ after which a flatter power law with slope 0.3 is followed. With the turnover being beyond the parameter space covered in this work, these values are chosen such that the extrapolated RLF agrees best with literature results (see Ch. 4.5.2 for more discussion). As low power compact sources are very numerous (the incidence function is steep), this turnover is important as the incidence is a probability distribution. Therefore, it must be bounded at its extremes and must integrate to one, which are conditions that a simple power law cannot satisfy. We apply the same extrapolation to the data in the $0.285 < z < 0.4$ redshift bin.

The grey shaded region in Fig. 4.5 represents the uncertainty on the RLF. Given the method to obtain the RLF in this work via Eq. 4.3, the major sources of uncertainty reside in the accuracy of: the SMF, the modelling of incidence distributions, the stellar mass and jet power determination. Having already motivated the choice of SMF for this work above, carefully modelled the incidence of compact and complex radio AGN in Ch. 4.2 and justified the relatively low (~ 0.1 dex) uncertainty in the stellar mass (see Fig. 3.13), we deem the major uncertainty in the RLF originates from the determination of jet power itself (see detailed discussion in Chs. 3.5.2 and 4.5.1). We account for this by adding a ± 0.3 dex scatter on the jet power but keeping all aforementioned extrapolation parameters the same. The choice of ± 0.3 dex is motivated by the scatter (over the range of jet powers probed by this work) in the inferred jet power distribution presented in Hardcastle et al. (2019), which takes into account relevant jet physical processes and source environments.

Given the assumptions above, we reproduce well the full range of the RLF, in line with the results of Sabater et al. (2019) shown by square markers on Fig. 4.5 (see also Kondapally et al. 2022, for more comparison with recent literature RLFs). Overall, we find that compact (complex) radio AGN dominate in number density below (above) $\log(L_{150\mathrm{MHz}}/[\mathrm{W\ Hz}^{-1}]) \sim 24.5$, equivalent to a jet power of $\log(Q/[\mathrm{W}]) \sim 36.7$ (see similar results in Baldi 2023). Additionally, the number density of radio AGN with typical luminosities $\log(L_{150\mathrm{MHz}}/[\mathrm{W\ Hz}^{-1}]) \sim 23.5-25$ is dominated by galaxies of mass $\log(M_*/M_\odot) \sim 11.4$, with broad $\log \lambda_{\mathrm{Jet}}$ distribution (~ -3.8 to -2.6). Given that radio AGN preferentially lie in higher mass galaxies, the predominant contribution by moderately massive galaxies to the RLF is *not* just due to the shape of the SMF (which has a break around $\log(M_*/M_\odot) \sim 11.4$), unlike the mass-independent X-ray AGN incidence and the XLF results by Aird et al. (2013). Above $\log(L_{150\mathrm{MHz}}/[\mathrm{W\ Hz}^{-1}]) \sim 25$, radio AGN with host $\log(M_*/M_\odot) > 11.4$ and $\log \lambda_{\mathrm{Jet}}$ reaching up to -1.7, dominate in number density, whereas below $\log(L_{150\mathrm{MHz}}/[\mathrm{W\ Hz}^{-1}]) \sim 23.5$, the main contributors are radio AGN in galaxies with $\log(M_*/M_\odot) < 11.4$ and reaching down to $\log \lambda_{\mathrm{Jet}} \sim -4.1$.

By transforming the observed radio luminosity into a jet power using Eq. 2.9, i.e. computing a jet kinetic luminosity function, $\rho(Q, z)$, and integrating as in Eq. 4.4, we also derive the average jet kinetic heating rate per Mpc^{-3}, Ω_{kin}, in other words the kinetic luminosity density of the whole AGN population:

$$\Omega_{\mathrm{kin}}(z) = \int Q\ \rho(Q, z)\ \mathrm{d}Q\ \mathrm{W\,Mpc}^{-3}. \tag{4.4}$$

Propagating the uncertainties from the grey shaded region on Fig. 4.5, we obtain $\log \Omega_{\mathrm{kin}}/[\mathrm{W\ Mpc}^{-3}] = 32.15^{+0.18}_{-0.34}$ in the redshift range $0 < z < 0.285$, which is in a similar

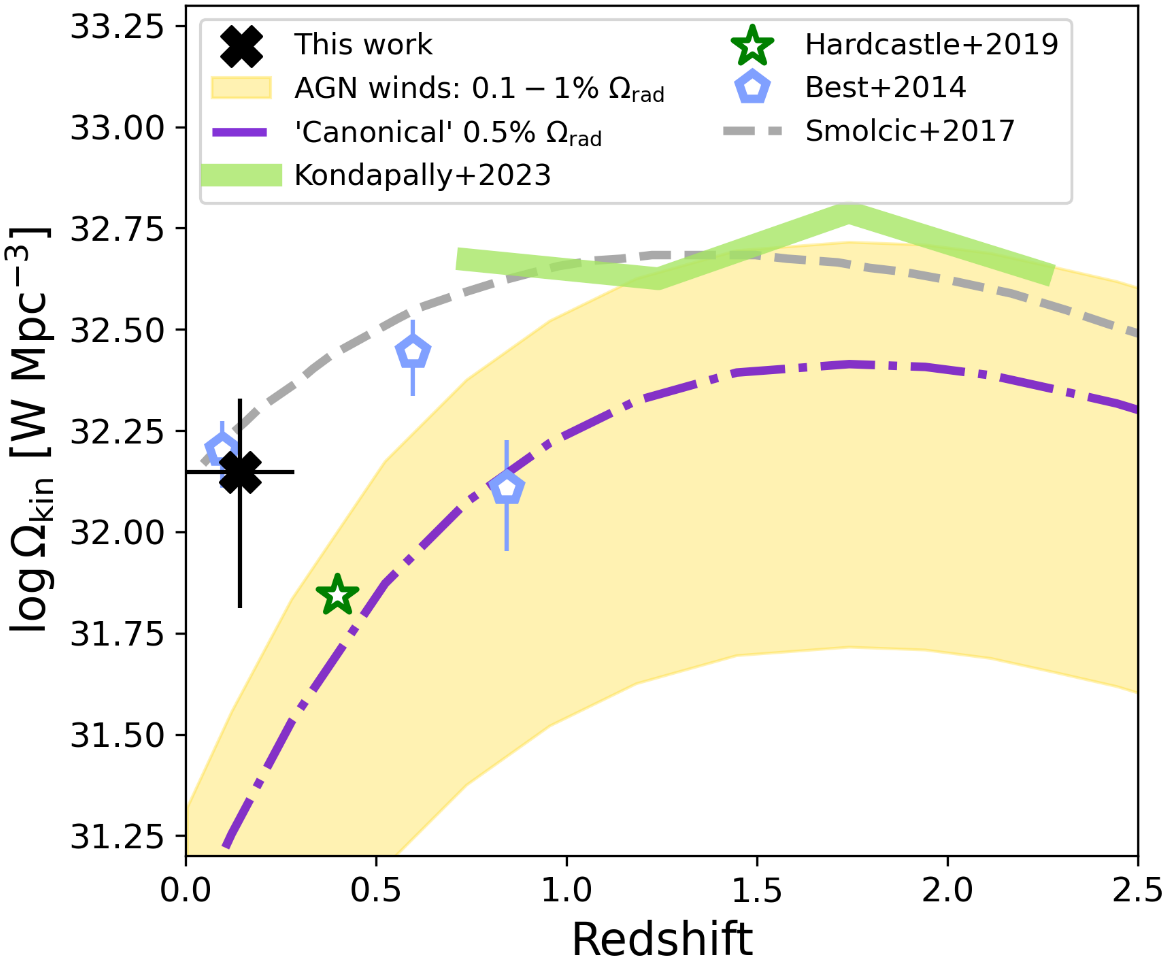


Figure 4.6: Kinetic luminosity density of (radio) AGN as a function of redshift obtained from recent observational work. Results of this work are shown by the thick black cross. The yellow shaded region also shows the $0.1 - 1\%$ of bolometric luminosity density from Aird et al. (2015) which represents a rough indicator of the 'canonical' feedback energy that may be delivered by radiatively driven AGN winds.

range to past observational works as shown in Figure 4.6 (Best et al. 2014; Smolčić et al. 2017; Hardcastle et al. 2019; Kondapally et al. 2023).

Apart from the kinetic luminosity density, we can roughly estimate the total radiative wind-driven outflow energy density by scaling the (well-constrained) AGN bolometric luminosity density as a function of redshift (see e.g. Merloni & Heinz 2008; Aird et al. 2015) by a factor $0.1 - 1\%$. We show this average quasar-driven wind energy density as a yellow shaded region in Fig. 4.6. This range indicates the typical expectation, observationally, for the fraction of AGN radiative output that can efficiently couple to the surrounding medium (henceforth, 'coupling efficiency'), producing feedback effects from multi-phase winds. However, this ‘coupling efficiency’ is widely debated and can span orders of magnitude from $\sim 0.001\%$ to tens of percent of the AGN bolometric luminosity (L_{bol}), as discussed in detail in Harrison et al. (2018); Heckman & Best (2023), with observational and simulation-based estimates tending to lower and higher values, respectively. A large scatter in the ‘coupling efficiency’ is thought to result from numerous factors such as observational limitations in accurately measuring outflow geometries and/or energetics, variability in AGN luminosity and/or outflow power, varied sub-grid physics implementation and resolution of different feedback simulations, uncertainty in the fraction of feedback energy in kinetic form, among many others (see e.g. Costa et al. 2014; Harrison et al. 2018; Weinberger et al. 2018; Ward et al. 2024; Harrison & Ramos Almeida 2024, and references therein for more details). For example, the seminal work of Di Matteo et al. (2005) using hydrodynamical simulations (Springel et al. 2005) of galaxy mergers find that their results can reproduce the $M - \sigma$ relation (Ferrarese & Merritt 2000) given a ‘coupling efficiency’ of 5% L_{bol}. On the other hand, Heckman & Best (2023) review a wide range of recent observational results spanning the molecular, warm-ionised and highly-ionised phases of radiatively driven feedback in a variety of AGN, finding a total ‘coupling efficiency’ of 0.5% L_{bol}. Therefore, throughout this Chapter, we assume 0.5% L_{bol} as the ‘canonical’ fraction of radiative output from multi-phase winds (shown with a purple dot-dashed line in Fig. 4.6) and adopt this reference value when (qualitatively) comparing radiative versus kinetic energetics results.

The global estimates presented in Fig. 4.6 show that kinetic feedback from (jetted) AGN dominates over any plausible inventory of radiatively driven feedback modes. This holds for galaxies with mass comparable to that of the Milky Way and above, at low redshift (see Ch. 4.5 for more discussion). In the following section we proceed to analyse how this dominant feedback mode distributes over systems (galaxies, halos) of different mass, providing fresh insights into the mechanics of AGN feedback in the large scale structure.

4.4 Global energetics of radio AGN kinetic feedback

Having shown that the RLF computed via the incidence distributions agrees with the observed one, in this section we proceed to compute reliable estimates for the global energetics of radio AGN kinetic feedback.

4.4.1 Average jet power of massive galaxies

The average jet power, $\overline{Q}(M_*)$, released by massive galaxies is obtained by computing the expectation value of the modelled incidence distribution functions, shown in Fig. 4.1, as a function of stellar mass. For the compact radio AGN case, we take the average value (-1.4) of the power law slope of the incidence function across the two redshift bins and also determine the power law normalisation, given by the best-fit (and intrinsic scatter) to the observed incidence distributions in each redshift bin as a function of stellar mass (see Fig. 4.2, left). The observed distributions (Fig. 4.1) are constrained by our data only over a limited range in λ_{Jet}; thus we proceed with the extrapolation as described in Ch. 4.3 in the range $-8 < \log \lambda_{\mathrm{Jet}} < 0$. For the complex radio AGN case, we take the average values of all double power law parameters except for the normalisation, which we similarly determine from the best-fit as a function of stellar mass (see Fig. 4.3) and extrapolate as described in Ch. 4.3. The combined weighted average of $\overline{Q}(M_*)$ in each of the two redshift bins is the quantity plotted in Fig. 4.7, indicating the average jet power as a function of mass in the full redshift range, $0 < z < 0.4$.

Figure 4.7 shows the average jet power output of massive galaxies, considering LOFAR-detected and undetected sources (via the incidence), as a function of M_*. The intrinsic scatter on the best-fit normalisation as a function of stellar mass (Figs 4.2 and 4.3) is propagated through the calculation of $\overline{Q}$ to find the lower and upper uncertainty on this quantity, shown by the shaded regions. Light and dark blue shaded regions mark the compact and complex contributions, respectively, with the black-hatched region defining their sum, considering all massive galaxies. $\overline{Q}$ is also converted to an average jet kinetic energy released, $\overline{E_{\mathrm{Jet}}}$ (rightmost y-axis), with a multiplication by the look-back time between $0.0 < z < 0.4$. As a comparison, 0.1% (red dotted), 0.5% (purple dot-dashed) and 5% (red dot-dashed; Di Matteo et al. 2005) of the average radiative energy output as a function of M_* ($\overline{E_{\mathrm{rad}}}$) are also shown[2].

Figure 4.7 highlights that the average jet power of a population of massive galaxies in the local Universe ranges from $\log(\overline{Q}/[\mathrm{W}]) \sim 33.7 - 37.5$ and is dominated by compact (complex) radio AGN at stellar masses below (above) $\log M_*/M_\odot \sim 11.5$. More specifically, the total kinetic energy released by complex radio sources grows more steeply with stellar mass than that released by compact sources, reflecting the different M_* dependence of the incidence normalisation. The average λ_{Jet} is $\sim 10^{-5} - 10^{-3}$ across the mass range of this study.

When comparing to the radiative energy output of AGN, we see that the average jet power of the entire population also increasingly dominates over the 'canonical' fraction (0.5%) of average radiative power useful for feedback as stellar mass increases. Due to the observed near mass-independence of the X-ray AGN incidence (Aird et al. 2012), the radiative output as a function of M_* has a shallower (linear) slope than the kinetic output released by both compact and complex radio AGN. For the assumed 'canonical' value of 0.5% L_{bol}, the radiative mode of feedback only begins to dominate at low stellar masses, $\log M_*/M_\odot < 10.6$.

---

[2] $\overline{E_{\mathrm{rad}}}$ is computed by firstly integrating the specific black hole radiative power distribution, $p(\lambda_{\mathrm{Edd}} \mid M_*, z)$, to find the average λ_{Edd}, using Model C of Aird et al. (2013) at the average redshift ($z = 0.27$) of the radio AGN sample. Then the corresponding bolometric luminosity as a function of M_* is converted to $\overline{E_{\mathrm{rad}}}$, multiplying by the look-back time to $z = 0.4$.

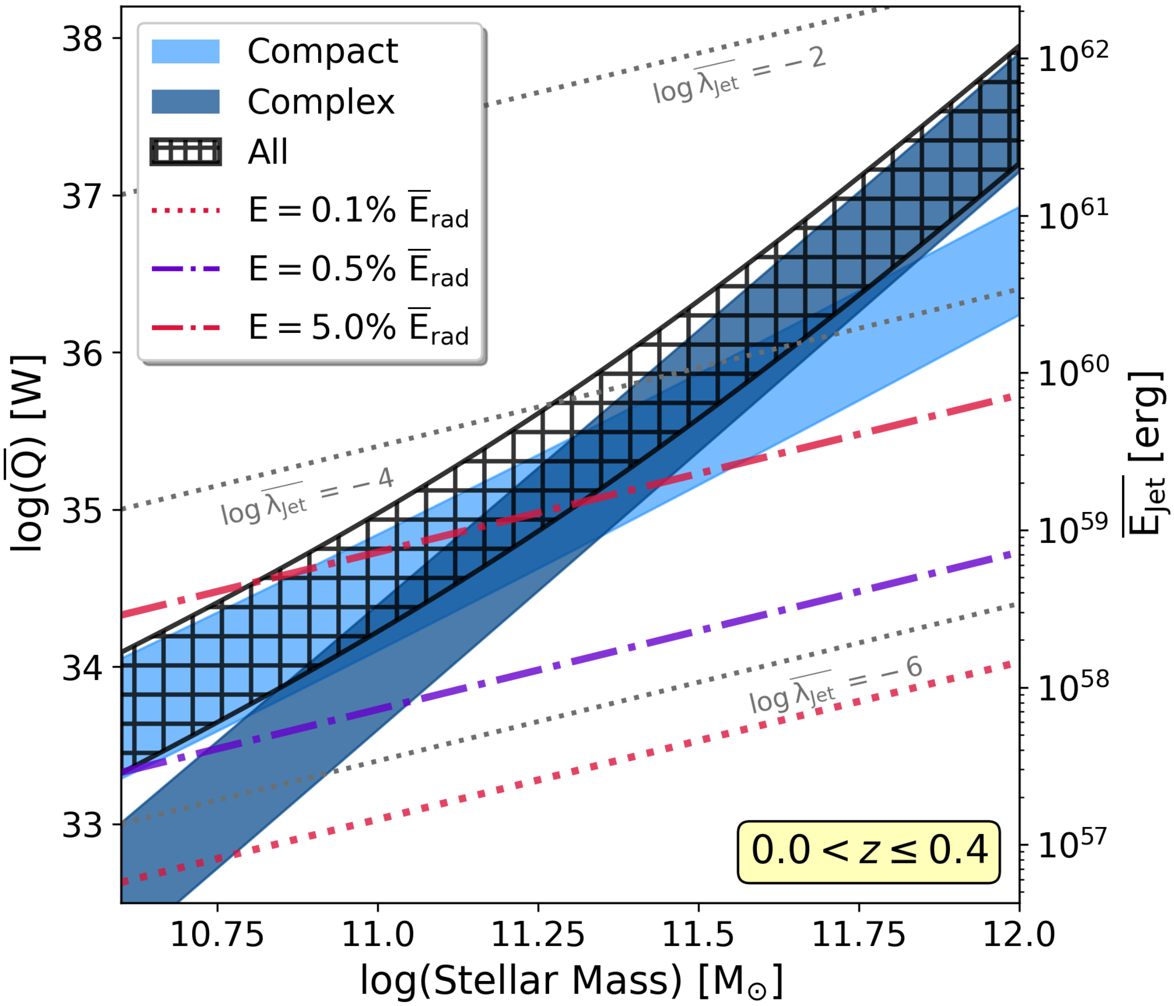


Figure 4.7: Average jet power released by radio AGN as a function of stellar mass for the population of all massive galaxies between $0 < z < 0.4$ (black-hatched area). The secondary y-axis (right) shows the corresponding average jet kinetic energy released, obtained by multiplying $\overline{Q}$ with the look-back time. The contributions from compact and complex radio AGN are shown with light and dark blue shaded regions, respectively. Red dotted and dot-dashed lines indicate the 0.1% and 5% of the average radiative energy output as a function of M_*, respectively. The purple dot-dashed line represents a 'canonical' 0.5% of the average radiative energy output, which is used for comparison to the kinetic output throughout the Chapter.

4.4.2 Disruptive kinetic feedback in massive galaxies and dark matter halos

We are now in the position to compare $\overline{E_{\rm Jet}}$ to the binding energy, $U_{\rm bin}$, of the host galaxy (on small scales) and host dark matter halo (on large scales). A total integrated kinetic energy injection in excess of a galaxy/halo binding energy would suggest that jetted radio AGN may have the ability to disrupt the stellar body and gas distribution of their host galaxies or halos, or at least deeply influence it. Therefore, we define this ratio as the 'small-scale (or large-scale) disruptive kinetic feedback efficiency', $\mathcal{F}_{\rm D,\ gal/halo} = \overline{E_{\rm Jet}}/U_{\rm bin,\ gal/halo}$.

In order to estimate the galaxies' binding energy, we use the result of Shi et al. (2021), who compile a sample of 752 objects from the literature, ranging over a wide variety of galaxy types, that have measurements for their M_*, effective radius[3] ($R_{\rm e}$) and the dynamical velocity ($V_{\rm e}$) at $R_{\rm e}$. They find a strong empirical correlation between $V_{\rm e}R_{\rm e}^{0.25}$ — effectively the fourth-root of galaxy binding energy, $U_{\rm bin,gal}$ — and the galaxy stellar mass:

$$\begin{aligned} U_{\rm bin,gal} &\approx \frac{GM_{\rm dyn,e}^2}{R_{\rm e}} \equiv \frac{R_{\rm e}V_{\rm e}^4}{G} \\ &= 4.62 \times 10^{48}\ {\rm erg}\ \left(\frac{V_{\rm e}R_{\rm e}^{0.25}}{{\rm km\ s^{-1}\ kpc^{0.25}}}\right)^4, \end{aligned} \tag{4.5}$$

where $M_{\rm dyn,e}$ is the dynamical mass within $R_{\rm e}$ and

$$V_{\rm e}R_{\rm e}^{0.25} = 15.7\ {\rm km\ s^{-1}\ kpc^{0.25}} \left(\frac{M_*}{M_0}\right)^{0.134} \left(1+\frac{M_*}{M_0}\right)^{0.272}, \tag{4.6}$$

with $M_0 = 2.5 \times 10^7\ {\rm M_\odot}$.

For halos, we use simple physical arguments to obtain the binding energy of the baryons in the halo as follows:

$$U_{\rm bin,halo} = \frac{3}{5}\ \frac{G\ f_{\rm gas}M_h^2}{R_{200c}}, \tag{4.7}$$

where the radius, R_{200c}, at 200 times the critical density of the Universe at a given redshift, $\rho_{200c}(z)$, is

$$R_{200c} = \frac{3}{4\pi} \left(\frac{M_h}{\rho_{200c}(z)}\right)^{\frac{1}{3}}. \tag{4.8}$$

Note that the gas fraction, $f_{\rm gas}$ is the ratio of the cosmological baryon and total matter density (Ω_b/Ω_M) and its universal value is ~ 0.16 (e.g. see review of Eckert et al. 2021, and references therein).

Figures 4.8 and 4.9 show the small-scale and large-scale disruptive kinetic feedback efficiency as a function of stellar mass and halo mass for the massive galaxy population (black-hatched area), with respective contributions from compact (light blue) and complex

---

[3] The effective radius of a galaxy is defined as the radius within which half of the total light is emitted.

(dark blue) radio AGN. At the horizontal black dashed line the jetted kinetic energy equals the host galaxy or halo binding energy. In Figure 4.9, we use the simulation results from Girelli et al. 2020 (their Eq. 6 with best-fit parameters from Table 1) to produce a mapping between M_* and M_h (the so-called Stellar to Halo Mass Relation, SHMR), where $M_h = M_{200c}$.

Thus far, we have assumed a universal hot gas fraction, however, it is known that $f_{\rm gas}(M_h)$ is not constant. Groups are found to be more hot gas depleted compared to the universal value, the one typically observed in clusters. To account for this, we adjust Eq. 4.7 using the best-fit relation[4] from Eckert et al. (2021, Eq. 11):

$$f_{\rm gas,500} = 0.079^{+0.026}_{-0.025} \times \left(\frac{M_{500c}}{10^{14} M_\odot}\right)^{0.22^{+0.06}_{-0.04}}. \tag{4.9}$$

We show the resulting $\mathcal{F}_D - M_h$ trend with a grey-hatched area in Fig. 4.9, albeit only for $\log M_h/M_\odot > 13$, which is where observational measurements for the gas fractions exist.

Upon computing the small-scale (galaxy-wide) disruptive kinetic feedback efficiency, shown in Figure 4.8, we observe that compact radio AGN have significant $\mathcal{F}_{\rm D,\ gal}$, ranging from $20-80\%$, whereas complex radio AGN have negligible feedback efficiency ($\mathcal{F}_{\rm D,\ gal} \approx 0.02$) for Milky-Way-like galaxies, but reach significant values for massive galaxies, exceeding 100% around $\log M_*/M_\odot \sim 11.5$. In comparison, upon computing the large-scale (halo-wide) disruptive feedback efficiency, shown in Fig. 4.9, we observe that $\overline{E_{\rm Jet}}$ for compact and complex radio AGN reaches at most $2-30\%$ of $U_{\rm bin,halo}$ at the lowest mass scales, and around $0.02-0.2\%$ at the highest mass scales.

4.4.3 Preventative kinetic feedback in halos

We also compare $\overline{E_{\rm Jet}}$ to the thermal (cooling) energy, $E_{\rm th}$, of the hot gas in the host halo, to measure the extent to which the cooling processes within the gaseous halo can be offset by jet heating, thus preventing the accumulation of fresh cold fuel for star-formation. $E_{\rm th}$ (as a function of halo mass) is obtained by multiplying the bolometric X-ray [$0.01 - 100$ keV] luminosity, $L_{\rm bol}$ —the total X-ray emission of hot gas, predominantly from thermal bremsstrahlung, recombination and two-photon decay— by the look-back time between $0.0 < z < 0.4$. $L_{\rm bol}$ as a function of halo mass is derived from the scaling relations[5] presented in Figs. 3 and 4 of Lovisari et al. (2021): $L_{\rm bol} \propto T^{3.04}$ and $M_{\rm 500c} \propto T^{1.76}$, where T is the intracluster medium (ICM) temperature.

If the total integrated kinetic energy injection is in excess of the halo thermal energy, it is conceivable that the gas could be maintained in a hot state, preventing current (and future) star formation. Therefore, we define this ratio as ‘preventative kinetic feedback

---

[4] Note that the $f_{\rm gas} - M_h$ relation uses M_{500c} (total mass within r_{500c}), so we use Eq. 6 and A2 from Bocquet et al. (2016) to convert M_{200c} from Girelli et al. (2020) to the required M_{500c}.

[5] Note that the $k_bT - M_h$ relation uses M_{500c} (total mass within r_{500c}), so we again convert M_{200c} to the required M_{500c}. We also scale the resulting $L_{\rm bol}$ to the correct cosmology and redshift interval via $E(z) = H(z)/H(z=0)$, where H is the Hubble parameter.

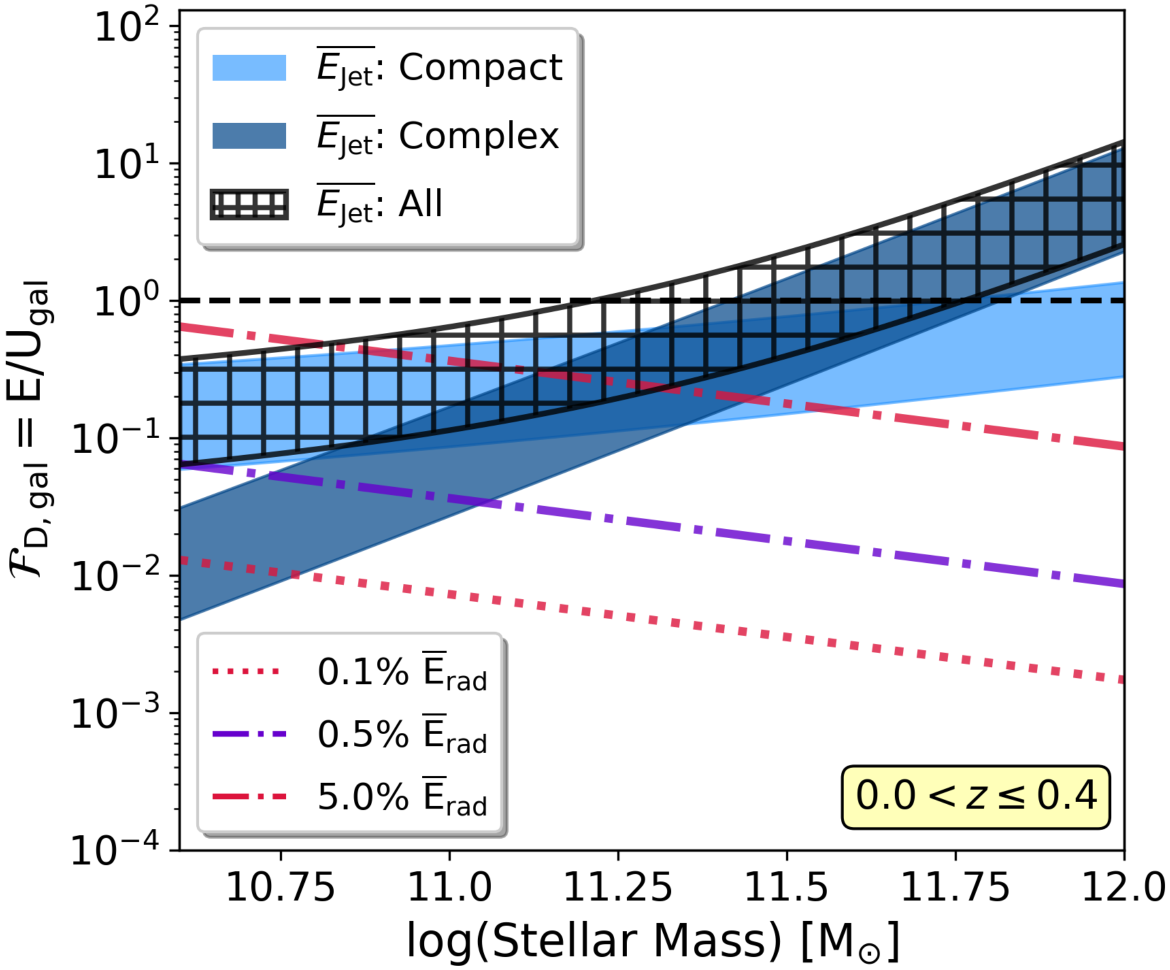


Figure 4.8: Small-scale disruptive kinetic feedback efficiency (i.e. ratio of jet kinetic energy divided by host galaxy binding energy) as a function of stellar mass for the massive galaxy population (black-hatched area) in the redshift range $0 < z < 0.4$. The respective contributions from compact and complex radio AGN are shown in light and dark blue shaded regions, respectively. The red and purple lines are the same as in Figure 4.7. The horizontal black dashed line marks the equality of the jetted kinetic energy and the host galaxy binding energy.

efficiency', $\mathcal{F}_{\rm P} = \overline{E_{\rm Jet}}/E_{\rm th}$. Note that the scaling relations derived by Lovisari et al. (2021) are fit to observed groups and clusters with $M_h > 10^{13}\ M_\odot$, so their application to the galaxy regime is merely an extrapolation, and one should be cautious not to over-interpret the results in this regime.

Figure 4.10 shows $\mathcal{F}_{\rm P}$ as a function of halo (and stellar) mass for the massive galaxy population (black-hatched area), with respective contributions from compact (light blue) and complex (dark blue) radio AGN. The horizontal black dashed line marks the equality of

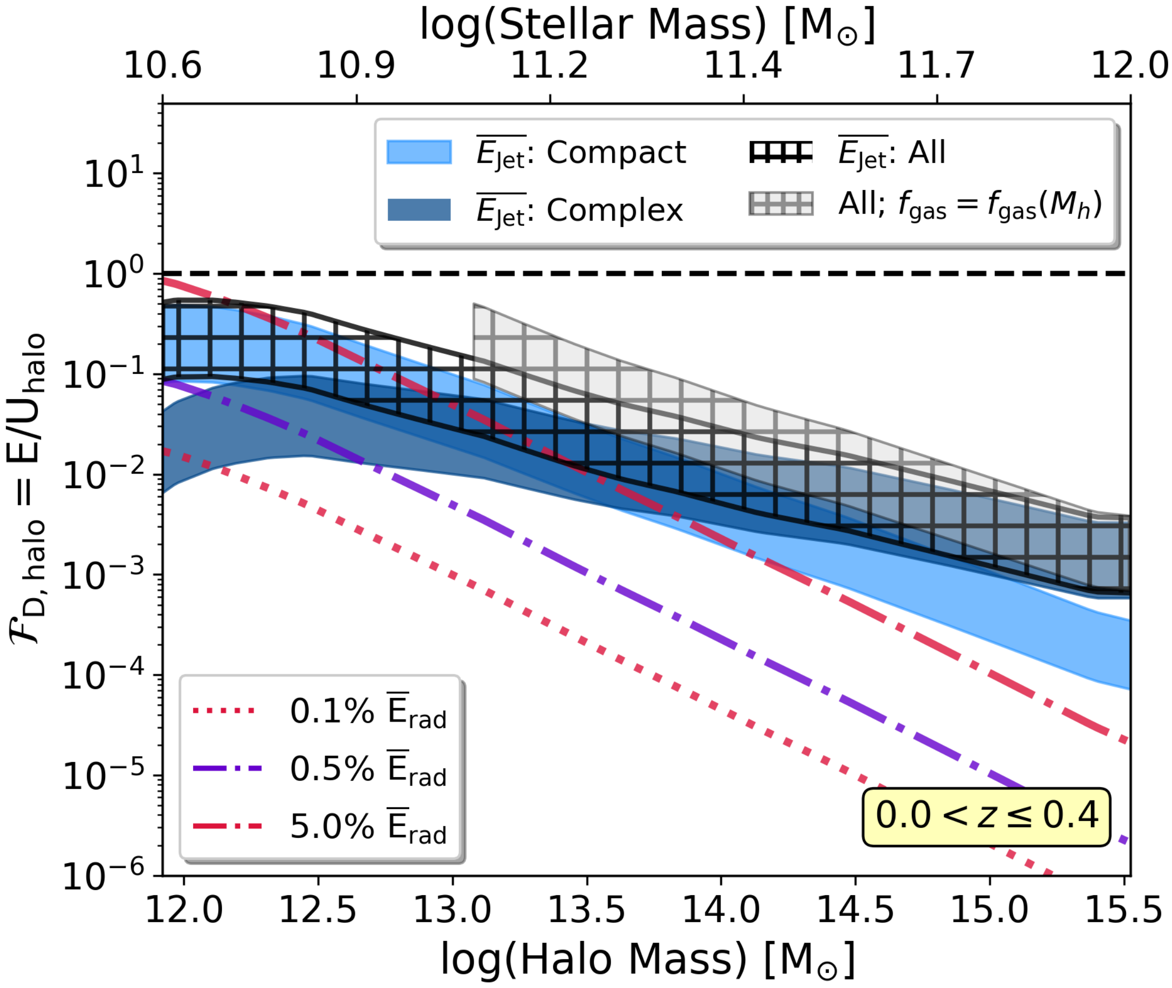


Figure 4.9: Large-scale disruptive kinetic feedback efficiency (i.e. ratio of jet kinetic energy divided by host halo binding energy) as a function of halo mass for the massive galaxy population (black-hatched area) in the redshift range $0 < z < 0.4$. The respective contributions from compact and complex radio AGN are shown in light and dark blue shaded regions respectively. The grey-shaded area shows the large-scale disruptive feedback efficiency on the baryonic halo mass (assuming a mass-dependent gas fraction). The red and purple lines are the same as in Figure 4.7. The horizontal black dashed line marks the equality of the jetted kinetic energy and the host halo binding energy.

the jetted kinetic energy and the thermal cooling energy. We deduce that, for the galaxy and small group regime, the jet kinetic energy exceeds the total thermal energy of the cooling gas in the halo, and so the preventative feedback effect of the collective population of radio AGN affects the global thermodynamics of these halos. On the other hand, for the large group and cluster regime ($\log(M_h/M_\odot) > 13.5$), $\mathcal{F}_P$ drops to $\sim 3 - 60\%$, suggesting that radio AGN do not inject sufficient energy to globally affect those larger halos.

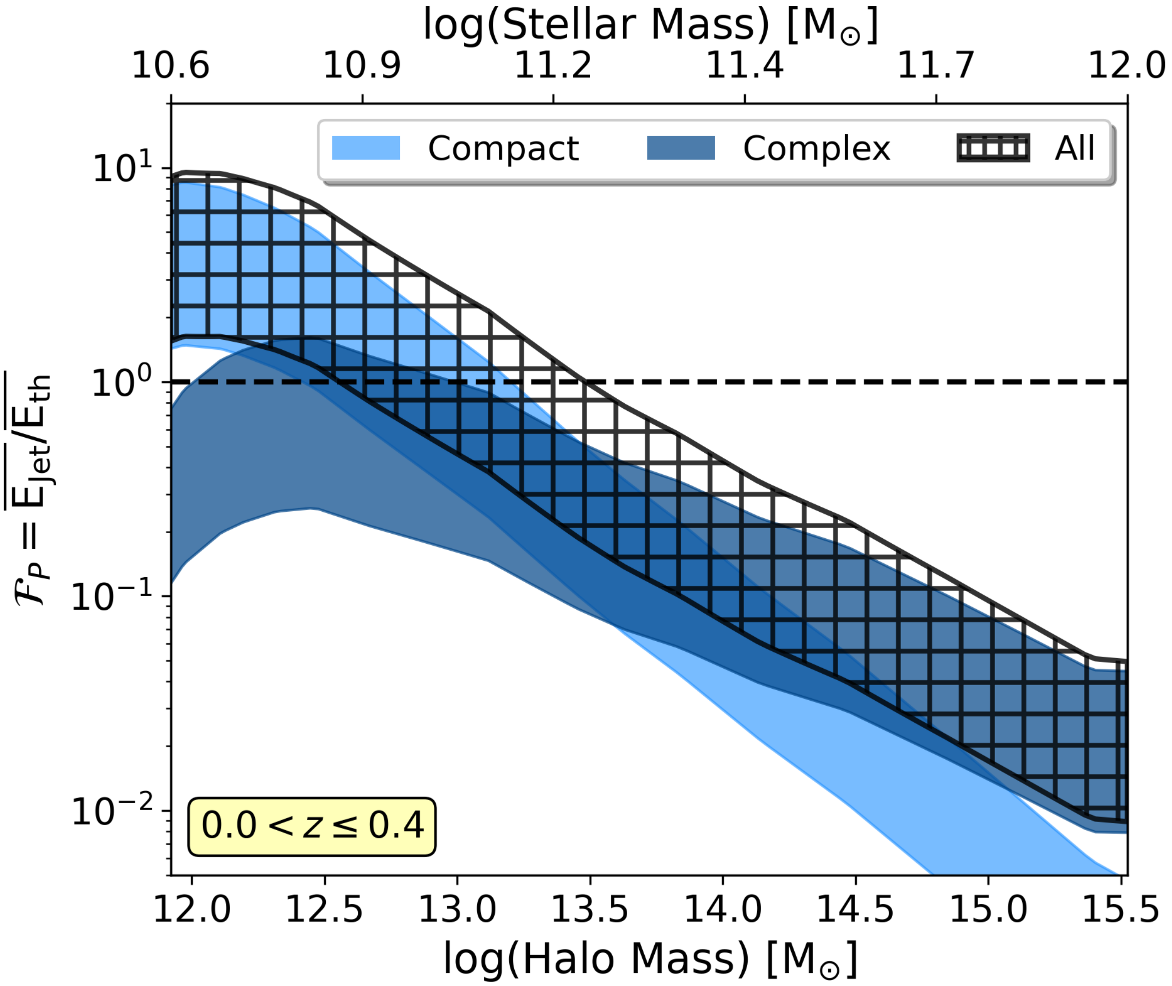


Figure 4.10: Preventative kinetic feedback efficiency, that is, the ratio of the average jet kinetic energy to the thermal (cooling) energy, as a function of halo mass for the massive galaxy population (black-hatched area). The respective contributions from the compact and complex radio AGN are shown in light and dark blue shaded regions. The horizontal black dashed line marks the equality between the heating provided by the AGN and the cooling of the halo gas.

Finally, in order to explore the local impact of preventative kinetic feedback, we compute the 'equivalence radius', R_{eq}, defined as the radius at which the integrated thermal cooling luminosity equals that of the average jet power: $\overline{Q} = L_{\mathrm{bol}}(R = R_{\mathrm{eq}})$. To do so, we express the X-ray surface brightness (SB) profile of the CGM of halos as a function of radius (R) using a β-profile (Cavaliere & Fusco-Femiano 1976):

$$S_{\mathrm{X}} = S_{\mathrm{X},0}\left[1 + \left(\frac{R}{R_{\mathrm{c}}}\right)^2\right]^{-3\beta+\frac{1}{2}}, \tag{4.10}$$

where $S_{\mathrm{X},0}$ is the central SB normalisation (set by $\int_0^{R_{500c}} S_{\mathrm{X}}\ 2\pi R\ dR = L_{\mathrm{bol}}$), R_c is the core radius, and β is the power law slopes outside R_c. An important assumption in this computation is that the average jet kinetic energy is fully contained within the spherical volume of radius R_{eq}. Additionally, the value of R_{eq} depends on the assumed SB profile. Given that there is currently no clear consensus in the literature about a 'universal SB profile' especially in the inner regions where non-gravitational processes such as AGN feedback can be present and are time-dependent (see e.g. Vikhlinin et al. 2006; Arnaud et al. 2010; Hudson et al. 2010; McDonald et al. 2014; Ghirardini et al. 2019; Käfer et al. 2019; Lehle et al. 2024, Eckert et al. in prep.), the simple but flexible β-model is deemed adequate to make global energetics statements in the context of this Chapter. Figure 4.11 presents several $R_{\mathrm{eq}}/R_{500\mathrm{c}}$ curves, derived using varying canonical SB profiles for groups and clusters, with the following parameter values: $\beta = 0.4, 2/3$ and $R_c = 0.02R_{500\mathrm{c}}, 0.1R_{500\mathrm{c}}$. These choices are discussed further in Ch. 4.5.

We also compute the 'normalised jet impact radius' of the sample of 'G9 radio AGN' from Chapter 2, by dividing the (projected) physical jet radii (R_{Jet}), defined[6] as `LOFAR_Maj`/2 [in units of kpc], by $R_{500\mathrm{c}}$. Figure 4.11 shows this 'normalised jet impact radius' as a function of halo mass, for the compact (light blue upper limits), complex (dark blue filled squares) and 'small' complex (white squares with dark blue edges) radio AGN, as well as for the subset of FRIIs (white crosses) and giant radio galaxies (GRGs; blue diamonds). R_{Jet} for compact radio AGN is an upper limit on the true physical size of the jets as the source is unresolved at the LOFAR-eFEDS $8'' \times 9''$ resolution. Similarly, we define a complex source to be 'small' if it has a physical size < 40 kpc (i.e. smaller than typical massive galaxy scales; van der Wel et al. 2014), roughly corresponding to the resolution limit. Sources which have $R_{\mathrm{Jet}}/R_{500\mathrm{c}} \sim 1 - 5\%$ are at very low ($z < 0.1$) redshifts.

Overall, taking the canonical SB profile with $\beta = 2/3$ and $R_c = 0.1R_{500\mathrm{c}}$ for the purpose of this discussion (used for computing the red curve on Fig. 4.11), R_{eq} reaches of order 10% and 1% of the halo $R_{500\mathrm{c}}$ for groups and clusters, respectively. This confirms that, on average, the jet heating in these more massive systems cannot offset the total thermal cooling energy, as the 'equivalence radii' of the jets are simply too small compared to the full size of the dark matter halo. On the other hand, the parameter space where $R_{\mathrm{Jet}} \leq R_{\mathrm{eq}}$ (red shaded region), i.e. where the jetted energy is fully contained within the sphere with radius R_{eq}, defines a region where the jets can still exert significant *local* impact on the thermodynamical heating and cooling balance. In fact, the majority of the 'G9 radio AGN' in this halo mass regime lie in the region marked by $R_{\mathrm{Jet}} \leq R_{\mathrm{eq}}$, meaning that jet heating is able to offset cooling flows in the cores of large groups and even the most massive clusters (see Ch. 4.5.3 for more discussion).

---

[6] `LOFAR_Maj` is the FWHM of the major axis of the source, in degrees. For sources with distant multi-components, the physical size is determined as the largest linear size, as described in Ch. 2.3.3.

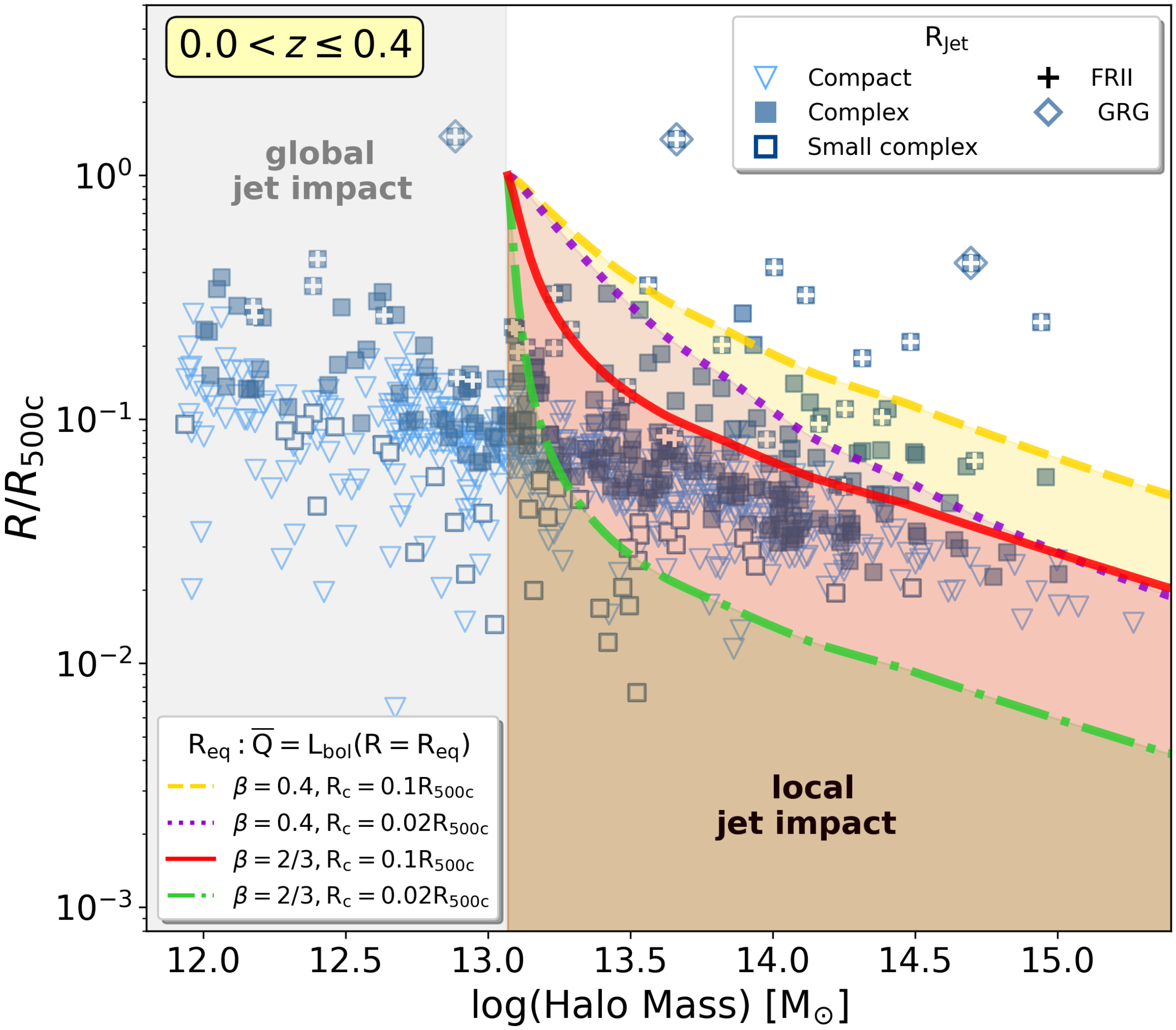


Figure 4.11: Curves showing the 'equivalence radius' ($R_{\rm eq}$), that is, the radius at which the integrated thermal cooling luminosity equals the average jet power, as a function of halo mass, normalised by $R_{\rm 500c}$. The red, yellow, green and purple curves denote $R_{\rm eq}$ for a set of β-profiles with canonical parameters: $\beta = 0.4, 2/3$, $R_c = 0.02R_{\rm 500c}, 0.1R_{\rm 500c}$. Given that the definition of $R_{\rm eq}$ requires $R_{\rm Jet} \leq R_{\rm eq}$, the region below each curve (shaded) denotes the parameter space where the radio jets can exert a significant impact on the *local* thermodynamical heating and cooling balance of the halo gas. The different subsets of G9 radio AGN from Chapter 2 are also shown with markers defined in the legend (see text for details).

4.5 Discussion

4.5.1 Interpretation and determination of $\overline{Q}(M_*)$

It is commonly assumed that it is the large, complex radio sources that exert the most powerful feedback on their surroundings. However, when considering the global energy budget from all radio AGN (Fig. 4.7), it is the numerous compact radio AGN, which are often, but not exclusively, less luminous in the radio (recall incidence, Fig. 4.1, and RLF, Fig. 4.5), that dominate the average jet power for all but the most massive galaxies ($\log M_*/M_\odot < 11.5$).

Generally, $\overline{Q}$ scales with M_* roughly as a power law with index ~ 2.5 (or $\overline{Q} \propto M_h$, for our choice of SHMR). Such a relation is similar in slope to that found in the semi-analytical models (SAMs) of Somerville et al. (2008), but is around $1-2$ orders of magnitude lower in normalisation (given our simplistic conversion of $M_{\mathrm{BH}} = 0.002 M_*$). However, our normalisation agrees well with past observational work by Allen et al. (2006) and Best et al. (2006) (see Fig. 11 in Somerville et al. 2008, for a summary). Importantly, Allen et al. (2006) conduct detailed *Chandra* X-ray spectral analysis to determine the 'Bondi' accretion rates (using also galaxy velocity dispersion to estimate black hole masses) and jet powers from cavity measurements of nine nearby, X-ray luminous elliptical galaxies, and find that a tight correlation exists between these two parameters. As a result, the authors conclude that the 'Bondi' formalism, the one usually implemented in SAMs, provides a suitable description of the accretion mechanisms present in their sample of luminous elliptical galaxies. Thus, our empirical $\overline{Q}(M_*)$ measurement may help constrain the radio mode feedback efficiency in SAMs and simulations (e.g. κ_{radio}, κ_{R} in Somerville et al. 2008; Croton et al. 2016, respectively), a free parameter which is usually fixed to match typical AGN-galaxy observational trends.

As mentioned in Ch. 4.3, the most important caveat of this work remains the uncertainty in the determination of jet power from the observable radio luminosity via Eq. 2.9. We refer the reader to Ch. 3.5.2 where we discuss this issue in detail. However, we note here that the difference in the normalisation of $Q - L_R$ relation (typically by the 'uncertainty factors' denoted by f_W or f_{cav}, detailed for example in Heckman & Best 2014) can explain the discrepancy between the $\log \Omega_{\mathrm{kin}}$ derived by Hardcastle et al. (2019) and the results of this Chapter and of Smolčić et al. 2017 (see their Figure 6 showing the effect of f_W on $\log \Omega_{\mathrm{kin}}$).

The second most important caveat is the unknown origin of radio emission in low-luminosity compact radio AGN (see the recent review of Panessa et al. 2019, and references therein). In this work, we assume that the radio luminosity observed in our sample of radio AGN, after thorough cleaning from possible star-formation related emission (see Fig. 2.12), is dominated by (unresolved) jetted emission. However, for those unresolved compact sources, the radio emission may originate from shocks, wind or accretion coronae (which in turn may also be viable feedback mechanisms; e.g. Brinkmann et al. 2000; Laor & Behar 2008; Zakamska & Greene 2014; Panessa et al. 2019; Kawamuro et al. 2022). Eq. 2.9 is most probably not valid for these emission processes, and their scaling with black hole mass, stellar mass, or other physical properties, is likely different from jetted emission. Nevertheless, this is only important for the lowest radio luminosity sources as past a certain threshold only a jetted origin is energetic enough to explain the radio emission.

A further minor uncertainty is the possible redshift evolution of the SMF and incidence within the finite redshift bins, resulting in an evolution of the duty cycle. However, there is only minor redshift evolution from $0 < z < 0.285$ and $0.285 < z < 0.4$ (see Figs. 4.2 and 4.3) and this is taken into account by the weighted average to obtain $\overline{Q}$; furthermore, the SMF does not significantly evolve over $0 < z < 0.4$ (see Fig. 5 from Ilbert et al. 2013).

Lastly, since low frequency LOFAR observations are able to detect radio emission from older electron populations, we cannot easily distinguish active versus remnant jets (without the use of radio data at other frequencies), especially in compact sources. This, therefore, may hamper the interpretations of the measured kinetic energy as a tracer of current feedback in galaxies (e.g. recently triggered or quenched star-formation), as it is likely a result of the *cumulative* jetted events across the lifetime of the source (potentially even from numerous triggering episodes).

4.5.2 Importance of the radio luminosity function synthesis

Observed luminosity functions are often used as metrics to calibrate hydrodynamic and semi-analytic simulation outputs. Yet, such simulations can have vastly different AGN accretion and feedback (‘sub-grid’) prescriptions and still successfully reproduce the same LFs due to degeneracies in the different fine-tuned input parameters (e.g. Vogelsberger et al. 2014; Schaye et al. 2015; Croton et al. 2016; Kaviraj et al. 2017; Springel et al. 2018; Davé et al. 2019; Habouzit et al. 2021). In addition, it is particularly difficult to predict radio continuum emission due to the complex physics and computing power limitations for resolving the sub-pc to pc scales involved (e.g. evolving the distribution of electrons and the magnetic fields self-consistently), and simultaneously simulating large volumes. Often such radio predictions are added in post-processing using empirical relations between the (known) black hole accretion rates and the (unknown) radio luminosities, for example in recent work by Slyz et al. (2015); Thomas et al. (2021) who manage to reproduce the observable RLF from Horizon-AGN and SIMBA cosmological hydrodynamical simulations, respectively. Nevertheless, these works highlight the still uncertain parameter choices in their prescriptions and the lack of physically motivated radio AGN model to couple jets with different accretion modes self-consistently (Thomas et al. 2021, see also Raouf et al. 2017).

Although our work cannot provide a detailed physical understanding of the mechanisms of jet powering and how exactly this ties into all accretion mode, galaxy and environmental properties, the mass-dependent jet powering mechanism in massive galaxies, as seen from the incidence distributions, is a key ingredient for understanding kinetic AGN feedback. Moreover, decomposing the RLF into stellar mass and radio morphology classes may be helpful to disentangle degeneracies in AGN feedback simulations.

Figure 4.5 also highlights the need for more deep and large-area radio surveys to probe the low and high luminosity end of the RLF, respectively. This would help in constraining the turnover of the radio AGN incidence at $\log \lambda_{\rm Jet} < -4$, a key physical constraint to understand the triggering of jets in the radiatively inefficient accretion mode.

Future work on larger samples that allow for a separation of the RLF and global energetics into quiescent and star-forming host galaxy subsets will be essential to understand the effect of star formation on jet powering (see also discussion about HERGs/LERGs in Best & Heckman 2012, and Thomas et al. 2021 for HERG/LERG RLF from SIMBA simulations). In Chapter 3, we already investigated the incidence of radio AGN in different host galaxy types, finding that the fraction of quiescent galaxies hosting radio AGN was similar to that of star-forming galaxies. However, lack of statistics prevented more detailed conclusions to be drawn, but given the differences in the SMF of quiescent and star-forming galaxies (Moustakas et al. 2013), deeper multiwavelength data on larger survey fields might reveal interesting conclusions.

4.5.3 Interpretation of the global energetics

As we have shown above, kinetic feedback dominates over radiative feedback in the local Universe for massive galaxies. This is a direct consequence of the total kinetic energy input of both compact and complex radio AGN sources scaling more steeply than the average radiative energy of AGN with galaxy mass (which follows $\overline{E_{\rm rad}} \propto M_*$, due to the near mass-independent X-ray AGN incidence; Aird et al. 2012), meaning that the 'radiative' mode of feedback only begins to dominate at low stellar masses, $\log M_*/M_\odot < 10.6$ (see purple dot-dashed line in Fig. 4.7). In fact, this agrees well with recent work of Petter et al. (2024), who find that powerful jet heating significantly dominates over quasar winds for $M_h \gtrsim 10^{13}\ h^{-1}\ M_\odot$, i.e. in the group and cluster regime (at $z < 2$, for $h = H_0/100$ km s^{-1}Mpc^{-1} $= 0.6766$; see their Fig. 11). This is interesting as they use a completely independent method to derive the clustering and halo masses via halo occupation distribution (HOD) modelling for sample of radio AGN from Best et al. (2023a), and later use this information to derive the energetics.

Our result also agrees with recent observational work by Buchner (2024); Kondapally et al. (2023); Heckman & Best (2023). Buchner (2024) combines distribution functions and scaling relations to derive average outflow rates as functions of mass and cosmic time, concluding that massive galaxies at $z < 0.3$ are predominately prevented from growing further by jet heating. Similarly, Kondapally et al. (2023) estimate a kinetic heating rate as a function of radio luminosity for different subsets of the radio AGN population (including quiescent low-excitation radio galaxies and all radio-excess AGN), as well as comparing to a suite of semi-analytical and hydrodynamical simulations, finding that AGN jets play a dominant role in AGN feedback at $z \lesssim 2$. These results are further supported by the work of Heckman & Best (2023), who compare the energy injection from massive stars and supernovae, radiation pressure and winds driven by AGN, and AGN radio jets, finding that the amount of $E_{\rm kin}$ for jets is an order of magnitude larger than from AGN winds at least up to $z \sim 1$. The authors further show that the maximum kinetic energy injection by jets occurs around $z \sim 1$, a lower redshift than the peak of star-formation and radiative

AGN activity ('cosmic noon', $z \sim 2-3$). They derive a time-integrated (i.e. across the entirety of cosmic history) total kinetic energy per unit volume due to jets of $U_{\rm Jet} = 2.6 \times 10^{57}$ erg Mpc^{-3}. Taking simply our local Universe estimate of $\log \Omega_{\rm kin}/[\rm W\ Mpc^{-3}] \sim 32.15$, propagated across the entirety of cosmic history, we derive a value of $U_{\rm Jet} = 6.1 \times 10^{56}$ erg Mpc^{-3}, which is of similar order (albeit lower, as the increasing $\Omega_{\rm kin}$ up to $z \sim 1$ is not considered in the calculation).

In terms of the small-scale 'disruptive' feedback, we showed in Fig. 4.8 that the compact radio AGN do not have enough jet kinetic energy to surpass the binding energy of galaxies, but $\overline{E_{\rm Jet}}$ reaches significant fractions of $U_{\rm bin,gal}$ for higher and higher stellar masses (see also Heckman & Best 2023). This means that, although energetically the gas cannot be unbound from such galaxies, it may be significantly disrupted in its kinematics and distribution, potentially impacting both the star formation and the central gas supply for fuelling the AGN (e.g. McNamara et al. 2014; Morganti et al. 2015). Factors that can affect the extent of gas disruption include the morphology, collimation, entrainment and mass-loading of the jet (De Young 1986; Bicknell 1986; Bowman et al. 1996; Hubbard & Blackman 2006; O'Shea et al. 2025), as well as the structure of the surrounding ISM or inter-group medium (IGrM; e.g. Tanner & Weaver 2022; Dutta et al. 2024; Hardcastle & Krause 2013; English et al. 2016; Croston et al. 2019; Gaspari et al. 2020; Morris et al. 2022; Mingo et al. 2022), but further detailed discussion is out of the scope of this Chapter. As for the complex sources, $\mathcal{F}_{\rm D,gal} > 1$ for the highest masses, $\log(M_*/M_\odot) > 11.5$. However, from the analysis in Ch. 4.4.3, around 86% of complex LOFAR radio AGN have physical sizes (at the LOFAR-eFEDS resolution) greater than ~ 40 kpc (see Fig. 3.11), extending beyond typical galaxy scales (van der Wel et al. 2014). This suggests that a large fraction of the jet kinetic energy (depending on some of the factors mentioned above) may be deposited outside the galaxy, and therefore a comparison with the galactic binding energy may not provide physically meaningful results for all the complex radio AGN.

This is why we also compute the large-scale disruptive feedback efficiency (Fig. 4.9), which shows that the average jet kinetic energy of both compact and complex radio AGN is largely insufficient at all mass scales to unbind all the gas from the host halo: the binding energy of dark matter halos is simply too large compared to the kinetic energy that radio AGN (at low redshift) can provide. At galaxy and group scales $\mathcal{F}_{\rm D,\ halo}$ can reach up to $\sim 30\%$ (considering a mass-dependent $f_{\rm gas}$), but a steep decline with M_h means that for the cluster regime $\mathcal{F}_{\rm D,\ halo}$ is only $\sim 0.2\%$. This is in line with the universal gas fractions observed in clusters (whereas groups are preferentially gas-depleted; e.g. Eckert et al. 2021), as the deep potential wells maintain the primordial ratio of cosmological baryon-to-total-matter densities.

Thus far, we have shown that, *globally*, radio AGN in the local Universe do not exert enough small- and large-scale disruptive feedback to their host galaxies and halos. Nevertheless, radio AGN may be an important source of 'preventative' feedback as shown by Figures 4.10 and 4.11. Given that $U_{\rm bin} \propto M_h^2/R_{\rm 200c} \sim M_h^{1.67}$ and $L_{\rm bol} \propto M_h^{1.73}$ (Lovisari et al. 2021), $U_{\rm bin}$ scales quasi-linearly with $L_{\rm bol}$, i.e. with the total thermal energy of hot halos. Therefore, $\mathcal{F}_{\rm P}$ follows a similarly steep evolution with M_h as $\mathcal{F}_{\rm D,\ halo}$, but now the normalisation is higher, as shown in Fig. 4.10, reflecting the simple fact that for massive halos, the thermal energy of the hot gas is smaller than their binding energy.

In the galaxy and small group regime, we measure $\mathcal{F}_{\mathrm{P}} \gtrsim 1$, meaning the jet energy can be greater than the E_{th}. This suggests that the heating provided by the jet can efficiently offset the cooling in these smaller halos. In the large group and cluster regime, on the other hand, $\overline{E_{\mathrm{jet}}}$ reaches only $\sim 3 - 60\%$ of the total thermal energy of the cooling gas in the halo, suggesting that radio AGN cannot impact the global thermodynamical equilibrium of these systems. This can be explained by Fig. 4.11, where we show that $R_{\mathrm{eq}}/R_{500\mathrm{c}}$ declines from of order 10% at group scales to of order 1% for cluster scales, meaning that as the halo mass increases, the radial impact of the jets decrease (see also e.g. Eckert et al. 2021).

However, even though the reach of the jet is small compared to the virial radius of the halo, we know from X-ray observations of clusters (and simulations, e.g Croton et al. 2006; Bower et al. 2006; Somerville et al. 2008; Croton et al. 2016) that preventative feedback is a necessity to prevent ‘catastrophic cooling flows’ (Fabian 1994; Peterson et al. 2004; McNamara & Nulsen 2012), which would be expected given the short radiative cooling timescales compared to the cluster ages. In fact, taking again the canonical SB profile with $\beta = 2/3$ and $R_c = 0.1R_{500\mathrm{c}}$ as an example, most G9 radio AGN lie in the region where $R_{\mathrm{Jet}} \leq R_{\mathrm{eq}}$. Therefore, low-redshift radio AGN living in groups and even in the most massive clusters, do have the power to exert significant *local* impact on the thermodynamical balance in the central cores of their host halos. Interestingly, it is mainly the FRII and GRG sources that preferentially lie above the red curve ($R_{\mathrm{Jet}} > R_{\mathrm{eq}}$), which is not surprising given that the spatial distribution of their jet kinetic energy is edge- rather than core-brightened, meaning that they deposit the bulk of their energy further from the halo core. This, therefore, could impact their ability to effectively offset strong central cooling flows, and thereby also affect the fuelling of the central AGN, although this remains to be further investigated, as the feeding and feedback cycle is complex and time-dependent (e.g. Gaspari et al. 2020; Mingo et al. 2022).

Moreover, the R_{Jet} distribution of the observed radio AGN traces very well the trends of the red canonical R_{eq} curve, clustering just below the curve. This may be a coincidence of the given LOFAR resolution, or be a potential indication of the natural heating and cooling equilibrium reached by the jet and the inner regions of the gaseous halo. Although a comparison of individual realisations of jetted AGN (the data points), which may be variable in time, and population-averaged quantities (the curves) should be done with caution. Future high resolution LOFAR VLBI studies (Morabito et al. 2022; Sweijen et al. 2022), capable of resolving the detailed jet structures on large samples of radio AGN, will be needed to shine light on this matter.

Past works have also made local and global statements about the heating-cooling balance in the environments of radio AGN. For example, Hardcastle et al. (2019) integrate the Schechter profile of the local cluster luminosity function obtained by Böhringer et al. (2014) to get a cooling luminosity of 2×10^{31} W Mpc^{-3} and conclude that the derived heating rate (see green star on Figure 4.6) from their sample of 23,344 LOFAR radio AGN can offset (in statistical terms) the radiative cooling in these systems (see also Smolčić et al. 2017; Butler et al. 2019; Croton et al. 2016). Similarly, Dunn & Fabian (2006) use detailed X-ray and radio observations of X-ray cavities and spatially coincident radio bubbles to conclude that the average radius to which radio bubbles could offset X-ray cooling ($r_{\rm heat}$) was $r_{\rm heat}/r_{\rm cool} = 0.86 \pm 0.11$, (where $r_{\rm cool}$ is the radius within which the cooling time equals 3 Gyr) and 10/16 clusters had $r_{\rm heat}/r_{\rm cool} \gtrsim 1$ (see also McNamara & Nulsen 2007, for a review).

4.5.4 Jet energy deposition efficiency, cluster profiles and halo occupation distributions

Throughout this Chapter, we assume that 100% of the jet kinetic energy couples to the surrounding medium, via thermal dissipation, sound waves, shocks, turbulence, release of cosmic rays and other mechanisms which can also indirectly transfer the jet energy to the ISM and/or circumgalactic medium (CGM; e.g De Young 1986; Bicknell 1986; Wagner et al. 2012; Wykes et al. 2013; Zhuravleva et al. 2014; Jacob & Pfrommer 2017; Perucho 2019; Hlavacek-Larrondo et al. 2022, and references therein). Without going into details of these physical processes, this must be a reasonable assumption if the jet fully decelerates within the host galaxy/halo. In fact, other than two giant radio galaxies (see Ch. 2.3.3), all G9 radio AGN, from which the energetics results in this work are computed, have $R_{\rm Jet}/R_{\rm 500c} < 1$, so the aforementioned assumption is reasonable, at least when considering the (large) halo scales. We also note that the distinction between preventative and disruptive feedback is somewhat artificial and the total jet kinetic energy is not distributed in a mutually exclusive way between the two; the physics of these two processes are closely interlinked.

Regarding Figure 4.11, our choices for $\beta = 0.4, 2/3$ are justified as they represent well the ranges of SB profile slopes found in past literature (e.g. Sanders et al. 2025; Vladutescu-Zopp et al. 2025, and references therein). The range of core radii is also as of yet uncertain, and can be degenerate with β. We chose to present R_c in the range of $0.02 - 0.1R_{\rm 500c}$ as these are some canonical values for cool-core and non-cool core clusters (e.g. Wang et al. 2023). However, it seems that the combination of $\beta = 2/3$ and $R_c = 0.02R_{\rm 500c}$ (green curve) represents a SB profile that is unphysically centrally concentrated, as most observed radio AGN jets lie above the green curve. Using a more complex Vikhlinin et al. (2006) model for the SB, with variable inner ($R < R_c$) power law slope, α, would also show similarly low radii of equivalence for $\alpha > 0$, as more of the luminosity would be concentrated in a smaller volume.

Interestingly, Wang et al. (2023) quantifies the cool-core condensation radius ($R_{\rm ccc}$: radius within which the cooling time equals the turbulence eddy turnover time[7]) and quenching cooling flow radius ($R_{\rm qcf}$: radius within which the cooling time is 25 times the free fall time[8]) for a range of massive, nearby clusters ($1.3 \times 10^{14} < M_{\rm 500c}/M_\odot < 16.6 \times 10^{14}; 0.03 < z < 0.29$). They find typical values of $0.01 < R_{\rm ccc}/R_{\rm 500c} < 0.05$ and $0.02 < R_{\rm qcf}/R_{\rm 500c} < 0.13$, both of which span over the observed parameter space of $R_{\rm Jet}/R_{\rm 500c}$ for those radio AGN in the cluster regime. This highlights, in an independent

---

[7] $R_{\rm ccc}$ is a measure of the balance between feeding and feedback processes, generating turbulent condensation rain and related chaotic cold accretion (Wang et al. 2023; Gaspari et al. 2018)

[8] $R_{\rm qcf}$ encompasses the region of thermally unstable cooling (Wang et al. 2023; Voit et al. 2015)

manner, the effectiveness of jet heating offsetting cooling in the central regions of massive clusters.

Finally, we have taken all radio AGN so far to be central, and not satellite, galaxies in their host halos. Halo occupation distribution (HOD) models predict the number of satellites to increase as a function of halo mass, and it is currently unknown how the occupation fraction of radio AGN changes between centrals and satellites (e.g. Berlind et al. 2003; Zheng et al. 2005; Comparat et al. 2023). Radio AGN triggering in satellite galaxies is also likely to work differently to centrals due the local distribution and/or dynamics of the gas (e.g. gas stripping at the outskirts of clusters), and different formation histories, potentially impacting the incidence distributions and thereby our statements on the global energetics. For example, de Vos et al. (2024) (and related simulation work by Rihtaršič et al. 2024) find that the observed LOFAR radio AGN fraction peaks near the cluster core (as is expected for reasons described below) but then *declines* before rising again in the cluster outskirts (∼10 R_{500}), potentially due to lower velocity dispersion at these radii allowing for more mergers to occur and potentially trigger radio AGN (see also e.g. Ramos Almeida et al. 2011, 2012; Pierce et al. 2022, for more discussion on the importance of mergers for radio AGN triggering). However, Best et al. (2007) clearly showed that brightest group or cluster (BGG, BCG) galaxies have a higher probability to host radio AGN compared to other galaxies of the same mass (see also e.g. Burns 1990; Sun 2009; Smolčić et al. 2011), likely due to the more effective condensation of cold clouds from the hot halo or direct hot gas fuelling near the centres of the halo (Hardcastle et al. 2007; Gaspari et al. 2020). In fact, upon matching the G9 radio AGN sample to the GAMA Groups (`G3CGalv10`) and Friends of Friends (FoF; `G3CFoFGroupv10`) catalogues (Robotham et al. 2011), 78% of sources are associated with the BCG. Therefore, treating our radio AGN as (mostly) centrals appears justified. Future work computing the incidence of radio AGN in centrals and satellites separately will be crucial to rigorously test this effect and provide a more in depth interpretation of Fig. 4.11.

4.6 Summary

In this Chapter, we have used the radio AGN incidence as a function of specific black hole kinetic power, stellar mass and radio morphology to quantify the average jet power of massive galaxies in the local Universe and interpret this in the context of AGN kinetic feedback energy balance on galaxy and halo scales.

As in Chapter 3, we show that the incidence of radio AGN is mass dependent, whereby higher mass galaxies are more likely to host radio AGN across the λ_{Jet} range. However, the compact and complex radio AGN follow different incidence distributions (Fig. 4.1). The former follows a steep power law distribution with slope ~ -1.4 dominated by jets of lower power, whereas the latter follows a double power-law-like distribution with shallower faint-end slope (~ 0.2), but reaching to high jet powers.

We then synthesise the radio AGN luminosity function by convolving the radio AGN incidences with the stellar mass function, allowing us to decompose, for the first time, the

RLF as a function of stellar mass and radio morphology (Fig. 4.5). We find that in the luminosity range $\log(L_{150\mathrm{MHz}}/[\mathrm{W\ Hz}^{-1}]) \sim 23.5-25$, radio AGN with $\log(M_*/M_\odot) \sim 11.4$ dominate the RLF in number density. We also find that the compact (complex) radio AGN contribute dominantly in the range below (above) $\log(L_{150\mathrm{MHz}}/[\mathrm{W\ Hz}^{-1}]) \sim 24.5$, equivalent to a jet power of $\log(Q/[\mathrm{W}]) \sim 36.7$. Our RLF and its integrated quantities, such as the average kinetic energy density Ω_{kin}, are in good agreement with past literature.

Importantly, we find that compact radio AGN dominate the average injected kinetic jet power for all but the most massive galaxies, $\log M_*/M_\odot < 11.5$, at $z < 0.4$. The total kinetic energy released by complex radio sources grows more steeply with stellar mass than that released by compact sources, and both scale more steeply than the average radiative energy (Fig. 4.7). Taking a 'canonical' fraction of radiative output from multi-phase AGN winds, 0.5% L_{bol}, we also see that this 'radiative' mode of feedback only begins to dominate at low stellar masses, $\log M_*/M_\odot < 10.6$. In an integrated sense, we quantitatively show that kinetic feedback dominates over radiative feedback in the local Universe (Fig. 4.6).

We then define three metrics to gauge the efficacy of kinetic feedback to either disrupt the galaxy-wide and halo-wide gas distribution, potentially driving it out of the system completely ('disruptive feedback'), or to prevent the gas from cooling and forming stars ('preventative feedback').

We find that compact radio AGN do not have enough jet kinetic energy to efficiently unbind gas from galaxies across the probed mass scale, although they reach a level that could be sufficient to significantly affect the local gas distribution. On the other hand, complex sources show $\mathcal{F}_{\mathrm{D,\ gal}} > 100\%$ for $\log M_*/M_\odot \sim 11.5$ (Fig. 4.8). Nevertheless, given that the majority ($\sim 86\%$) of complex sources have physical sizes larger than about 40 kpc, they extend on average past the stellar body of typical galaxies, meaning that the jet kinetic energy is mostly deposited on larger scales. Therefore, we consider also the efficacy of the halo-wide disruptive feedback and find that at no mass range are jets energetic enough to unbind gas from entire halos (Fig. 4.9). Nevertheless, $\mathcal{F}_{\mathrm{D,halo}}$ does reach $\sim 2-30\%$ at galaxy and group scales, which is in line with the possibility of reducing the gas fractions in these systems away from the cosmic value.

Lastly, we show that the jet kinetic energy of AGN may be sufficient to offset the cooling in halos (Fig. 4.10). In fact, in the galaxy and small groups regime, $\overline{E}_{\mathrm{Jet}}(M_h) \gtrsim E_{\mathrm{th}}$, meaning that the kinetic energy can impact the heating and cooling balance of the gas on global scales, i.e. throughout the halos. This is not the case at cluster scales, however, where $\mathcal{F}_{\mathrm{P}}$ is just a few percent. Nonetheless, by comparing the physical sizes of observed radio jets in our complete sample of radio AGN to the radius within which the integrated surface brightness of the halo is equal to $\overline{Q}$, we find that they preferentially populate the region where $R_{\mathrm{Jet}} \leq R_{\mathrm{eq}}$ (Fig. 4.11). Acknowledging the caveats present in the assumption of a simple β-profile for halos across the mass scale, we conclude that jetted AGN feedback can contribute significantly to the *local* heating of the gas in the central cores of groups and even the most massive clusters, where we do indeed observe strong cool-cores and jet-inflated bubbles or cavities.

Overall, we have shown that combining AGN incidence measures — enabled by the well-characterised and complete samples from Chapter 2 — with simple assumptions about host galaxy and halo properties provides a powerful approach for making meaningful physical statements about the global kinetic energy budget of the local Universe. Future work on expanding the samples both to high sensitivity and larger volumes, as well as drawing parallels between observational work and AGN feedback simulations, will be vital to complement this knowledge in the currently unattainable parameter and simulation space.

Chapter 5

The incidence of eROSITA X-ray AGN in the local Universe: from dwarf to massive galaxies

In this Chapter, we push studies of the X-ray AGN incidence as a function of λ_{Edd}, previously seen in Ch. 3, to an as-of-yet unexplored parameter space of low-mass galaxies with $\log M_*/M_\odot \leq 10$. We start by selecting an optical parent sample from LS10 consisting of ~ 5.35 million low- and high-mass ($\log M_*/M_\odot > 10$) galaxies with z-band fluxes of $z \leq 20$ mag and low redshifts $0.03 \leq z \leq 0.2$. We place particular emphasis on the detailed characterisation of our sample, including: (i) estimating unbiased physical galaxy properties through SED fitting; (ii) rigorous cleaning and validation of the X-ray aperture photometry and associations with optical host galaxy counterparts; and (iii) building a mass- and luminosity- complete sample to derive robust estimates of their λ_{Edd} distribution.

We perform X-ray aperture photometry, using X-ray data from the deepest all-sky scan of eROSITA (eRASS:4), centred on the parent galaxy optical coordinates. We find 874 X-ray emitting low-mass galaxies, more than 600 of them newly discovered, with some reaching 2–10 keV luminosities above 10^{43} erg s^{-1}. This unambiguously demonstrates that AGN are present even in low-mass galaxies, with a non-negligible fraction accreting vigorously. We also find $12,618$ X-ray AGN in high-mass galaxies.

By using a Bayesian framework that takes into account X-ray information from all parent sample galaxies, we constrain the specific accretion rate distribution, $p(\log \lambda_{\mathrm{Edd}}|M_*, z)$, across a wide range of λ_{Edd}. We robustly detect a break at high λ_{Edd}, possibly indicating Eddington-limited, self-regulated black hole growth. Integrating $p(\log \lambda_{\mathrm{Edd}}|M_*, z)$ above $\lambda_{\mathrm{Edd}} \geq 10^{-2}$ and $\geq 10^{-1}$, we derive the cumulative AGN fraction as a function of stellar mass. Interestingly, we find a peak in this cumulative AGN fraction around $\log M_*/M_\odot \sim 10 - 10.5$, for such moderately- and highly-accreting AGN, potentially highlighting a decrease in the efficiency of AGN fuelling at both very low and very high galaxy masses.

The contents of this Chapter will be published in Igo et al. (to be submitted to *A&A*).

5.1 Building the parent galaxy sample

This Chapter discusses the steps taken to build the parent sample of galaxies used to ultimately probe the incidence of X-ray detected AGN in both low- and high-mass galaxies. It includes: (i) a description of the optical selection criteria and removal of contaminants; (ii) the creation of a compilation of good-quality, extragalactic spectroscopic redshifts to supplement the photometric redshifts available for the optical survey used; and (iii) the calculation of galaxy properties, such as stellar masses and star formation rates, using two different methods.

We select our parent sample of galaxies from LS10, which includes photometry in the *g, r, i, z* bands and WISE forced photometry at the optical source coordinates, following Lang (2014); Lang et al. (2016), at 3.4 μm, 4.6 μm, 12 μm and 22 μm. The novel features of LS10, compared to previous releases, are its extended footprint, deeper coverage and added *i*-band observations. In addition to the observations completed by the Beijing-Arizona Sky Survey (BASS), the DECam Legacy Survey (DECaLS) and the Mayall z-band Legacy Survey (MzLS), the DECam eROSITA Survey (DeROSITAS; Zenteno et al. 2025), ensures 5σ depths across (almost) the entire eROSITA-DE footprint in the western Galactic hemisphere of 22.7, 23.2, 23.3, 22.5 mag in the *g, r, i, z* bands, respectively. For an overview of the depth and coverage of LS10 in the eROSITA-DE footprint, excluding the Galactic Plane (Galactic latitude $|\mathrm{b}| < 20°$), that is the area covering the entire extragalactic sky in the southern equatorial hemisphere with declination $< 32.375°$, see Figure 1 in Saxena et al. (2024).

To build our parent galaxy sample, we apply the following selection criteria to the LS10 sources:

1. Located in the western Galactic hemisphere ($179.9442° \leq$ galactic longitude $\leq 359.9442°$), to overlap with eROSITA-DE footprint (Merloni et al. 2024).

2. Has been observed with the *g, r, z, W1* bands (i.e. `NOBS_{g,r,z,W1}>0`).

3. Not associated with known problematic photometry[1]. This includes removing objects with: `MASKBIT` 0 (secondary detections), `MASKBIT` 1 (objects touching Tycho sources with `MAG_VT`$<$ 13 and Gaia stars with G$<$ 13), `MASKBIT` 12 (object touching a pixel in a *Siena Galaxy Atlas* large galaxy), `MASKBIT` 13 (object touching a pixel in a globular cluster), as well as `FITBITS` 1, 6, 8, 9, 10, 11, 12, 13.

4. Not a GAIA duplicate source (`TYPE!=DUP`).

5. Has high signal-to-noise (S/N) in four bands[2]:
`FLUX_{g,r,z,W1}*sqrt(FLUX_IVAR_{g,r,z,W1})`$>$3.

---

[1] DR10 bitmasks: `https://www.legacysurvey.org/dr10/bitmasks/`

[2] We denote the square root by `sqrt`.

6. Does not have high parallax nor proper motion (PM), i.e. a stellar astrometry cut: `PARALLAX*sqrt(PARALLAX_IVAR) < 5 &` `sqrt[(PMRA*sqrt(PMRA_IVAR))`^{2}`+ (PMDEC*sqrt(PMDEC_IVAR))`^{2}`] < 5`.
7. Does not have stellar colours according to Salvato et al. (2022b): $z - W1 > (0.8 \times (g - r) - 1.2)$
8. Has low Galactic absorption: `E(B-V)`< 0.1.
9. Satisfies the z-band magnitude cut: $12 < \mathrm{z\_mag} \leq 20$.
10. Satisfies the redshift cut: $0.03 \leq z \leq 0.2$ (spectroscopic, if existing, otherwise photometric).
11. Does not lie within the mask defined by large foreground galaxies from the Heraklion Extragalactic Catalogue (HECATE v2.0; Kyritsis et al. 2025).
12. Does not lie within a mask defined by the R_{500} of eRASS1 X-ray clusters nor in the known spurious over-dense regions flagged in the eRASS1 catalogue (see text below for details).

Figure 5.1 shows the sky density distribution of the parent sample that consists of $5,352,526$ galaxies obtained applying the above cuts, in galactic coordinates[3].

We select our sources in the z-band, as per criterion 9 above, as it offers more uniform sky coverage (see Fig. 1 in Saxena et al. 2024). The bright- and faint-end limits are chosen to avoid saturation effects and to remain competitive with Dark Energy Spectroscopic Instrument (DESI) Bright Galaxy Survey (BGS; Hahn et al. 2023) which selects objects with an r-band limit of 19.5 mag, respectively. To illustrate the effect of selection criteria $1-9$, we take four representative extragalactic (eROSITA-DE) LS10 sweep files (each $0.25° \times 0.25°$). An initial z-band magnitude cut retains around 1.1 million sources ($\sim$12%). Subsequent photometric and observational quality cuts (criteria 2, 3, 4, and 5) remove around 65% of these, leaving roughly $385,000$ sources. The non-stellar selection criteria (cuts 6 and 7) further reduce the sample to around $207,000$ sources. After additionally applying criterion 8, only $\sim$2% of the total LS10 sources within these regions contribute to the parent sample. This value is sensitive to Galactic latitude, with higher losses in areas of high extinction; in this example, 87% of the sources in the selected sweep files fall within low-extinction regions (`E(B-V)`< 0.1).

We focus on the low-redshift universe of $z \leq 0.2$, where we can build complete samples of low-mass galaxies (see Ch. 5.2.2), but do not explore $z < 0.03$, as the photometric redshifts (photo-zs) are known to be problematic in the very local Universe (e.g. Hearin et al. 2010; Dahlen et al. 2013). Zhou et al. (2021, 2023) compute photo-zs for all LS10 galaxies with good quality photometric information and show excellent agreement with

---

[3] We note that around 400 sources are not shown via the Mollweide projection as the eROSITA ‘western Galactic hemisphere’ cuts through the location of Sgr A*, $(l, b) = (359.9442°, -0.04616°)$, and not the origin of the Galactic coordinate system $(l, b) = (0, 0)$.

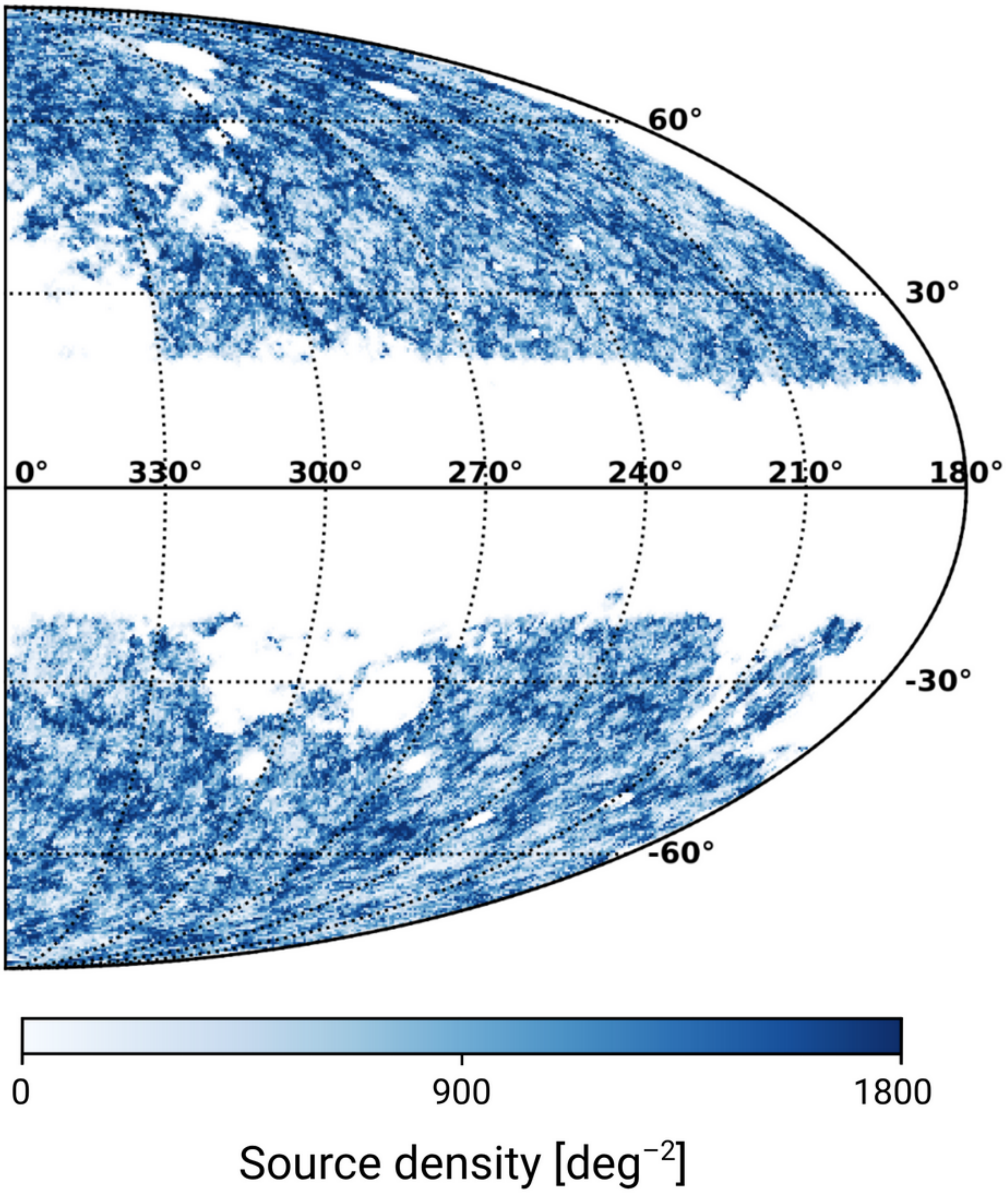


Figure 5.1: Sky map showing the source density of the parent galaxy sample, as defined in Ch. 5.1, in Galactic coordinates and Mollweide projection.

spectroscopic samples (e.g. for photo-zs without *i*-band compared to galaxies from GAMA there is a normalised bias of 0.017 and an outlier fraction of 1.2%, see Ch. 5.1.1 for formal definitions of these statistical metrics). In Chapter 5.1.1, we validate the quality of these photo-z explicitly for the case of our parent sample. For each galaxy in our sample, we assign the best possible redshift (`BEST_Z`) that exists in the following increasing priority order: photo-z without i-band, photo-z with i-band or spec-z.

Large, extended foreground galaxies can cause fragmentation of LS10 photometry. This is an effect where the Legacy Survey Tractor pipeline (Dey et al. 2019) breaks up the galaxy into smaller 'fragments', often due to enhanced HII regions or simply non-uniform galaxy isophote maps not allowing for typical parametric galaxy models to be well-fit without significant residuals. It then assigns an entry for each fragment in the catalogue with its own (erroneous) properties. To limit the effect of fragmentation in our parent galaxy sample, we mask out the D_{25} region around HECATE v2.0 galaxies, using the `HealSparse` Python library. The D_{25} region is defined as the 2-dimensional ellipse fitted to where the B-band brightness profile drops below 25 mag arcsec^{-2} (Kovlakas et al. 2021; Kyritsis et al. 2025). The size of this ellipse is encoded via the minor and major axes, along with the positional angle showing the projected orientation of each galaxy on the sky. This mask removes $< 1\%$ of the LS10 galaxies selected with criteria $1-9$ above. Similarly, the masking of sources associated with *Siena Galaxy Atlas* (SGA; Moustakas et al. 2023) galaxies, often suffering from fragmented optical data, aims to remove the matches to bright off-nuclear X-ray sources. These are outside the scope of this Chapter, which focuses exclusively on nuclear X-ray emission in low-mass galaxies (but see Ch. 5.3.3 for more details).

In Chapter 5.3.1 we describe how we compute X-ray fluxes via aperture photometry extracted at the locations of the optical galaxies. Given the relatively large PSF of eROSITA, nearby bright X-ray point-like and/or diffuse, extended sources, such as clusters of galaxies, can leak X-ray flux into our target apertures and cause biases. Therefore, we also create a `HealSparse` map of the R_{500} region around the 12,247 eRASS1 clusters (Bulbul et al. 2024). R_{500} defines the radius at which the local density equals 500 times critical density of the Universe and for clusters at low redshift $z < 0.2$ it ranges from $\sim$ 500 kpc to 1 Mpc. For clusters with no R_{500} information (due to bad X-ray surface brightness fits or mass determination) we set a value of 500 kpc. We mask all 12,247 clusters regardless of their redshift as even background sources ($z > 0.2$) could bias the X-ray aperture photometry. In addition, we also remove galaxies that lie in spurious X-ray regions based on eRASS1 source over-density analysis described in Section 5.2. of Merloni et al. (2024), including supernova remnants and PSF wings of X-ray very bright point sources. Overall, these effects flag a non-negligible $\sim$10% of the parent galaxy sample and are visible in Figure 5.1 (white ellipse mask regions with zero sources).

5.1.1 Extragalactic redshift compilation and validation of photometric redshifts

Knowledge of a galaxy's redshift is essential for more in-depth studies of galaxy evolution, as it allows for the accurate determination of the source luminosity and numerous host galaxy parameters, such as stellar mass or star formation rate. Therefore, we curate an extragalactic ($z > 0.002$) spectroscopic redshift compilation, aiming towards high completeness, including most of the largest catalogues in the literature[4], along with the latest redshifts from SDSS-V (Kollmeier et al. 2025) and 1229 unpublished redshifts from Balzer et al. (in prep.) using VIRUS on the Hobby-Eberly Telescope (Ramsey et al. 1998; Hill et al. 2021). Our compilation includes galaxies, AGN and QSOs that have good quality redshifts, as defined by the cuts described in Table 5.1. It builds on the compilation of galaxy-only redshifts assembled for the work by Kluge et al. (2024) (see their Appendix D and Table D.1 for further information and all the references). The quality cuts shown in Table 5.1 are especially important for AGN and QSO spec-zs as standard pipelines attribute erroneous redshifts to ∼ 5% of such sources (e.g. Aydar et al. 2025), whereas a higher accuracy is achieved for galaxy dominated spectra, due to precise absorption features (e.g. Calcium II H and K at 3969.59Å and 3934.78Å, respectively).

We first compile all available spec-zs and then create a unique table, listing only the 'best' spec-z entry per source, which is subsequently used to update the parent sample as explained above. The flagging algorithm to find the 'best' spec-z entry per source is as follows:

1. Individually sort large catalogues containing repeated target observations by S/N (e.g. SDSS, LAMOST) and take only the highest quality redshift. This is done with an internal best match with radius $2''$ to account for the fibre size of such surveys.
2. Compile all catalogues together to form the complete extragalactic spectroscopic redshift catalogue, containing duplicates.
3. Make an internal match using sky coordinates within a radius of $1''$ and sort duplicate sources into groups with a unique `GroupID` and `GroupSize`.
4. Compute the range in redshift (max-min) per group identified with `GroupID`.
5. Flag groups with large redshift ranges > 0.01: `z_discrepant_range_0pt01=1` (otherwise `0`).
6. If `GroupSize==2`
   (a) And group contains a spec-z from DESI DR1 (whether group redshifts are discrepant or not, i.e. `z_discrepant_range_0pt01=1` or `0`), keep DESI DR1 redshift.

---

[4] The public version of the extragalactic spectroscopic redshift catalogue developed here is available upon request.

(b) If not (a) and group contains a spec-z from SDSS in the priority order of latest release, that is SDSS-V, then IV, then III and earlier (whether group redshifts are discrepant or not), keep SDSS redshift.

(c) If not (a) nor (b) and `z_discrepant_range_0pt01=0`, keep first redshift of the group. Otherwise, if `z_discrepant_range_0pt01=1`, discard all redshift in the group from the unique compilation.

7. If `GroupSize>2` and `z_discrepant_range_0pt01=0`, keep entries as 6(a)-(c) above.
8. If `GroupSize>2` and `z_discrepant_range_0pt01=1`, do 'majority agreement' flagging: for each entry in a given `GroupID`, check how many other entries have $|\Delta z| < 0.01$ (i.e. are in agreement). Find the subset with the maximum number of redshifts in agreement and keep first entry. Mark the other sources not in the agreeing subset with `z_outlier=1`. If no majority agreement is found, all sources in the group are flagged as outliers and are discarded from the unique compilation.

The flagging algorithm results in 20,527,574 unique spec-z sources, with the top 20 contributors, as well as additional AGN and QSO catalogues added to the compilation by Kluge et al. (2024), described in Table 5.1. Note that this number refers to the number of unique sources over the all-sky (not eROSITA-DE footprint), as shown by the sky plot in Fig. 5.2. Additionally, the range of redshifts covered by this compilation is shown in Figure 5.3.

We append LS10 catalogue columns to the sources within the LS10 footprint and apply the same selection criteria as described in Ch. 5.1 to the compilation, except for the cut on the redshift. This results in around 3.57 million unique sources which can then be used to improve upon the photo-zs in the parent galaxy sample. Figure 5.4 shows the comparison of the photo-zs versus spec-zs, using now the same redshift cut as the parent sample, $0.03 < z \leq 0.2$, leaving around 1.03 million sources. Following the standard statistical metrics to assess the quality of the photo-zs (Ilbert et al. 2006), we find a bias of 0.0084 (defined as the mean of the normalised residuals: $\langle \Delta z \rangle = \frac{(z_{\rm spec} - z_{\rm phot})}{(1+z_{\rm spec})}$), a standard deviation from the normalised median absolute deviation of 0.015 (defined as $\sigma_{\rm NMAD} = 1.4826 \times {\rm Median} \frac{|z_{\rm spec} - z_{\rm phot}|}{(1+z_{\rm spec})}$) and a fraction of outliers of 0.5% (defined as $\eta = \frac{|z_{\rm spec} - z_{\rm phot}|}{(1+z_{\rm spec})} > 0.15$). The fraction of catastrophic outliers is $< 0.03\%$, defined arbitrarily as $\frac{|z_{\rm spec} - z_{\rm phot}|}{(1+z_{\rm spec})} > 1$. These excellent metrics thereby validate the use of the photo-zs for the rest of this Chapter.

Using this well-defined sample of 1.03 million sources with the same selection as the parent galaxies, we also quantify the potential systematic effect of sources scattering in and out of our redshift range. We find that 8.3% of sources with known spec-z between $0.03 < z \leq 0.2$ would be scattered out of the sample because the $z_{\rm phot} > 0.2$, and 2.3% with known $z_{\rm spec} > 0.2$ would be scattered in because their $0.03 < z_{\rm phot} \leq 0.2$. This will be improved with future spectroscopic surveys such as Legacy Survey of Space and Time (LSST) by the Vera C. Rubin Observatory (Ivezić et al. 2019), the 4-metre Multi-Object Spectroscopic Telescope (4MOST; de Jong et al. 2019) and future SDSS releases.

#	Survey	Reference	Number[a]	Quality Flag
1	DESI DR1	DESI Collaboration et al. (2025b)	15679799	`Z > 0.002 & ZWARN == 0 & COADD_FIBERSTATUS == 0`
2	SDSS-IV DR17	Abdurro'uf et al. (2022)	1492476	`Z > 0.002 & ZWARNING == 0 & SN_MEDIAN_ALL > 2 & SPECPRIMARY == 1`
3	Quaia (G<20.5)	Storey-Fisher et al. (2024)	693386	`redshift_quaia > 0.002`
4	SDSS-IV DR16	Ahumada et al. (2020)	681464	`zwarning == 0 & z > z_err & z_err > 0`
5	SDSS-V DR20	(priv. comm.)	212918	`MJD <= 60708 & Z > 0.002 & ZWARNING == 0 & SN_MEDIAN_ALL > 2 & SPECPRIMARY == 1`
6	Compilation[b]	Zou et al. (2019)	181775	`z > 0.002`
7	SDSS-III DR11-12	Alam et al. (2015)	177324	`z > 0.002 & zWarning == 0`
8	PRIMUS	Coil et al. (2011); Cool et al. (2013)	160862	`z > 0.002 & ZQUALITY >= 3` and not a star
9	2dF Galaxy Redshift Survey	Colless et al. (2001)	159491	`z > 0.002 & q_z >= 3`
10	3D-HST Survey[c]	Momcheva et al. (2016)	153681	`z > 0.002 & in_DESComplilation & des_flags >= 3`
11	WiggleZ Dark Energy Survey	Drinkwater et al. (2018)	125702	`z > 0.002 & q_z >= 3`
12	6dFGS DR3	Jones et al. (2009)	93866	`z > 0.002 & q_z == 4`
13	HETVIPS	Zeimann et al. (2024)	68118	`z > 0.002 & classification == GALAXY`
14	Milliquas v8	Flesch (2023)	60975	`Z > 0.002` (not photo-z, i.e. 1 decimal place redshift)
15	LAMOST DR10	(see caption[d])	54422	`z > 0.002 & z_err < z & z_err > 0`
16	GAMA DR4	Driver et al. (2022)	49977	`Z > 0.002 & NQ > 2`
17	VERONCAT (corrected)	Véron-Cetty & Véron (2010); Flesch (2013)	47565	`z > 0.002`
18	2dFLenS	Blake et al. (2016)	42756	`z > 0.002 & in_DESComplilation & des_flags >= 3`
19	VIPERS DR1	Garilli et al. (2014)	41775	2.6 < flag < 4.6; also 22.6–24.6 (secondary), 12.6–14.6 (AGN), 212.6–214.6 (AGN, secondary)
20	Las Campanas Redshift Survey	Shectman et al. (1996)	21907	`z > 0.002`
	⋮	⋮	⋮	⋮

Table 5.1: Top 20 contributors to the extragalactic redshift compilation used for this Chapter, with columns for the survey name, literature reference, number of sources and quality flag used for selection. Entries appearing after the triple dots are additional AGN and QSO catalogues added in this work that are not documented in Kluge et al. (2024). (Table and caption continued on next page)

#	Survey	Reference	Number^a	Quality Flag
	⋮	⋮	⋮	⋮
	SDSS-V DR19	(priv. comm.)	10810	`Z > 0.002 & ZWARNING == 0 & SN_MEDIAN_ALL > 2`
	VVDS DR2	Le Fèvre et al. (2013)	10584	`z>0.002 & f_z>1 & z!=9.99`
	OzDES DR2	Lidman et al. (2020)	10276	`z>0.002 & qop>2 & qop!=6`
	AGES	Kochanek et al. (2012)	8704	`z>0.002`
	COSMOS z-compilation	Khostovan et al. (2025)	8552	`specz>0.002 & Confidence_level>=80`
	2SLAQ	Cannon et al. (2006); Croom et al. (2009)	3207	`z>0.002 & q_z>2, z2S>0.002 & q_z2S==1`
	2Qz	Croom et al. (2004)	411	`z1>0.002 & (q_z1==11 or q_z1==21)`
	VANDELS DR4	Garilli et al. (2021)	381	`zsp>0.002 & q_zsp!=0 & q_zsp!=1 & q_zsp!=10 & q_zsp!=11 & q_zsp!=20 & q_zsp!=21 & q_zsp!=210 & q_zsp!=211 & q_zsp!=220`
	SDSS-IV DR16Q	Lyke et al. (2020)	227	`Z>0.002 & ZWARNING==0 & SN_MEDIAN_ALL>2`
	VUDS DR1	Tasca et al. (2018)	204	`zspec>0.002 & zflags>1 & zflags!=11 & zflags!=21 & zflags!=31 & zflags!=41 & zflags!=32 & zflags!=33 & zflags!=34`
	SDSS-V DR18 eFEDS	Aydar et al. (2025)	84	`Z > 0.002 & ZWARNING == 0 & SN_MEDIAN_ALL > 2 & SPECPRIMARY == 1`
	QUBRICS	Boutsia et al. (2020)	14	`zspec>0.002`
	z>5.6 QSO compilation	Fan et al. (2023)	406	

Table 5.2: (continued from previous page) *a*: Number of sources refers to the unique entries that made it into the final compilation after the flagging algorithm (see text). *b*: Note that the Zou et al. (2019) entries come from a compilation of several spectroscopic surveys: 2dFGRS (Colless et al. 2001), 2SLAQ (Cannon et al. 2006), 6dFGS (Jones et al. 2004, 2009), CFRS (Lilly et al. 1995), CNOC2 (Yee et al. 2000), DEEP2 (Davis et al. 2003; Newman et al. 2013), SDSS DR14 (Abolfathi et al. 2018), VIPERS (Garilli et al. 2014; Guzzo et al. 2014), VVDS (Le Fèvre et al. 2005; Garilli et al. 2008), WiggleZ (Drinkwater et al. 2010; Parkinson et al. 2012), and zCOSMOS (Lilly et al. 2007). *c*: The cuts `in_DESComplilation & des_flags >= 3` refer to good-photometric quality sources present in the compilation from Gschwend et al. (2018). *d*: Documentation for the LAMOST DR10 catalogue can be found here: `https://www.lamost.org/dr10/v2.0/doc/lr-data-production-description`.

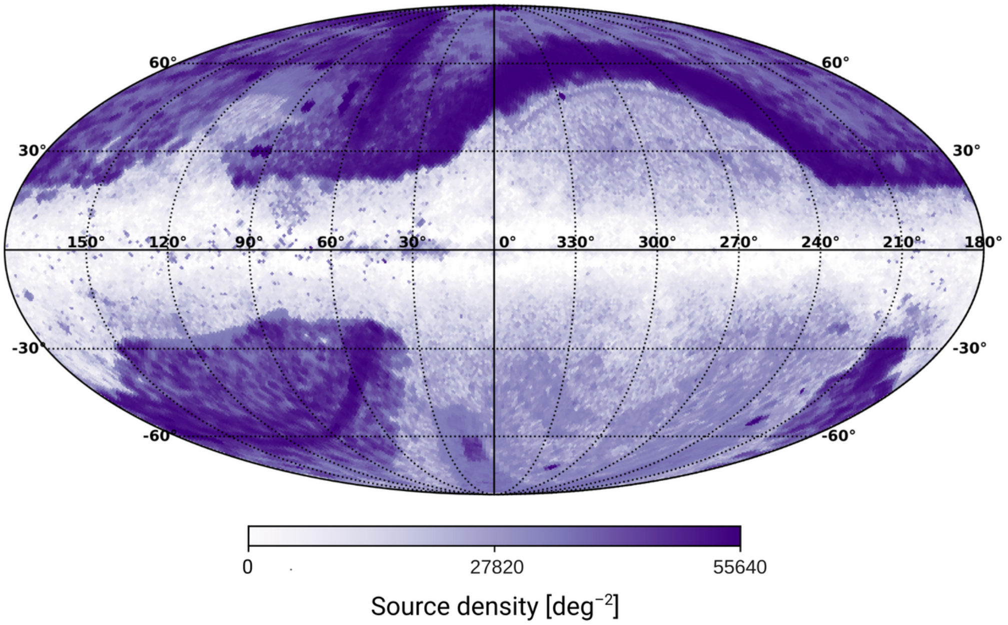


Figure 5.2: Sky map showing the source density of the unique extragalactic redshift compilation in Galactic coordinates and Mollweide projection.

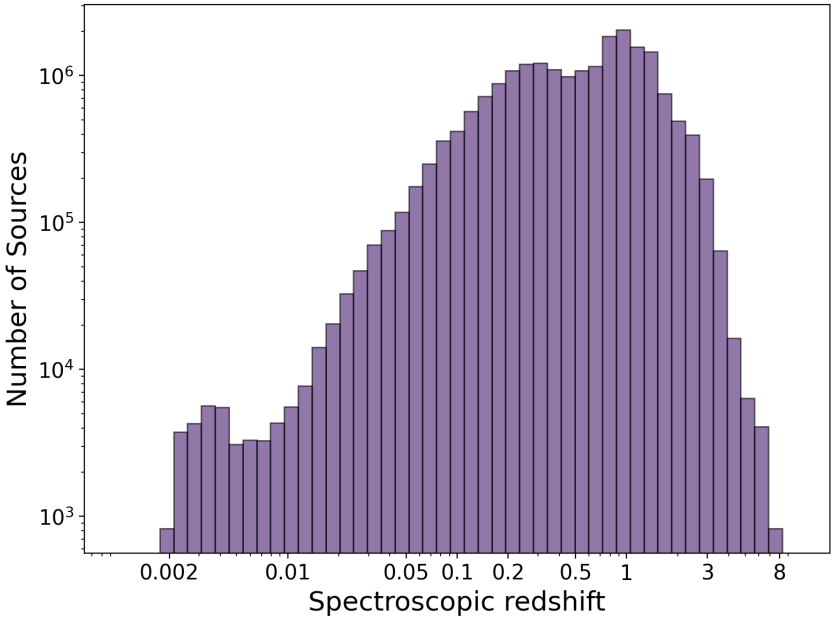


Figure 5.3: Histogram showing the redshift distribution of the unique extragalactic redshift compilation.

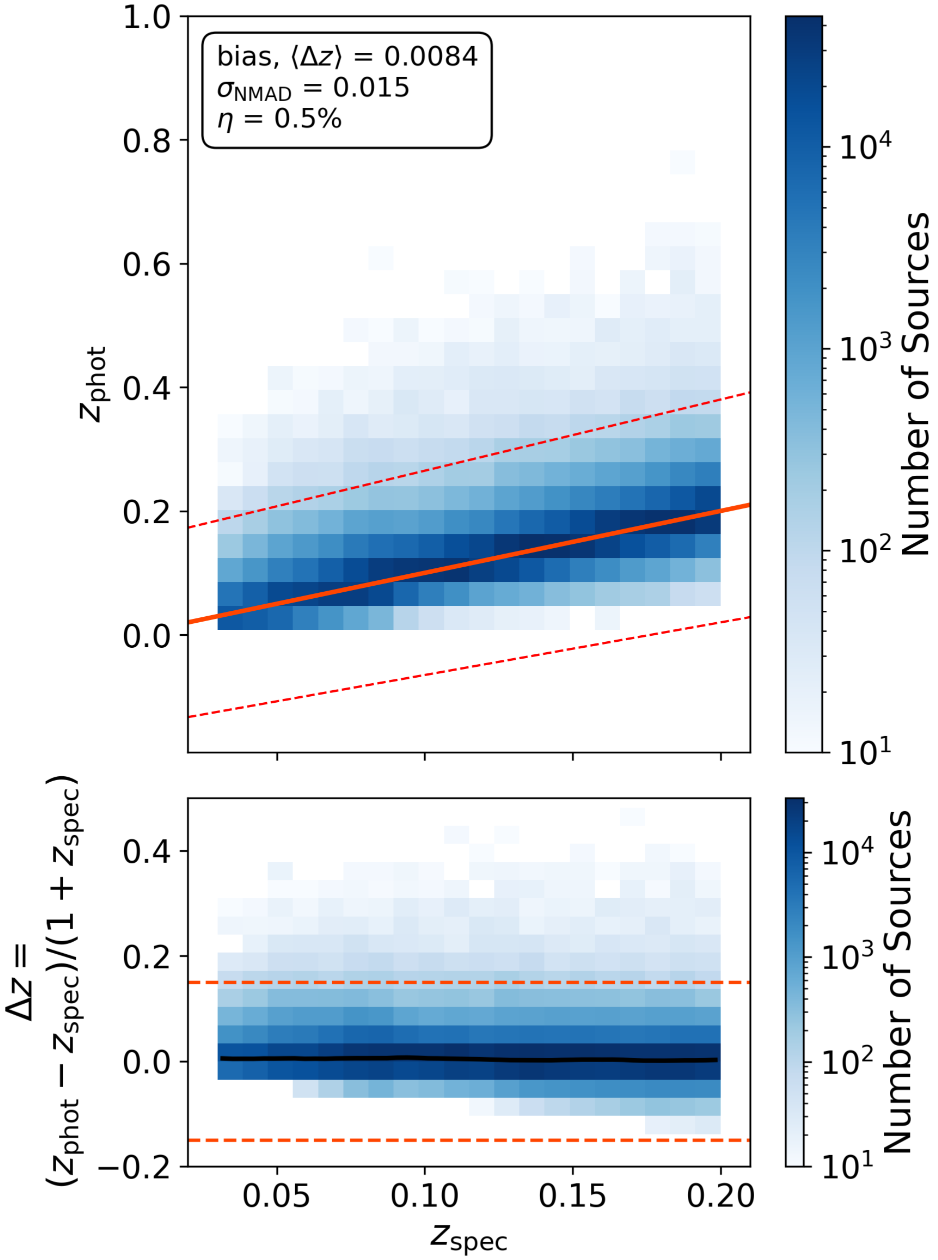


Figure 5.4: Validation of the subset of photometric redshifts from the parent sample of galaxies with spectroscopic redshift from the extragalactic redshift compilation. Minimal outlier fraction (η) of 0.5% and almost negligible bias of 0.0084 show the excellent quality of the photo-zs. The solid red line marks the 1:1 relation and the dashed red lines are used to define η (see text for details). Darker blue colours indicate higher number of sources, as shown by the colour bar.

5.2 Calculating host galaxy properties: stellar masses and star formation rates

Using the accurate photo-zs, we compute host galaxy properties using the Photometric Analysis for Redshift Estimation code (LePHARE; Arnouts et al. 1999; Ilbert et al. 2006, 2009). This is a template-fitting code that fits the observed spectral energy distribution (SED), which in our case is composed of six photometric bands (when available) *g, r, i, z, W1, W2*, with stellar population synthesis models. It works by redshifting synthetic galaxy (and/or AGN, stellar) SEDs across the desired redshift range, applying dust attenuation (and optionally emission lines) and convolving with the filter response functions to simulate observed magnitudes in the required photometric bands, effectively building a large grid of model magnitudes. Then, the observed and modelled magnitudes are compared through χ^2-minimisation to find the best fit and associated physical parameters. Note that since the grid of magnitudes is pre-computed, the computing time per source is significantly reduced (compared to other SED fitting methods, see Ch. 5.2.1).

We use the Bruzual & Charlot (2003) stellar evolution models, taking a Chabrier (2003) initial mass function (IMF) and Calzetti et al. (2000) dust extinction curves to introduce a reddening through varied $E_{(\mathrm{B-V})}$. The modelled parameter grid includes two metallicity values ($Z = 0.008, 0.02$), one sub-solar and one solar; two star formation histories, exponential decline [$\mathrm{SFR}(t) \propto \exp(-t/\tau)$] and delayed exponential [$\mathrm{SFR}(t) \propto t \times \exp(-t/\tau)$], with star formation timescales (τ) equal to 0.1, 0.3, 1, 3, 5, 30 Gyr and 1, 3 Gyr, respectively. Importantly, we build galaxy-only SED models, which do not include models to explain the mid-infrared (MIR) emission from galactic and nuclear dust heated by star-formation or the AGN (e.g. Mullaney et al. 2011; Mor & Netzer 2012; Dale et al. 2014; Lyu et al. 2017) nor the big blue bump of the AGN accretion disk (e.g. Richards et al. 2006) (see Ch. 5.2.1 for how we treat sources that may require these components). We choose to build the model grid in flux and not magnitude space, such that even sources with bad flux estimates (e.g. negative entries), but large errors, can provide an informative constraint on the spectrum. The *k*-correction for each input photometric band (used in Ch. 5.2.2 to calculate absolute magnitudes) is derived from the apparent magnitude at the nearest redshifted band. Lastly, we do not apply emission line templates for the SED fitting with LePHARE, but do account for this in later re-fits with a more sophisticated Bayesian SED code (see Ch. 5.2.1) as their inclusion can be important in deriving unbiased physical galaxy parameters (e.g. Mobasher et al. 2015; Santini et al. 2015).

In order to validate the stellar masses computed with LePHARE, we match (using optical coordinates) to the GAMA DR4 catalogue (Driver et al. 2022), which provides robust stellar mass estimates derived via SED fitting using extensive photometric coverage from the UV to the far-IR. We find 71,962 sources with $|(\texttt{BEST\_Z} - z_{\mathrm{GAMA}})|/(1 + z_{\mathrm{GAMA}}) < 0.01$ and both LePHARE- and GAMA-computed $\log M_*/M_\odot > 7.5$. The top panel of Figure 5.5 shows the difference in stellar masses derived from GAMA and from LePHARE, as a function of GAMA stellar mass. The grey solid line marks the running median (offset) with the shaded region indicating the standard deviation of the difference in mass. It is

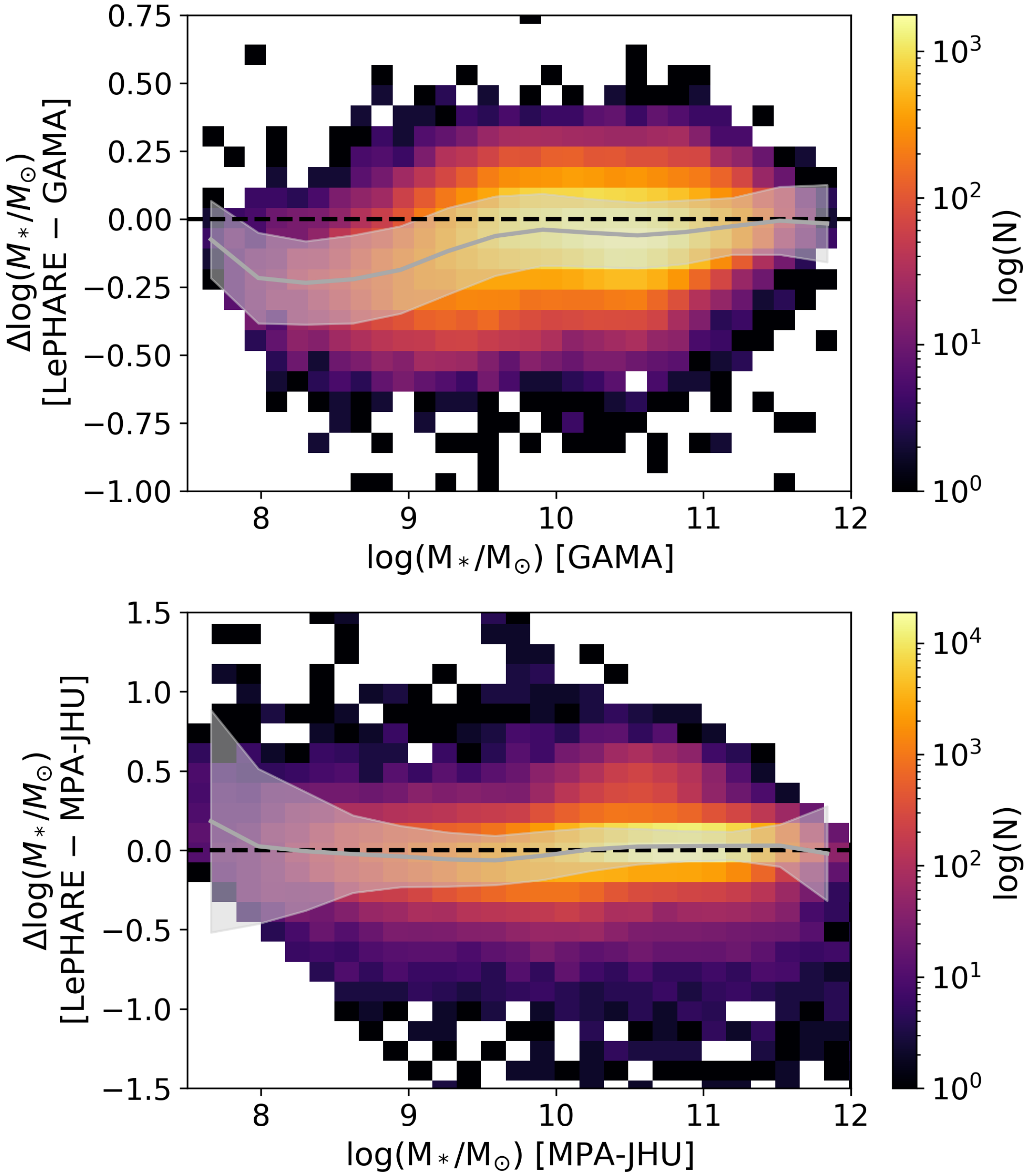


Figure 5.5: Difference in stellar mass as a function of stellar mass between measurements from the GAMA and MPA-JHU surveys and those computed with LePHARE for sources in common from our parent sample (with redshifts in agreement). The grey solid line marks the running median and the shaded region shows the standard deviation of the difference in mass.

clear that down to around $\log M_*/M_\odot \sim 9$ the stellar masses have a small downward offset of ~ -0.05 dex, however, this increases to ~ -0.18 dex between $\log M_*/M_\odot \sim 7.5 - 9$. The overall scatter (average standard deviation of the offset) is 0.14 dex, in agreement with the findings of Zou et al. (2019). Similarly, on the bottom panel of Figure 5.5, we do the comparison using the MPA-JHU DR8 catalogue (Brinchmann et al. 2004; Kauffmann et al. 2003; Tremonti et al. 2004), with 258,516 matched galaxies. We find a small positive offset of ~ 0.005 dex and scatter ~ 0.24 across the mass range. Overall, the stellar masses computed with LePHARE are (closely) consistent with previous work, albeit with large scatter.

Furthermore, we repeat the analysis to compare the LePHARE-derived star formation rates (SFRs) to GAMA and MPA-JHU. We conclude that the SFR computed with only six photometric filters are highly offset and scattered, so we do not use them further in this study (see Sect. 5.3.3 for how we set a conservative upper limit on the global star formation properties of the sample using the star-formation main sequence; Speagle et al. 2014).

5.2.1 Stellar mass estimates including an AGN and mid-IR emission component with `GRAHSP` and XGBClassifier

In order to further reduce the stellar mass offset and scatter of our final sample, we build a machine-learning-based classifier to identify sources with potential 'problematic' stellar mass estimation from LePHARE, due to the lacking MIR emission and AGN models mentioned above. For those problematic cases, we then recompute their physical galaxy properties with a more sophisticated fully-Bayesian SED fitting code called `GRAHSP` (Buchner et al. 2024). We note that computing stellar masses using `GRAHSP` for the whole parent sample of over five million sources is computationally unfeasible and largely unwarranted, given that AGN are relatively rare events and Figure 5.5 already shows low offset and scatter across the mass range (see also Aird et al. 2018, who adopted a similar approach).

We choose to use XGBClassifier (Chen & Guestrin 2016), which is a gradient-boosted decision tree algorithm, meaning that it builds an ensemble of decision trees sequentially, where each new tree is trained to correct the errors (residuals) of the previous ones using gradient descent on a given loss function. We train our XGBClassifier to assign a low 'mass reliability probability' to galaxies whose LePHARE-derived stellar masses differ from the `GRAHSP`-derived values by more than ±0.4 dex.

As a first step, we need to build a training sample on which we run SED fitting using `GRAHSP`. We do so by random sampling our parent galaxy sample in stellar mass and redshift space. We also note that for a classifier to perform well, it needs to learn from enough variety, meaning that there need to be enough outliers. Previous studies (e.g. Buchner et al. 2024, and references therein) have shown that when an AGN is present, but not included in the SED fitting, the stellar mass tends to be overestimated, since the multi-wavelength emission of the AGN is incorrectly attributed to the stellar component. An indication of an AGN being present can be given by nuclear X-ray emission, but several cleaning and validation procedures must be undertaken before attributing any X-ray detection by eROSITA to an AGN, as well as properly assigning it to its most likely host galaxy counterpart (a step especially difficult for the as-of-yet under-explored low-mass galaxy regime). This is why we have to make an iterative step in our analysis to first determine real X-ray emitting parent sample sources, as well as the correct associations to their host galaxy counterparts, such that we can feed our XGBClassifier with a well-understood sample of (potential) outliers. In Ch. 5.3.2, we describe these cleaning and validation steps that we perform on the parent sample galaxies. Using only their LePHARE-derived stellar masses, we find 892 secure X-ray detected sources among the low-mass galaxies ($\log M_{*,\mathrm{LePHARE}}/M_{\odot} \leq 10$), which are all added to the training sample in order to boost the potential outlier class. Then we sample (up to) 50 random galaxies in each $M_* - z$ bin for both the high-mass ($\log M_{*,\mathrm{LePHARE}}/M_{\odot} > 10$) X-ray detected galaxies (totalling to 583) and the full non-X-ray detected parent galaxies (totalling to 1552). Overall, our training sample is made up of 3,027 galaxies sampled over the entire $M_* - z$ distribution to ensure coverage of the entire parameter space. We note that the final sample numbers presented in Chapter 5.3.2 refer to the second iteration of these cleaning and validation procedures, using the LePHARE- and GRAHSP-derived stellar masses, as an outcome of the analysis in this Chapter.

We proceed to recompute the stellar masses of these 3,027 galaxies with `GRAHSP`. We use the same SED set-up as for the LePHARE runs, except for the following key differences (for which full details are given in Buchner et al. 2024). We account for continuum and line emission from ionised gas with the `nebular` module (Boquien et al. 2013, 2019), which contribute increasingly for galaxies with recently formed stars. We use the `biattenuation` module to account for dust attenuation from a Small Magellanic Cloud (SMC) attenuation curve (Prevot et al. 1984), which shows a steep rise with λ^{-1} with no strong 2175Å feature, indicating smaller dust grains compared to the galactic ISM average, and is parametrised using `E(B-V)`. As this attenuated optical light is then reprocessed and re-emitted (conserving energy) at IR wavelengths, we model this using the Dale et al. (2014) templates (`galdale2014`). Within the `biattenuation` module we also separately account for the effect of attenuation on the AGN models, as they are affected by both the galactic and nuclear attenuation. Importantly, we include several AGN components, including a big blue bump at optical/UV wavelengths using an empirically motivated smooth bending power-law parametrisation (`activatepl`), broad and narrow emission lines (`activatelines`) and torus emission associated with the reprocessed optical/UV emission by hot and cold dust emitting in the NIR and MIR, respectively (`activatetorus`). Photometric redshift errors (with or without i-band, depending on availability, see Ch. 5.1.1) are also ingested as priors in the SED fitting. `GRAHSP` efficiently samples this complex parameter space with a nested sampling Monte Carlo algorithm called *Ultranest* (Buchner 2021d), which additionally allows for finer sampling than traditional grid-based SED-fitting codes. The same *g,r,i,z,W1,W2* photometry, derived from the best-fit model fluxes in the LS10 catalogue, is used as in the LePHARE SED fitting, for consistency.

The results are shown in Figure 5.6. There are 344 sources with $|\Delta M_*| > 0.4$ dex (11%), confirming the conclusions of Ch. 5.2 that for most cases the simpler and faster SED fitting method with LePHARE produces unbiased results. However, it is also clear that the stellar

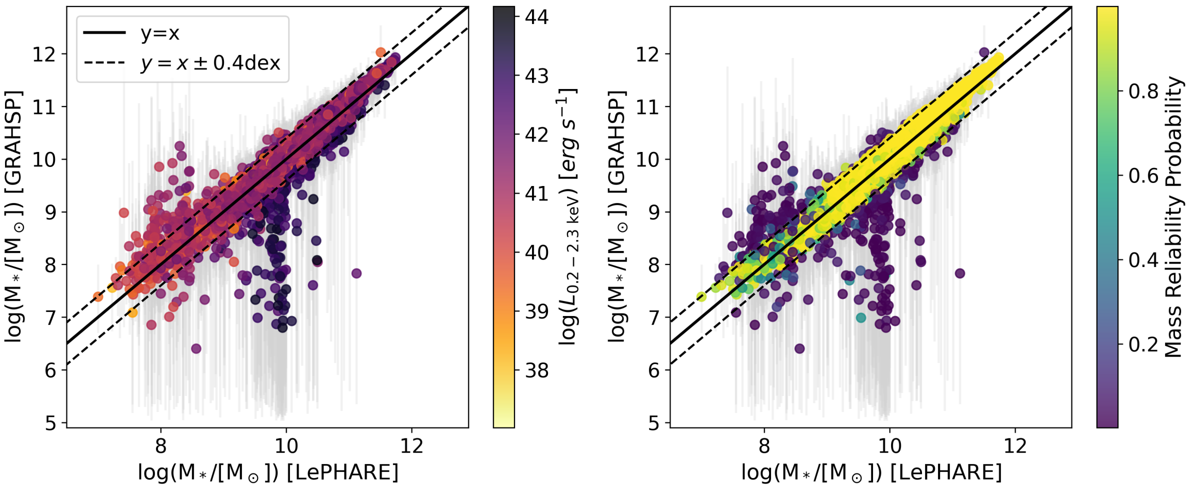


Figure 5.6: A comparison of the median stellar mass derived via LePHARE (x-axis) and `GRAHSP` (y-axis) SED fitting for the training sample of 3027 galaxies with solid and dashed lines indicating $y = x$ and $y = x \pm 0.4$ dex. The left and right panel are colour-coded as a function of observed $0.2 - 2.3$ keV X-ray luminosity and mass reliability probability (as computed by the XGBClassifier), respectively. 1σ uncertainties on LePHARE-derived stellar masses are often too small to be seen, whereas the 2σ uncertainties on GRAHSP-derived stellar masses are clearly visible (see text for more details on the handling of such large confidence intervals).

mass can be catastrophically wrong and off by several orders of magnitude for the outlier cases. Interestingly, the majority of outliers below the 1:1 line are highly X-ray luminous in the $0.2 - 2.3$ keV band and have overestimated stellar mass from LePHARE, as predicted above. Example SEDs derived from `GRAHSP` and LePHARE are shown in Figure 5.7, where the LePHARE model clearly fails to reproduce the high MIR emission coming from the AGN torus. However, being X-ray luminous does not necessarily mean that the stellar mass will be over-estimated, as shown by the numerous such sources lying within the scatter of the 1:1 line. It is important to correct the stellar masses of these X-ray luminous outlier objects as they fall exactly in our statistical sample (see Ch. 5.2.2) and boost the incidence of X-ray AGN in low-mass galaxies. At lower stellar masses, $\log M_{*,\ \mathrm{LePHARE}}/M_\odot < 9$, there are also a cloud of outliers above the 1:1 line, which are sources where `GRAHSP` fits a steeply rising MIR emission from hot galactic dust. As this model was not included in the LePHARE SED fitting run, the stellar masses of these objects are under-estimated. However, this dust-obscured star-forming galaxy fit is also found to be degenerate with a heavily obscured AGN solution with high nuclear $E_{(\mathrm{B-V}),\ \mathrm{AGN}}$ attenuation and high AGN luminosity, as traced by 5100Å emission. The overall impact on the stellar mass estimation of this degeneracy is non-trivial as stellar age is a confounding factor: a galaxy could have a lot of dust-obscured star formation now, but it might be young overall and not yet have built up much stellar mass. Conversely, a galaxy could be old and massive but currently

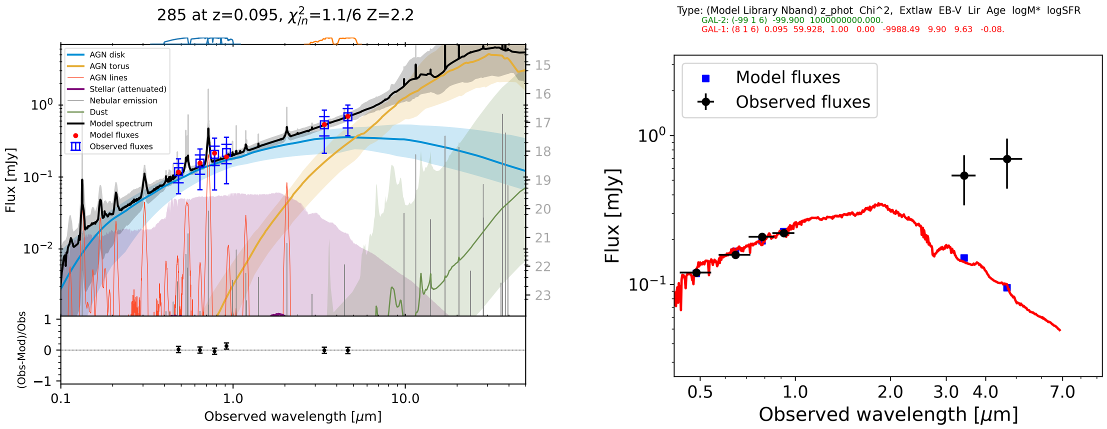


Figure 5.7: Best-fit SED results for a galaxy with strong AGN component computed with `GRAHSP` (left) and LePHARE (right). In comparison to the galaxy-only LePHARE model (red curve), `GRAHSP` finds the best-fit to include a strong AGN disk model (blue curve) and corresponding re-emission in the IR by the AGN torus (yellow curve). As a result, LePHARE derives a $\log M_*/M_\odot \sim 9.5$ whereas `GRAHSP` can only constrain a $2\,\sigma$ upper limit of $\log M_*/M_\odot \lesssim 9.6$.

forming fewer stars or hosting a dust-obscured AGN. We defer in-depth analysis of this issue to future work where we use more photometric bands that can help disentangle such degeneracies.

The advantage of machine-learning-based classifiers is that they can learn from a combination of several different features, referring in this case to the optical photometric properties and LePHARE fit statistics of the training sample galaxies (see below), allowing them to detect outliers with high completeness. In contrast, if we simply used a cut in X-ray luminosity, we would be recomputing stellar masses for sources which do not need it and simultaneously missing many outliers above the 1:1 line, as shown in Figure 5.6. In particular, the advantage of an XGBClassifier, in comparison to Logisitic Regression or Random Forests, is that it can handle non-linear feature interactions naturally, it is more efficient (thanks to the boosting meaning fewer trees are required), it includes regularisation to control over-fitting and handles imbalanced data well (Chen & Guestrin 2016).

The ultimate goal of any classifier is to maximise the precision[5] and recall[6], simultaneously. This is encoded in the F1 Score, the harmonic mean between the two[7]. For our scientific goal, where we define 'positive' detections (`class==1`) as outliers, we must prioritise maximising recall because false negatives (i.e. not identifying an outlier) are costly.

---

[5] Precision is defined as the ratio of the true positives over the sum of the true positives and false positives ($\frac{TP}{TP+FP}$), i.e. of all the instances the model predicted as positive, how many were actually positive?

[6] Recall is defined as the ratio of the true positives over the sum of the true positives and false negatives ($\frac{TP}{TP+FN}$), i.e. of all the actual positive instances, how many did the model correctly identify?

[7] The F1 Score is defined as: F1 Score $= \frac{2}{(1/\text{Precision})+(1/\text{Recall})} = 2 \cdot \frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}}$.

Accordingly, the methodological choices outlined below are designed with this priority in mind. We train the model on the following features, which we found to provide strong discriminatory power (see Fig. 5.8):

1. Colour indicators of the presence of an AGN or MIR dust emission component: *W1*-*W2*, *r*-*W2*, *g*-*r* magnitudes.
2. Stellar mass proxies: absolute *z*-band magnitude (`abs_Mag_z`), galaxy half-light radius for the best fitting galaxy type from LS10 (`SHAPE_R` in units of arcseconds).
3. Fit quality indicators: uncertainty in the redshift (with or without *i*-band and zero for spectroscopic redshifts: `redshift_err`), LePHARE χ^2 fit statistic (`log_LPH_CHI_BEST`).
4. Potential X-ray AGN indicator: $0.2 - 2.3$ keV luminosity (`log_Lx_soft` in units of erg s^{-1}).

For the 77 sources with missing *W2* flux, we replace the *W1*-*W2* and *r*-*W2* feature values with -999, such that the model learns the impact of the missing photometric point in such a key part of the SED, given the above findings. As is commonly done, we split the $3,027$ sources into training (80%) and test (20%) sets, and downsample 30% of the inliers to improve class balance and help the model learn that outliers are not so rare that they will never be assigned. After tuning the hyper-parameters via grid search, we settle on 300 estimators (weak learners), a maximum tree depth of 6, a learning rate of 0.05, a subsample fraction of 0.8, and a column subsampling rate of 0.7. We also apply a class-weighting factor (`scale_pos_weight`: the ratio of inliers to outliers) so that each outlier contributes `scale_pos_weight` times more to the loss function, thereby penalizing false negatives more strongly and increasing recall. Lastly, we choose our evaluation metric (or loss function) to optimize the area under the precision–recall curve (AUCPR). After training, we choose an optimal decision threshold of the mass reliability probability by maximising recall whilst keeping the precision>= 60% via AUCPR. We find the optimal decision threshold to be < 0.38, which we use to flag outlier sources (see Fig. 5.9).

Applying the model on the unseen test sample, using the optimal decision threshold above, 537 sources are classed as inliers and 69 as outliers, in the ratios shown by the confusion matrix in Figure 5.10. Therefore, the overall accuracy[8] of our classifier is (497+59)/606 $\sim$ 92%. It has an area under the Receiver Operating Characteristic Curve (ROC AUC) of 0.937, meaning the model can distinguish inliers from outliers very well, and has high recall for both inliers (93%) and outliers (86%). Meanwhile an AUCPR of 0.672 highlights that, though the model captures outliers well, the precision suffers, meaning that there are some false positives (inliers marked as outlier) included. However, as mentioned above, it is more important to catch all outliers than to potentially spend extra computational time recomputing inliers that have been erroneously marked as outliers. Figure 5.6 shows

---

[8] Accuracy = (Number of correct predictions) / (Total number of predictions) = (TP + TN) / (TP + TN + FP + FN)

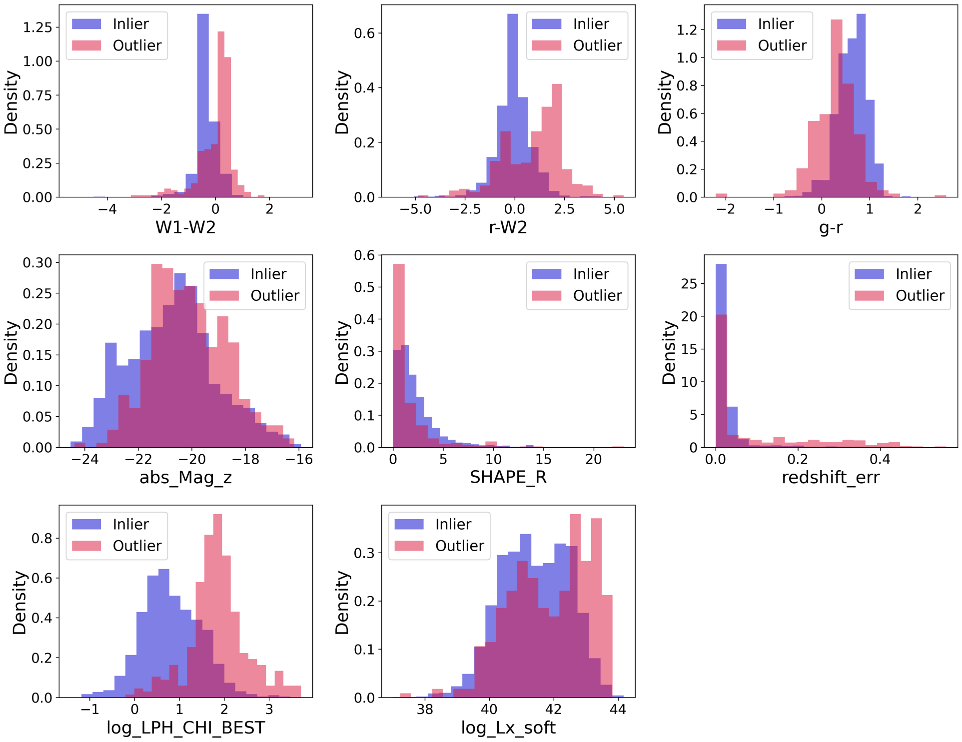


Figure 5.8: Histograms showing the difference in feature-space between the inlier ($|\Delta M_*| < 0.4$ dex) and outlier ($|\Delta M_*| > 0.4$ dex) sources in the training sample for the XGBClassifier (see text for more details of the meaning of the features and their units).

the stellar mass derived by LePHARE and `GRAHSP` colour coded by the probability of being an inlier, in other words the 'mass reliability probability', highlighting the excellent performance of the XGBClassifier. We find that the top three most important features to distinguish outliers are the $g - r$ and $r - W2$ magnitudes, as well as the χ^2 from the LePHARE fit; the $0.2 - 2.3$ keV luminosity is the second-to-last in the ranking.

Finally, we can apply the trained model on the full parent sample of galaxies and find $32,548$ with mass reliability probability < 0.38 to be recomputed with `GRAHSP`. Buchner et al. (2024) extensively validate the galaxy properties obtained by GRAHSP by comparing to a benchmark photometric dataset ('Chimera') where non-AGN pure galaxies are paired with optically-selected pure quasars at the same redshift. They show that the galaxy properties (e.g. M_*, SFR) obtained by GRAHSP on the Chimera sample show no measurable bias, even for galaxies with dominant AGN component or sources with few photometric bands, and very low outlier fraction (defined as when the estimated error bars lie completely outside a 1 dex wide band centred around the true value) of 5%. GRAHSP is also shown to estimate more realistic uncertainties as the fully Bayesian fit includes uncertainties in the model and the data, making the inference highly robust (Buchner et al. 2024). However, one consequence of this is that often the uncertainties on the galaxy properties

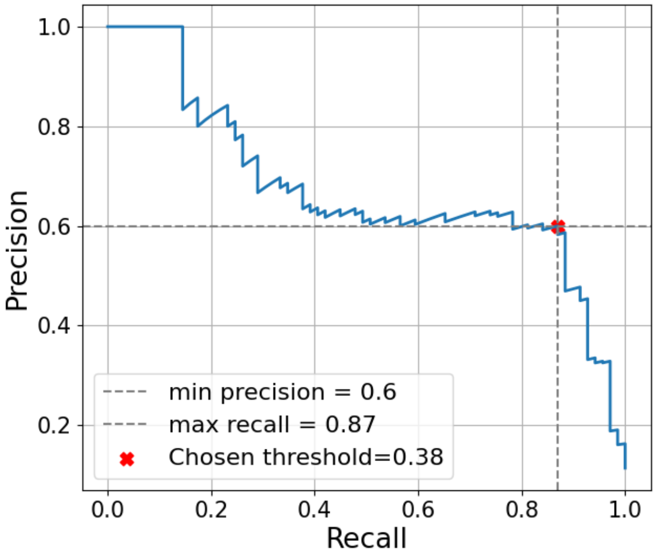


Figure 5.9: Precision-recall curve showing the optimal threshold chosen for selecting outliers from final galaxy sample.

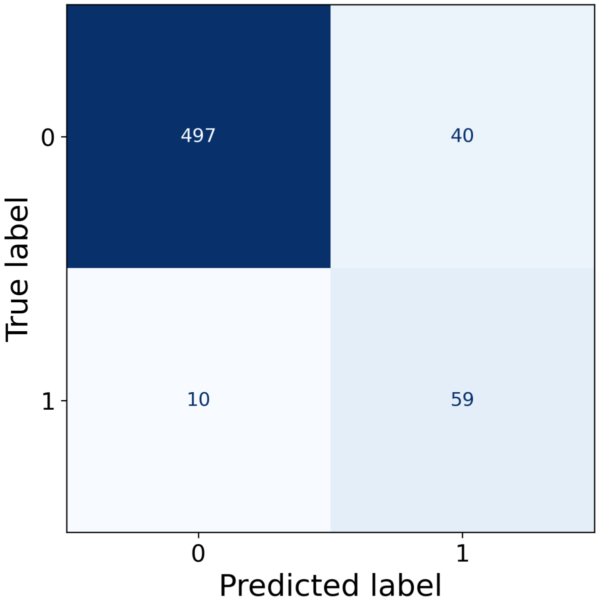


Figure 5.10: Confusion matrix showing the performance of the XGBClassifier on the test sample. The total accuracy of the model is 92% and there is a high recall for both inliers (93%; `class==0`) and outliers (86%; `class==1`).

are very large, especially for sources with few photometric bands, as is the case with our sample. For example, 9,423 out of 32,548 sources have stellar masses that are uncertain by more than 2.5 dex. We deem these sources unconstrained and set them as 2σ upper limits. Importantly, we note that this is a small minority of the overall parent sample consisting of over 5 million galaxies, so there is limited impact on the statistical incidence results presented in this work. Future studies will improve on this aspect by using more photometric bands and the full stellar mass posterior, to better account for the uncertainties in this parameter for all parent sample galaxies.

Overall, for the 32,548 parent sample galaxies deemed to have unreliable stellar masses, we replace the LePHARE-derived median stellar masses with those derived by GRAHSP and use the 2σ upper limit for the subset of sources with unconstrained GRAHSP-derived stellar masses. These are the final stellar masses used for the rest of this Chapter, including deriving the mass completeness limits in Ch. 5.2.2 and the incidence distributions in Ch. 5.4.2.

5.2.2 Stellar mass completeness

Given our magnitude-limited sample, we derive a redshift-dependent stellar mass completeness cut, to ensure that our later results will not be biased by our optical selection. To do so, we first k-correct the z-band magnitudes of our parent galaxy sample to derive an absolute z-band magnitude (M_z) and compute their mass-to-light ratios, M_*/L, where

$$L/L_\odot = 10^{-0.4(M_z - M_{\odot,z})}, \tag{5.1}$$

and the absolute magnitude of the Sun in the z-band is $M_{\odot,z} = 4.50$ (Willmer 2018). Given the variance in M_*/L with mass, whereby more massive galaxies tend to be quiescent, red and less luminous compared to lower mass, star-forming bluer galaxies, we split our sample into low-mass and high-mass galaxies to derive more physically motivated completeness limits.

Taking first the low-mass galaxy subsample, we bin our sample into small redshift bins of $\Delta z = 0.01$ and compute the corresponding absolute z-band magnitude limit and luminosity (Eq. 5.1), given by the survey selection: $z \leq 20$ mag. Then we take the upper 50th percentile of the M_*/L ratios in each Δz to signify the galaxies which are as luminous as, or less luminous than, the average for that redshift range. Multiplying this limiting luminosity and M_*/L ratio together and taking the upper 70th or 90th percentile gives the limiting mass, at 70% and 90% level, respectively, that a galaxy can have and still enter into the sample selection. The same process is repeated to derive the mass-completeness limit for the high-mass galaxy subset. Similar methods have been employed by, e.g., Pozzetti et al. (2010); Moustakas et al. (2013); Mountrichas et al. (2022); Guetzoyan et al. (2025).

Figure 5.11 shows the distribution in stellar mass and redshift of the parent galaxy sample (background blue shaded density grid) and the different mass-completeness curves for the low- and high-mass galaxy subsample. Since the low-mass galaxy mass-completeness function is rather steep compared to the change in stellar mass value, we adopt the 70%

limit (light green solid curve) to maximise source statistics. In comparison, all high-mass galaxies lie above the 90% mass-completeness curve (yellow dashed) and so are considered complete in the redshift range probed here. The X-ray detected low-mass (light red) and high-mass (grey) galaxies are also shown in Figure 5.11 and will be described in the following section. The sources with unconstrained GRAHSP-derived stellar masses are plotted with downward arrows at the upper 2σ confidence level.

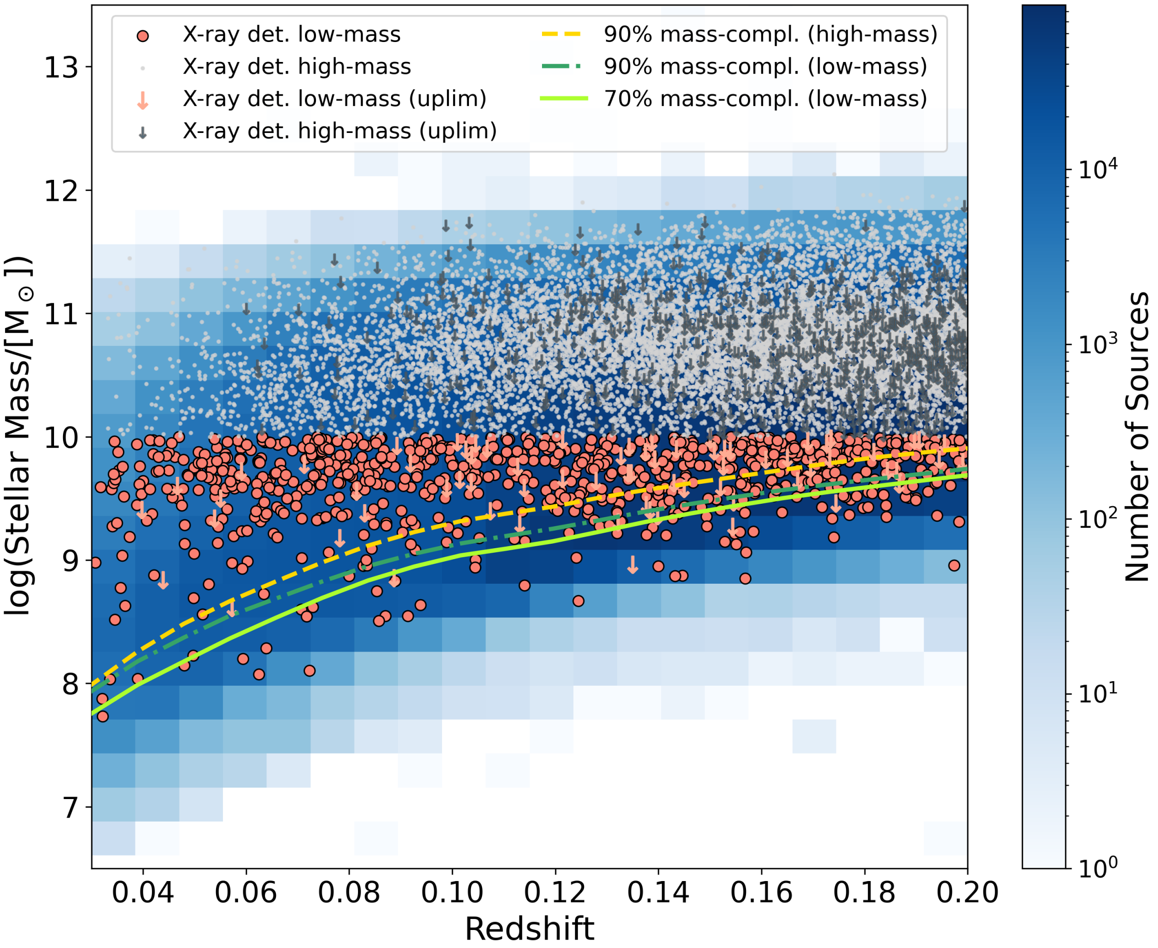


Figure 5.11: Stellar mass versus redshift of the parent sample of LS10 galaxies (background blue shaded density grid, with darker colours indicating higher number of sources as in the colourbar), overlaid with the X-ray detected low-mass galaxies (light red circles) and high-mass galaxies (small gray dots). 2σ upper limits in stellar mass are indicated with downward arrows. The 90% mass-completeness curve for the high- and low-mass galaxies is plotted in yellow (dashed) and dark green (dot-dashed), respectively. The 70% mass-completeness curve, used for computing the incidence in this work, is plotted with a solid light green curve.

5.3 eROSITA X-ray detected sample

We are interested in finding X-ray detected galaxies among our parent sample, so in this section we discuss the homogeneous determination of X-ray fluxes using aperture photometry centred on the location of the optical parent galaxy sample coordinates. This differs from past work dealing with low-mass galaxies which has typically cross-matched pre-existing optical and X-ray catalogues to find X-ray detected sources, after consideration of chance associations and spurious detections (e.g. Latimer et al. 2021; Birchall et al. 2020; Sacchi et al. 2024; Bykov et al. 2024). Although the methods are different, the problems are similar and we elaborate here on the extensive cleaning and validation procedures we took to make sure that: (i) the X-ray detection is real; (ii) the optical host is the statistically favoured counterpart; (iii) the X-ray flux is associated with the central AGN and not with other galaxy processes.

5.3.1 X-ray aperture photometry using `apetool`

We perform aperture photometry using the `apetool` (`v1.28 eSASSusers_240410.0.4`) task from the eROSITA Science Analysis Software System (eSASS; Brunner et al. 2022) centred on the optical coordinates of our galaxy sample to compute eRASS:4 X-ray counts in the main eROSITA band of $0.2-2.3$ keV. `apetool` computes the total counts (N) within the specified aperture, chosen here to be the 75% encircled energy fraction (EEF; i.e. the radius of the PSF at which 75% of the energy is contained), corresponding approximately to a radius of $\sim 30''$, for the survey-averaged PSF.

The background counts (C_{B}) are computed from the source-subtracted background map. More precisely, one of the outputs of the `ERMLDET` eSASS task is the source map, which encodes the results of the point-spread-function (PSF) fitting to determine source centroids and parameters, including multi-PSF fits to help with de-blending of nearby X-ray sources and optional convolution with an extent model (e.g. Gaussian or beta-model) for extended sources (Brunner et al. 2022). These source models can then be subtracted from the source map allowing `apetool` to estimate the source-free background level only. We choose a radius corresponding to the 37.5% EEF within which source subtraction is performed. This corresponds to roughly three times the 50th percentile of the eRASS:4 positional error (i.e. $\sim 10''$; shown in Fig. 5.12 with the yellow dashed circle). In practice, this means that X-ray sources whose centroids lie between the 37.5% and 75% EEF radii are not subtracted from the source map, and therefore contribute to a higher local background[9]. This is a key step to maintain a large enough aperture to encapsulate all the relevant source photons (especially for bright sources), but also filter out neighbouring spurious X-ray sources that are not associated with the target optical host and would otherwise erroneously enhance the detection probability. Given the broad eROSITA PSF wings and

---

[9] We note that as a result of setting this inner source-subtraction 37.5% EEF radius, we are not sensitive to off-centre X-ray emission from large, spatially extended galaxies, as is characteristic of ULXs. Although problems in the optical photometry pipeline leading to fragmented galaxies may result in off-centre X-ray detections in rare cases (see Ch. 5.3.3)

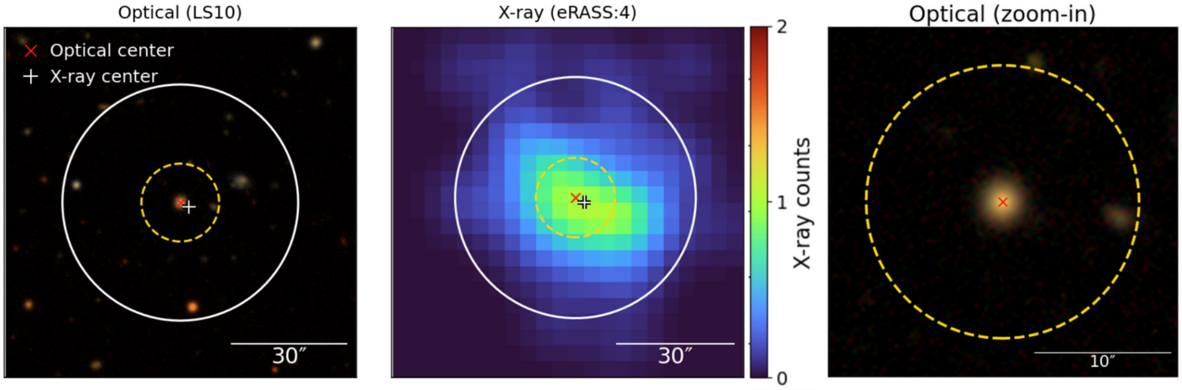


Figure 5.12: Example optical LS10 (left panel; right panel, zoom-in) and smoothed X-ray eRASS:4 (middle panel) images of an X-ray detected low-mass galaxy ($\log M_*/M_\odot \sim 9.55$ and spec-z of 0.18). The optical and X-ray centres are marked with a red and white cross, respectively. A colourbar indicates the number of X-ray counts. The white solid circle denotes the aperture used to extract X-ray photometry, it has a radius of $\sim 30''$. The yellow dashed circle with radius $\sim 10''$, marks the region outside which eRASS:4 sources contribute to the background counts within the total aperture.

the complex X-ray surface brightness profiles of extended sources, this subtraction may still be problematic around diffuse, extended or bright sources, which is why we adopted selection criterion 12 in Ch. 5.1 (masking of detected clusters and spurious sources) already in the parent sample definition.

The `apetool` algorithm computes a per-source Poisson tail probability (P_{thresh}) defined as the probability of observing $N \geq N_{\text{min}}$ counts given a Poisson distributed background with expected counts C_{B}, where N_{min} denotes the minimum number of counts required for detection at the chosen false-positive probability threshold (Georgakakis et al. 2008):

$$P_{\text{thresh}} = P(N \geq N_{\text{min}} \mid C_{\text{B}}) = \sum_{N=N_{\text{min}}}^{\infty} \frac{C_{\text{B}}^{N} e^{-C_{\text{B}}}}{N!}. \tag{5.2}$$

The smaller the value of P_{thresh}, the less likely it is that the total observed counts arise from a background fluctuation, and the more significant the detection of an astrophysical source. This probability can be used to define a minimum threshold for classifying a source as ‘X-ray detected’. We do so by firstly calculating the running median of P_{thresh} as a function of the one band (1B; $0.2 - 2.3$ keV) detection likelihood, `DET_LIKE_0` of the eRASS:4 source catalogues processed with pipeline version c030 (Lamer et al. in prep.). Then, using the eRASS1 ‘digital twin’ simulations by Seppi et al. (2022), we can infer the spurious detection fraction when opting for a `DET_LIKE_0` cut of ≥ 10 (D10) or ≥ 15 (D15). For D10, corresponding to a median value of $P_{\text{thresh}} \leq 1 \times 10^{-4}$, Seppi et al. (2022) compute a spurious fraction of $\sim$1% (see their Table 3). However, as there is a relatively large scatter on the running median, we choose to adopt a more stringent cut at D15, which corresponds to a median threshold of $P_{\text{thresh}} \leq 4 \times 10^{-6}$, similar to Georgakakis et al. (2008), with a spurious detection fraction $\sim 0.042\%$ (but still potentially reaching up to 1%, given the scatter)[10]. This threshold thereby sets a minimum number of source photons required, relative to the local background, to be considered a statistically significant X-ray detection. Although eRASS:4 is deeper than eRASS1 by a factor of ~ 4, the spurious fraction results from the eRASS1 'digital twin' still apply at a fixed `DET_LIKE_0` (Seppi et al., in prep.). This is consistent with the fact that, for a given source, the increased exposure time in eRASS:4 can lead to a correspondingly higher `DET_LIKE_0`.

We convert the aperture source counts (total minus background counts) to a net source count rate in the soft $0.2 - 2.3$ keV band by dividing by the EEF and the mean vignetted exposure time ($t_{\rm exp}$) at the galaxy position (computed by `apetool` from the exposure maps). Then, to calculate the soft X-ray flux ($F_{0.2-2.3\ \rm keV}$ in Eq. 5.3), we divide the count rate by a constant energy correction factor (ECF[11]) of 1.074×10^{12} counts cm^2 erg^{-1} (Table D.1 in Brunner et al. 2022), which is derived from an absorbed power law model with photon index, $\Gamma = 2.0$, and Galactic absorption column density $N_H = 3 \times 10^{20}$ cm^{-2}. The ECF also encodes information about the eROSITA effective area.

$$F_{0.2-2.3\ \rm keV} = (N - C_{\rm B})/(t_{\rm exp} \cdot EEF\ \cdot ECF). \tag{5.3}$$

Finally, we convert the soft X-ray flux to (rest-frame) $2-10$ keV luminosity, using the same photon index as above (i.e. we assume all detected sources are not affected by intrinsic obscuration) and the source redshift.

Running `apetool` on the parent galaxy sample resulted in 20,830 X-ray detections (i.e. $0 \leq P_{\rm thresh} \leq 4\times10^{-6}$). This is split into 4,121 X-ray detected low-mass, $\log(M_*/M_\odot) \leq 10$, galaxies and 16,709 X-ray detected high-mass, $\log(M_*/M_\odot) > 10$, galaxies.

5.3.2 Cleaning spurious associations using the eROSITA X-ray and counterpart catalogues

The $P_{\rm thresh}$ value computed for each galaxy gives an indication of the significance of the X-ray detection, however, we must still assess the reliability of the association of the X-ray photons with the host galaxy. In principle, the advantage of extracting X-ray photometry at the location of each optical galaxy is that it can homogeneously identify the X-ray emission originating from the source without requiring additional multi-wavelength catalogues or cross-matching. However, for the low-mass (and smaller physical size) galaxies in particular, this is complicated by the broad PSF of eROSITA, the rather large eROSITA positional uncertainty and the relatively high X-ray source density, dominated by more massive host galaxies, background AGN, stars and galaxy clusters.

Quantitively, we can show that we expect a high number of spurious associations in our parent sample apertures, especially for the low-mass sample. Taking the ~2.77 million

---

[10] We note that the full threshold is defined as $0 \leq P_{\rm thresh} \leq 4 \times 10^{-6}$, as `apetool` sets sources with zero aperture counts to $P_{\rm thresh} = -9.99$, which would otherwise be considered detections.

[11] `https://erosita.mpe.mpg.de/dr1/eSASS4DR1/eSASS4DR1_arfrmf/eROSITA_ECF_tutorial.pdf`

low-mass parent galaxy apertures, each with radius $10''$ (recall that the apertures themselves are $\sim 30''$ but sources outside $\sim 10''$ are not subtracted from the source map and so they contribute to the local background emission), gives a total area of 67 deg^2. Then, we find that the average sky density of eRASS:4 `DET_LIKE` > 10 extragalactic sources (excluding the south ecliptic pole and Galactic plane at $b \pm 20°$) is ~ 60 deg^{-2}. Therefore, if these $\sim$2.77 million apertures were randomly placed on the sky, one would expect 4020 chance associations, a very high number considering the actual number of detections at the positions of low-mass galaxies is $4,121$[12]. These crude estimates indicate that more than $\sim$90% of our 'detections' may be spurious associations; this clearly requires a careful cleaning process, which we proceed to describe.

The flowchart in Figure 5.13 describes the steps taken to validate the X-ray detections found via `apetool`, as well as the optical host associations. We explain the procedure in detail below, focusing on low-mass galaxies (high-mass galaxies are treated in the same way, except we do not visually inspect any sources). We note that identifying sources with unreliable LePHARE-derived stellar masses using our machine-learning classifier required prior knowledge of X-ray emission from a cleaned sample of galaxies (recall Ch. 5.2.1). Therefore, an iteration of this flowchart was first completed using only the LePHARE-derived stellar masses. Figure 5.13 shows the second iteration using the final LePHARE and GRAHSP stellar masses for which some sources changed from low- to high-mass subsets (or vice versa).

The first step is to see whether the parent sample galaxy has a match within $30''$ to a source in the entire eRASS:4 X-ray catalogue (multiple parent sample galaxies could match to the same eROSITA source at this point). This results in $3,979$ matches from the low-mass galaxy sample. The 142 unmatched sources are not necessarily spurious as `apetool` may be able to detect sources below the detection threshold of the catalogue and there is a small minority of real (bright) X-ray detections that are not present in the eRASS:4 catalogue. Therefore, all 142 unmatched sources are visually inspected to recover real X-ray sources associated with the target low-mass galaxies. Visual inspection is done by Z. Igo and A. Merloni, and the answer to the following four questions must be positive for the source to be considered an X-ray-detected low-mass galaxy: (i) is there a visually discernable agglomeration of X-ray photons indicating an X-ray source (i.e. the counts in the aperture are not scattered due to high local background levels)?; (ii) is the X-ray emission centred on the low-mass galaxy?; (iii) Are there no other possible contaminating X-ray sources that could have erroneously lowered the $P_{\rm thresh}$ value making it a false detection?; (iv) is the optical photometry of good quality, meaning that there are no artifacts, nearby bright objects such as stars or fragmentation? Figure 5.12 shows an example optical LS10 and X-ray eRASS:4 cutout of an X-ray detected low-mass galaxy.

The result of visually inspecting the 142 sources in this subset shows that in many

---

[12] An alternative test is to first match the $\sim$2.77 million low-mass parent sample to the $\sim$1.39 million eRASS:4 `DET_LIKE` > 10 extragalactic X-ray sources within a radius of $10''$, finding 5,582 'real' matches. Then we shift the low-mass parent sample apertures by $60''$ in their declinations, remove real low-mass galaxies within $10''$ of the shifted positions ($\sim 21,000$ removed), and rematch this catalogue to the same $\sim$1.39 million X-ray sources, providing us with 4,883 'fake' matches.

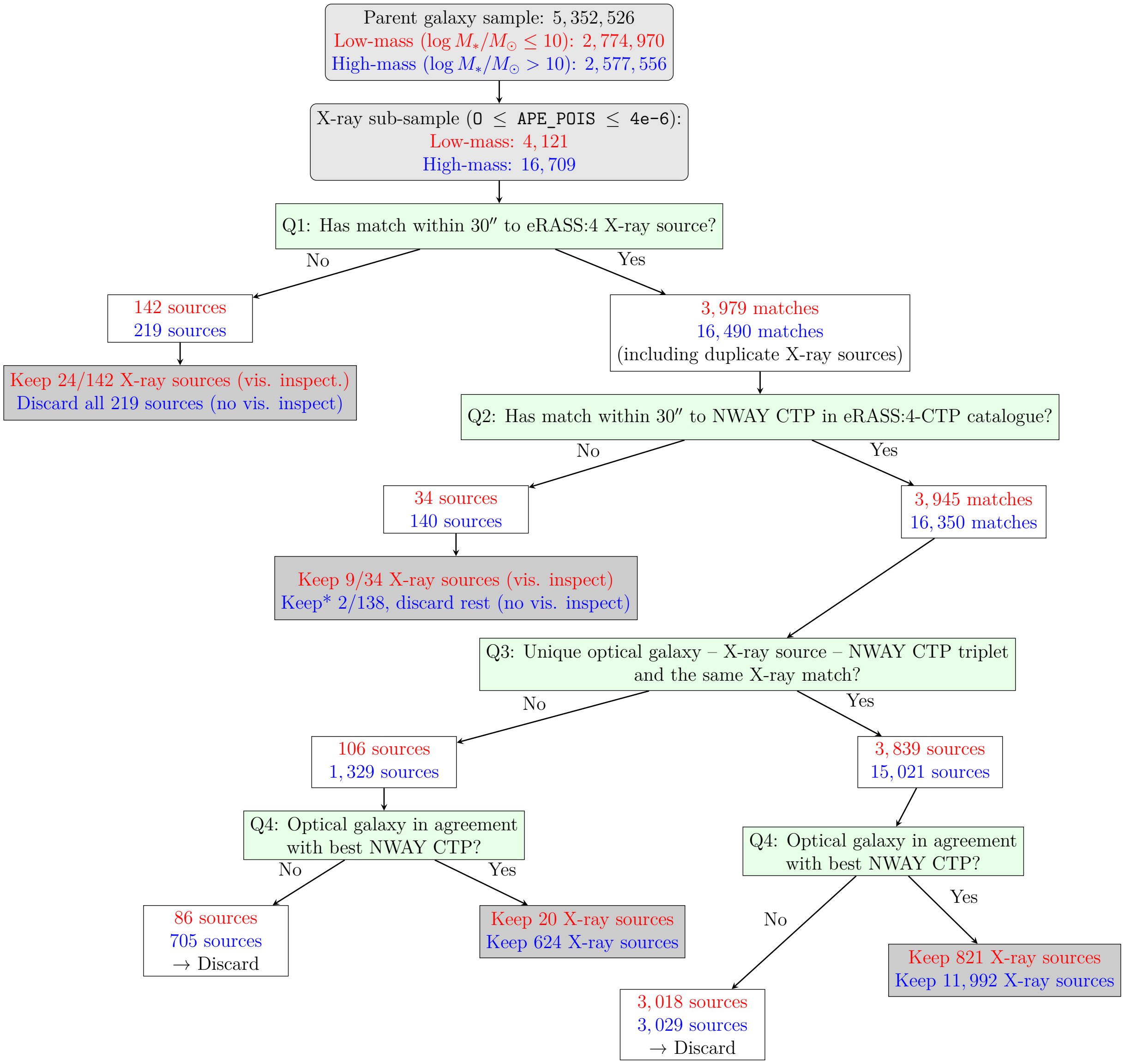


Figure 5.13: Flowchart showing the decision tree used to verify the validity of the X-ray emission and the association with the host galaxy. *Two sources without matches to an eRASS:4 counterpart (CTP) that were kept in the original visual inspection iteration using only LePHARE-derived stellar masses (see Ch. 5.2.1) changed from low-mass to high-mass and thus are kept as part of the final high-mass sample.

cases these sources are in regions of high X-ray background; are located near very bright X-ray sources which leak photons into the aperture; or simply the `ERMLDET` algorithm has failed to catalogue nearby X-ray sources properly, leading to them not being subtracted from the source map and thereby erroneously lowering the P_{thresh} value. Visual inspection helps save 24 X-ray sources passing the criteria defined above, although the large majority of these source lie close to the P_{thresh} cut-off, meaning that they are faint.

The second step of the flowchart involves associating the $3,979$ low-mass galaxies, with matches in the eRASS:4 catalogue, to the LS10-eRASS:4 counterpart catalogue (Salvato et al. in prep.). This step makes use of the Bayesian cross-matching algorithm NWAY (Salvato et al. 2018) which not only uses astrometric (distance) information, but also multi-wavelength priors learned from known X-ray sources and their host galaxy counterparts[13] (Salvato et al. 2022b, 2025). Out of 34 sources with no match within $30''$ to the eRASS:4 counterpart catalogue, 9 sources remain as X-ray emitting low-mass galaxy candidates as per the visual inspection criteria defined above. These are typically faint, diffuse and isolated low-mass galaxies with clear X-ray detections, that were likely under-represented in the training sample used to derive the X-ray prior or in some cases lie just outside the declination cut used in the eRASS:4-CTP catalogue.

For the other $3,945$ sources with a match, we split the sample into unique triplets of parent low-mass galaxy – eRASS:4 X-ray source – LS10 counterpart (making sure the match is to the same X-ray source from step 1 and 2) and duplicated match triplets, where multiple low-mass galaxies match to the same X-ray source or LS10 counterpart galaxy. For both of these subsets, we then check if the optical coordinates of the low-mass parent galaxy are in agreement with the LS10 optical counterpart assigned by the catalogue. If so, we keep these sources as X-ray detected low-mass galaxy candidates and if not, they are discarded from the sample. In the case of duplicated match triplets, this step also efficiently identifies the most favoured match, leaving only a unique optical-host-X-ray pair where the counterparts are in agreement. For the unique triplets we also make a sanity check to compare the total aperture counts derived using our method and those quoted in the eRASS:4 X-ray catalogue (using the same aperture radius) and they are in good agreement, as expected.

Given the high probability of spurious matches, as shown by the shifted aperture analysis above, it makes sense that $3,104$ ($3,018+86$; see “No” branches from Q4 in Fig. 5.13) sources are discarded this way. Simply put, if there is a nearby galaxy which is behaving exactly like a known X-ray emitter, it is the most likely origin of the detected X-ray emission; the target low-mass galaxy is at-most constrained by an upper limit in X-ray emission. However, it is clear statistically that there is a high spurious association as: (i) 82% of best-match LS10 eRASS:4 counterparts are closer in position to the X-ray source than the low-mass galaxy; (ii) 97% of them are within three times the positional error (`POS_ERR`) of the X-ray source; (iii) 90% of them have `p_i` > 0.9; (iv) 92% of them are

---

[13] If a source was heavily influenced by these additional priors the `bias_LS10_Xray_proba` will be $>> 1$. Additional columns returned by NWAY are the best-match flag (`match_flag = 1`), a probability for the match being the correct one (`p_i`) and a probability of the source in question having any counterpart at all in the search region (`p_any`).

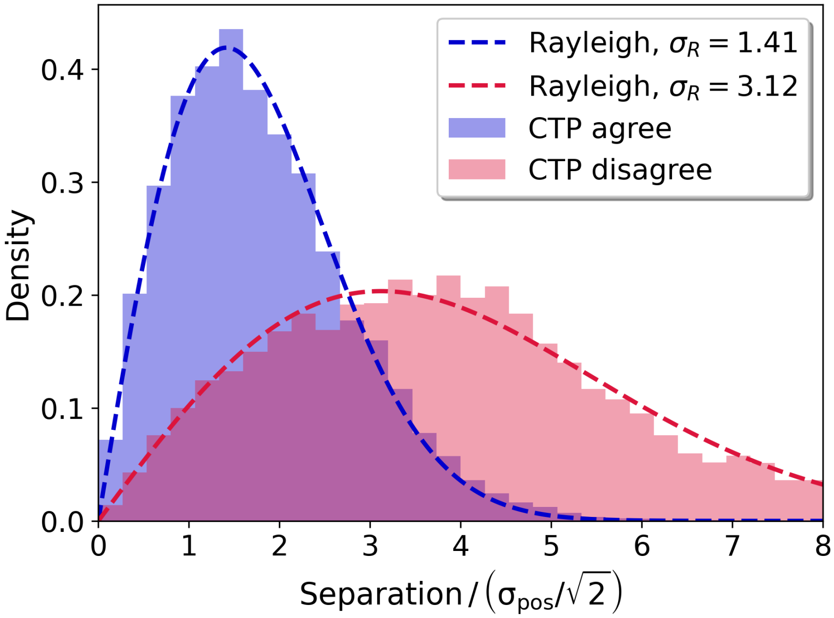


Figure 5.14: Rayleigh curve fit to the histogram of the separation between the parent galaxy and the nearest eRASS:4 source normalised by the one-dimensional eRASS:4 positional error for the sources in agreement (blue) and disagreement (red) with the eRASS:4 CTP catalogue (see Figure 5.13).

significantly influenced by the X-ray prior, meaning that their LS10 catalogue properties match the host galaxy properties of real X-ray emitters; and (v) the LS10 eRASS:4 counterparts lie in typical regions where bluer, more massive quasars are in the $g - r$ versus $z - W1$ colour-colour plot (see Fig. 18 in Salvato et al. 2022b), along with the majority of these having best-fit LS10 galaxy `TYPE=PSF`. Figure 5.14 shows the histogram of the separation between the parent galaxy and the nearest eRASS:4 source normalised by the one-dimensional eRASS:4 positional error for the sources in agreement (blue) and disagreement (red) with the eRASS:4 CTP catalogue. Sources in agreement can be well fitted by a Rayleigh distribution with $\sigma_R \sim 1$, as would be expected from cross-matching catalogues with Gaussian distributed astrometric errors (Pineau et al. 2017; Salvato et al. 2022b). Sources in disagreement show a very broad distribution, indicating the unreliability of the counterpart association, confirming our approach in Figure 5.13. Additionally, examining the NWAY information of the low-mass target sample in these cases reveals very low counterpart probabilities, meaning that they are not even close secondary counterparts (which would otherwise justify treating them as X-ray upper limits). All sources labelled as ‘discarded’ in Figure 5.13 are masked out from the parent sample and no longer used in future analysis.

Overall, we can finally confirm 874 (out of $4,121$) X-ray detected low-mass galaxies, of which 192 are in dwarf galaxies (i.e. $M_* < 10^{9.5} M_\odot$). This means that only 21% of the initial detections from the aperture photometry method are found to be secure associations, consistent with the high level of spurious contamination predicted from our simple estimates above (see also the extensive cleaning from spurious sources and counterpart associations done in e.g. Birchall et al. 2020; Latimer et al. 2021; Sacchi et al. 2024; Bykov et al. 2024). In terms of the high-mass galaxies, we find 12,618 X-ray candidates, using the same method (see Fig. 5.13). This corresponds to 76% of the original X-ray detections, showing that the high-mass galaxies suffer much less from contamination compared to the low-mass sample. Figure 5.15 shows the distribution of rest-frame $2-10$ keV luminosities (left panel) and observed $0.2-2.3$ keV fluxes (right panel) for the sample of X-ray detected low-mass (light red) and high-mass (grey) galaxies as a function of redshift.

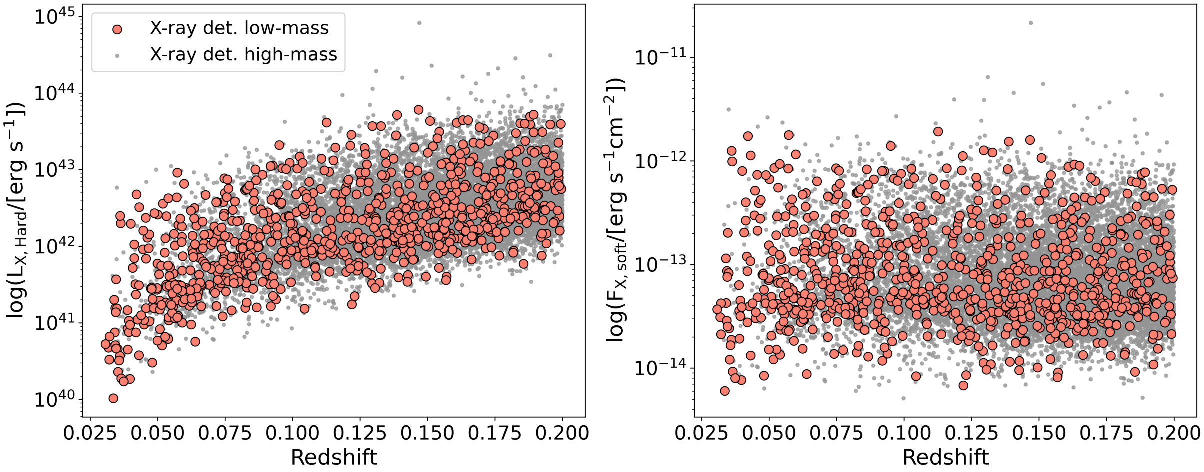


Figure 5.15: Left: Distribution of rest-frame $2-10$ keV luminosity versus redshift of the X-ray detected low-mass (light red) and high-mass (grey) galaxies from the parent sample. Right: Distribution of observed $0.2-2.3$ keV flux versus redshift of the same X-ray detected sources.

5.3.3 Origin of X-ray emission: galactic processes or accretion onto a central black hole?

While nuclear X-ray emission is one of the most effective indicators of a central accreting black hole, the galaxy can also produce X-ray emission from other processes that are unresolved at the spatial resolution of eROSITA. For example, galactic X-ray emission can come from low- and high-mass X-ray binaries (XRBs). The X-ray emission from the former scales with the stellar mass of the galaxy (tracing the long-lived, older stellar population) and that of the latter with the star formation rate (tracing the short-lived luminous stellar population; see review by Fabbiano 2006). Lehmer et al. (2016) were the first to parametrise this scaling in the form of:

$$L_{\mathrm{X,G}} = \alpha(1+z)^{\gamma}M_{*} + \beta(1+z)^{\delta}SFR^{\theta}. \tag{5.4}$$

This has since been refined by Aird et al. (2017), who use a similar method to the one described in Ch. 5.4.1 to find a peaked probability distribution as a function of luminosity at $L_X < 10^{42}$ erg s^{-1}. This peak, when analysed as a function of stellar mass, is attributed to the 'X-ray main sequence of star-formation'. We use the best-fit parameters to Eq. 5.4 found by Aird et al. (2017), which are the following: $\log\alpha = 28.81 \pm 0.08$, $\gamma = 3.90 \pm 0.36$, $\log\beta = 39.50 \pm 0.06$, $\delta = 0.67 \pm 0.31$ and $\theta = 0.86 \pm 0.05$.

Hot gas can also emit a faint and diffuse X-ray background via thermal bremsstrahlung. This is expected to be around an order of magnitude fainter than the XRB component, as the soft X-ray emission scales only as $L_{\mathrm{Gas}} = (8.3 \pm 0.1) \times 10^{38}$ SFR/[$M_\odot$ yr^{-1}] (Mineo et al. 2012, see also Fig. 8 of Lehmer et al. 2016 for a comparison). The relation from Aird et al. (2017) inherently includes this component as the authors do not attempt to separate XRB and hot-gas related X-ray emission.

As discussed in Chapter 5.2, the SFRs derived with only six photometric bands using LePHARE are not reliable. Therefore, we choose to estimate the SFR using the stellar mass of the galaxies, which we have now validated also for AGN-dominated sources using `GRAHSP`. It has been found that low-mass galaxies tend to be located in bluer, more star-forming galaxies (e.g. Kauffmann et al. 2003; Baldry et al. 2004; Papaderos et al. 2008), so for a given stellar mass we adopt a SFR that is on the main sequence of star formation at the galaxy's redshift (Eq. 28 of Speagle et al. 2014) and use it in Eq. 5.4 to compute the X-ray emission from galactic processes. Figure 5.16 shows the comparison of the $2-10$ keV X-ray emission from the parent sample of galaxies versus the expected emission from galactic processes. All X-ray detected low-mass galaxies lie above the 3:1 black solid line, meaning that their X-ray emission is more than a factor three greater than that expected from galactic processes and thus AGN-dominated. Therefore, we do not expect significant effects on the AGN incidence distributions, especially for the higher X-ray luminosity sources (but see Chapter 5.4.1 on how we account for this contribution explicitly in the Bayesian formalism).

Another luminous source of galactic X-ray emission are the so-called 'ultra-luminous X-ray sources' (ULXs; e.g. Fabbiano 1989; Feng & Soria 2011; Gúrpide et al. 2021; Walton et al. 2022, Weber et. al. in prep.). These are detected as non-nuclear X-ray point sources with luminosities exceeding the Eddington limit of a $10M_\odot$ black hole (approximately $> 10^{39}$ erg s^{-1}). As mentioned in Ch. 5.1, we filter for large, extended SGA and HECATE galaxies in the local Universe which are likely hosts of ULXs and are most subjected to photometric fragmentation. Unfortunately, even after the relatively stringent quality cuts listed in Ch. 5.1 we still find fragmented LS10 sources in our parent sample, a handful which could be associated with ULXs. After visually inspecting the optical and X-ray cutouts for all $4,121$ low-mass X-ray detected sources, we find around 50 sources with fragmented optical photometry, leading to potentially wrongly determined stellar mass. Extrapolating this to the full parent sample, we expect $\sim 1\%$ of sources to be affected. Therefore, we flag such problematic sources in the low-mass X-ray detected sample for future catalogue releases but do not attempt to correct their photometry as it will not

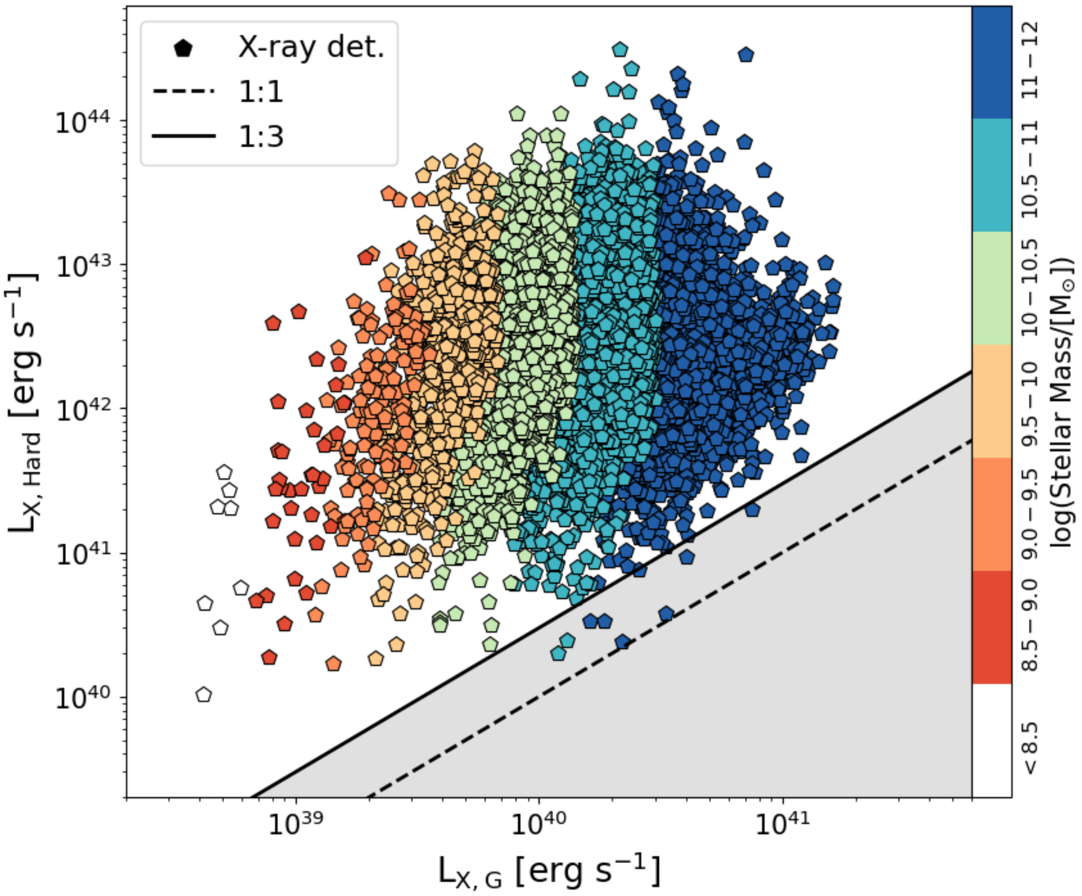


Figure 5.16: A comparison of the $2-10$ keV X-ray emission from parent sample of galaxies (colours denote a given stellar mass range as shown by the colourbar) versus the expected emission from galactic processes, namely unresolved X-ray binaries, derived using Eq. 5.4 with a SFR equal to the Main Sequence value for a given stellar mass. The black dashed and solid lines represent the 1:1 and 3:1 relations, respectively.

affect the statistical incidence results that are the main focus of this work[14].

5.4 The distribution of specific black hole accretion rates in low- and high-mass galaxies

We use a Bayesian methodology to compute the distribution of specific black hole accretion rates across the mass range covered by the parent galaxy sample. We then show how the AGN fraction, or duty cycle, varies with stellar mass and how this compares to previous

---

[14] We note also that 80% of the fragmented low-mass X-ray detected sources are ultimately discarded by the decision process shown in Fig 5.13 and so are not considered as real X-ray sources for the incidence distributions.

studies. From here on, we consider only mass-complete samples of galaxies (up to 70% for low-mass galaxies and ≫90% complete for high-mass galaxies; recall Ch. 5.2.2).

5.4.1 Methodology

Equivalently to Chapter 2, we use Eq. 2.8 to define a proxy for the Eddington ratio that we denote as the specific black hole accretion rate, $\lambda_{\rm Edd}$. We use the same constant scaling between $M_{\rm BH}$ and M_*, that is $M_{\rm BH} \sim 0.002\ M_*$, assuming the mass of the bulge is equal to M_* (Marconi & Hunt 2003), even though this relation is still uncertain in the low mass regime (e.g. Greene et al. 2020). Similarly, although $k_{\rm bol}$ has been found to increase with increasing Eddington ratio (see e.g. Vasudevan & Fabian 2009; Duras et al. 2020; Gupta et al. 2025), we use a constant value of 25 for all sources in our sample (see e.g. Marconi et al. 2004; Hopkins et al. 2007; Vasudevan & Fabian 2009; Duras et al. 2020). These choices are to ensure that our $\lambda_{\rm Edd}$ distribution remains as close as possible to the observable parameter space of $L_{\rm X}/M_*$.

We adopt the Bayesian framework described in Aird et al. (2017); Georgakakis et al. (2017); Aird et al. (2018), where our knowledge of $\lambda_{\rm Edd}$ can be described by a probability distribution:

$$p(\lambda_{\rm Edd}|D_i)d\lambda_{\rm Edd} = \mathcal{L}(N_i|\lambda_{\rm Edd}, b_i, t_i, z_i)\ \pi_A(\lambda_{\rm Edd}|M_{*,i}, z_i)d\lambda_{\rm Edd} \tag{5.5}$$

where D_i is the observed data from source i, $\mathcal{L}(N_i|\lambda_{\rm Edd}, b_i, t_i, z_i)$ is the likelihood of observing N_i counts from a source with specific accretion rate $\lambda_{\rm Edd}$, and $\pi_A(\lambda_{\rm Edd}|M_{*,i}, z_i)$ acts as a prior, describing the true underlying distribution of specific accretion rates of AGN in galaxies with stellar mass $M_{*,i}$ and redshift z_i.

The likelihood of observing N_i X-ray counts can be described by a Poisson process, thus,

$$\mathcal{L}(N_i|\lambda_{\rm Edd}, b_i, t_i, z_i) = \frac{\mu_i^{N_i}\ e^{-\mu_i}}{N_i!} \tag{5.6}$$

where the underlying (non-integer) total expected counts, μ_i, are defined as:

$$\mu_i = k_i(z_i) \cdot \lambda_{\rm Edd} \cdot M_{*,i} \cdot t_i + b_i, \tag{5.7}$$

where b_i are the (non-integer) background counts and $k_i(z)$ is a constant factor for each source at a given z_i that converts between specific accretion rate $\lambda_{\rm Edd}$ and observed counts including the EEF and ECF. In fact, the above expectation value can be modified to explicitly include the contribution from galactic processes (such as hot gas and X-ray binary emission, recall Ch. 5.3.3), via

$$\mu_i' = [k_i(z_i) \cdot \lambda_{\rm Edd} \cdot M_{*,i} + l_i \cdot L_{\rm X,G}] \cdot t_i + b_i, \tag{5.8}$$

where now $l_i L_{\rm X,G}$ represents the underlying (non-integer) counts from such galactic processes. Thus, we can rewrite the likelihood as a binomial expansion of the different contributors to the expected counts:

$$\begin{aligned}\mathcal{L}(N_i|\lambda_{\rm Edd}, L_{\rm X,G}, b_i, t_i, z_i) = \\ \sum_{C_{\rm A}=0}^{N_i} \sum_{C_{\rm G}=0}^{N_i - C_{\rm A}} \Bigg(&\frac{(k_i(z_i)\cdot\lambda_{\rm Edd}\cdot M_{*,i}\cdot t_i)^{C_{\rm A}}}{C_{\rm A}!} e^{-C_i\cdot\lambda_{\rm Edd}\cdot M_{*,i}\cdot t_i} \\ &\frac{(l_i\cdot L_{\rm X,G})^{C_{\rm G}}}{C_{\rm G}!} e^{-l_i\cdot L_{\rm X,G}} \\ &\frac{b_i^{C_{\rm B}}}{C_{\rm B}!} e^{-b_i}\Bigg), \qquad (5.9)\end{aligned}$$

where $C_{\rm A}, C_{\rm G}, C_{\rm B}$ are nuisance parameters representing the (integer) counts due to AGN, galactic and background processes, with the requirement that the total observed counts, N_i, fulfil: $N_i = C_{\rm A} + C_{\rm G} + C_{\rm B}$. The summations in Eq. 5.9 are equivalent to marginalising over all possible values of these nuisance parameters. We use Eq. 5.4 to derive a prior, $\pi_G(L_{\rm X,G}|M_{*,i}, \tilde{\rm SFR}_i)$ on the galactic contribution to the X-ray luminosity using the stellar mass of the galaxy and a SFR equal to the value of the Main Sequence of star-formation at the galaxy's stellar mass, as discussed in Ch. 5.3.3. We assume that the logarithm of $L_{\rm X,G}$ follows a normal distribution with an intrinsic scatter $\sigma = 0.2$ dex, meaning $\log L_{\rm X,G} \sim \mathcal{N}(\log L_{\rm X,G}, \sigma)$, and marginalise over this quantity.

The overall likelihood function for all the galaxies in a given stellar mass – redshift bin ($\mathbf{D}_{\rm bin}$) is:

$$\begin{aligned}\mathcal{L}(\mathbf{D}_{\rm bin}) &= \prod_{i=1}^{n_{\rm source}} \int_0^\infty p(\lambda_{\rm Edd}|D_i)\, d\lambda_{\rm Edd} \qquad (5.10)\\ &= \prod_{i=1}^{n_{\rm source}} \int_0^\infty \int_0^\infty \mathcal{L}(N_i|\lambda_{\rm Edd}, L_{\rm X,G}, b_i, t_i, z_i)\cdot \\ &\quad \pi_G(L_{\rm X,G}|M_{*,i}, \tilde{\rm SFR}_i)\, \pi_A(\lambda_{\rm Edd}|M_{*,\rm bin}, z_{\rm bin})\, dL_{\rm X,G}\, d\lambda_{\rm Edd}\end{aligned}$$

Recall that the goal is to derive the intrinsic probability distribution function of $\lambda_{\rm Edd}$, which we now rewrite as,

$$\pi_A(\lambda_{\rm Edd}|M_{*,\rm bin}, z_{\rm bin})\, d\lambda_{\rm Edd} = p(\log\lambda_{\rm Edd}|M_{*,\rm bin}, z_{\rm bin})\, d\log\lambda_{\rm Edd}, \qquad (5.11)$$

to indicate a probability density per unit $\log\lambda_{\rm Edd}$ in a given stellar mass ($M_{*,\rm bin}$, sampled in log-space) and redshift ($z_{\rm bin}$) bin. $p(\log\lambda_{\rm Edd}|M_*, z)$ represents the intrinsic distribution of $\lambda_{\rm Edd}$ that produces observed X-ray counts associated with the AGN and we synonymously refer to it as the incidence of X-ray AGN as a function of specific black hole accretion rate. We can then model this distribution as a step-function in discrete $\log\lambda_{\rm Edd}$ bins (which is effectively equivalent to a series of Gamma functions used in Aird et al. 2017, 2018, for a small enough $\log\lambda_{\rm Edd}$ bins). This model is flexible as it does not assume any functional form for $p(\log\lambda_{\rm Edd}|M_{*,\rm bin}, z_{\rm bin})$ and is only constrained by a prior that prefers a smooth variation[15] across $\log \lambda_{\rm Edd}$ bins and the requirement that this probability distribution function integrates to unity.

Using this parametrisation, we can reduce the overall likelihood function in Eq. 5.10 to a series of likelihood terms, w_{ij}, from each source i (independent of whether it is X-ray detected or not) multiplied by the value of the step function, θ_j, in a given $\log \lambda_{\rm Edd}$ bin j:

$$\mathcal{L}(\mathbf{D}_{\rm bin}) = \prod_{i=1}^{n_{\rm source}} \sum_j \theta_j \cdot w_{ij}. \tag{5.12}$$

This expression represents a mixture likelihood, where the likelihood of the data for each source is modelled as a weighted sum over all $\log \lambda_{\rm Edd}$ bins, with the mixture coefficients given by the population fractions θ_j. We fit this hierarchical Bayesian model using `CmdStan`, the command-line interface to the `Stan` statistical modelling language (Carpenter et al. 2017), which samples the population parameter space via Markov Chain Monte Carlo (MCMC) techniques and assesses convergence based on maximum likelihood estimation. During MCMC sampling, successive sets of population parameters ($\boldsymbol{\theta}$) are proposed and evaluated through the ratio of their posterior probabilities (proportional to the product of the mixture likelihoods across all sources) such that parameter combinations that better reproduce the observed data are explored more frequently. Overall, using the X-ray information from every parent sample galaxy, significantly increases our ability to constrain the $p(\log \lambda_{\rm Edd}|M_{*,\rm bin}, z_{\rm bin})$ distribution, also in the low-$\log \lambda_{\rm Edd}$ regime, where there are no longer any formal X-ray detections.

5.4.2 Incidence of X-ray AGN as a function of $\lambda_{\rm Edd}$

Figure 5.17 shows the $p(\log \lambda_{\rm Edd}|M_*, z)$ distributions in six stellar mass bins (rows) in the range of $\log M_*/M_\odot = 8.5-12$ and three redshift bins (columns), as found by the Bayesian modelling described in Ch. 5.4.1. The upper bound of each of the three redshift bins is defined such that the galaxies are mass-complete (to the 70% level) at $\log M_*/M_\odot = 8.5$, 9 and 9.5. The solid black curves indicate the best-fit $p(\log \lambda_{\rm Edd}|M_*, z)$ distribution using the galaxy prior as defined by the modified likelihood in Eq. 5.9, with solid coloured shaded regions indicating the 1σ uncertainties. The grey dashed histogram shows the observed $\log \lambda_{\rm Edd}$ distribution for only the X-ray detected sources (Ndet) normalised by the total number of galaxies (N) in that stellar mass and redshift bin (Ndet and N are stated in the top right corner of each panel). This observed histogram is not corrected for incompleteness and therefore differs significantly from the shape of $p(\log \lambda_{\rm Edd}|M_*, z)$ at low specific accretion rates. This demonstrates the advantage of the flexible Bayesian framework: by incorporating the full X-ray luminosity posterior distributions of all galaxies, we are able

---

[15] Smoothing is applied by uniformly sampling the first two $\log \lambda_{\rm Edd}$ bins (θ_1, θ_2) and then recursively defining the following $\log \lambda_{\rm Edd}$ bins via: $\theta_j = \theta_{j-1} + \Delta\theta_{j-2}$. Each increment $\Delta\theta$ is drawn from a standard normal prior $\Delta\theta \sim \mathcal{N}(0, 1)$, meaning that the subsequent θ bin can scatter 1 dex around the previous bin value (with no preference on the direction). The smoothing prior exerts its strongest influence at the very low and very high ends of $\log \lambda_{\rm Edd}$, where the data provide little information to constrain the posterior, while in the well-sampled intermediate range the posterior is primarily determined by the likelihood.

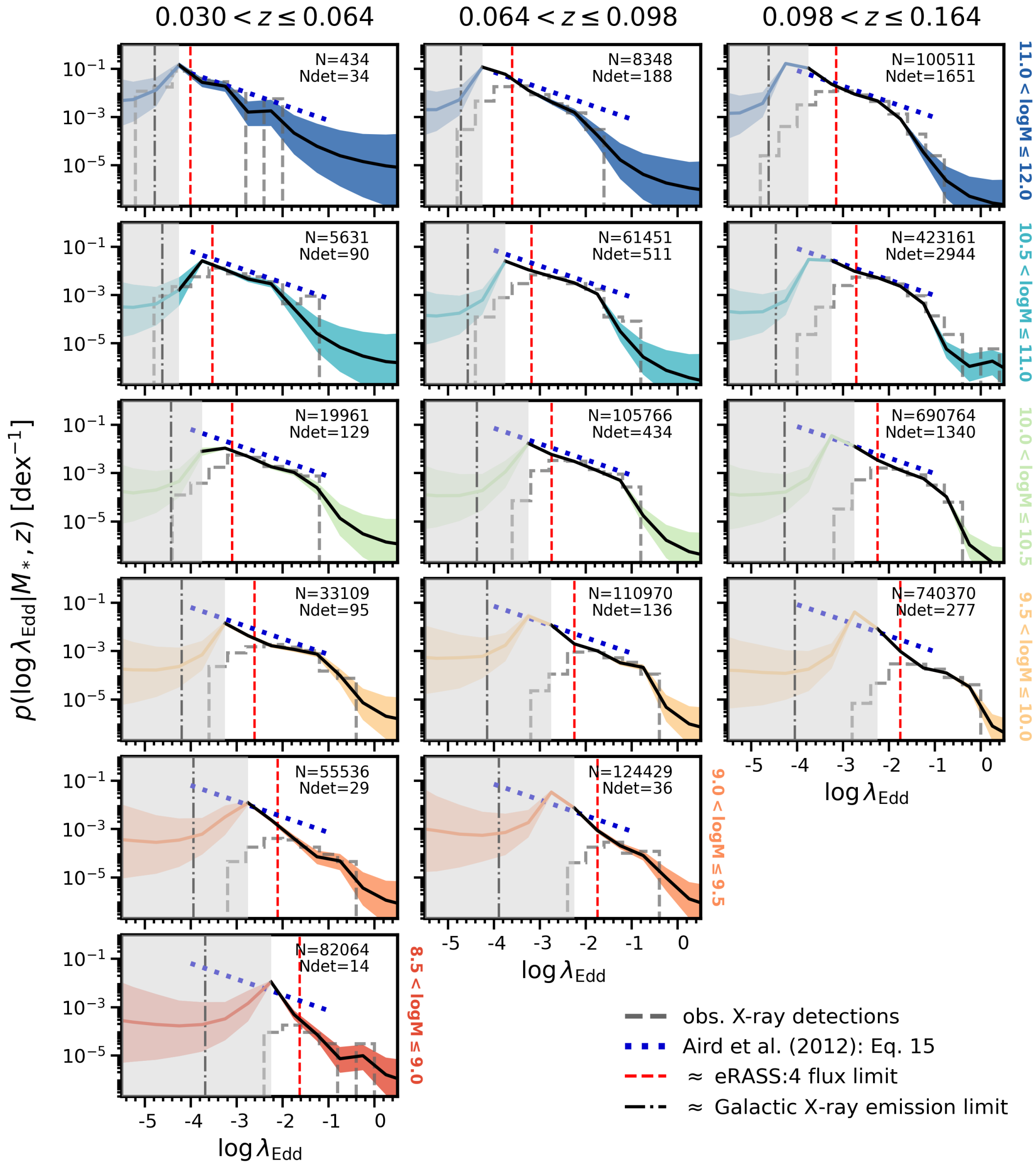


Figure 5.17: The incidence of X-ray AGN as a function of $\lambda_{\rm Edd}$, in different mass (rows; in units of $M_\odot$) and redshift (columns) bins: $p(\log \lambda_{\rm Edd}|M_*, z)$. The black curves and solid coloured shaded regions mark the best-estimate and 1σ confidence interval, respectively, as found by our Bayesian methodology. The grey dashed histogram represents the observed distribution of X-ray detected sources and the extrapolated results from Aird et al. (2012) are shown in blue dotted lines. Dashed red and dot-dashed black vertical lines indicate the approximate eRASS:4 flux limit and galactic X-ray emission contamination limit, respectively; grey shaded regions mark the low-λ regime which we cannot reliably constrain (see text for details).

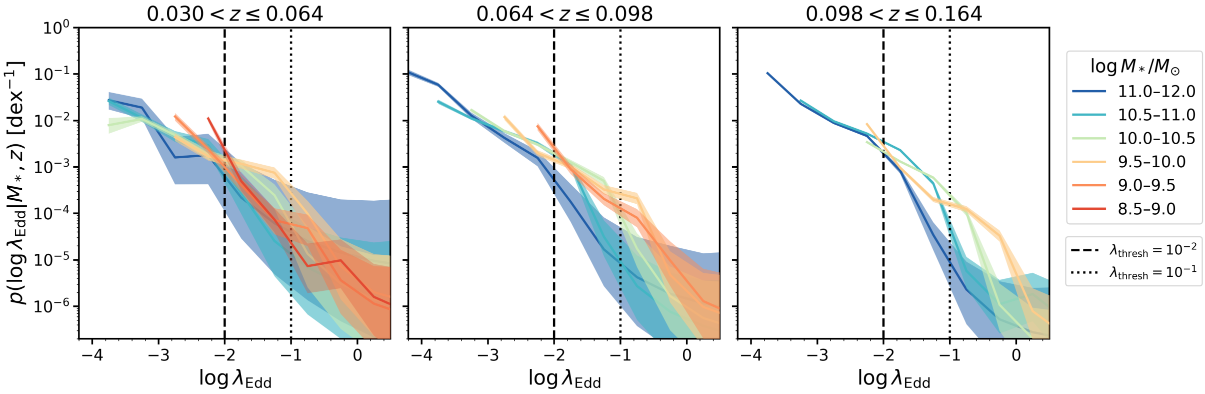


Figure 5.18: The same $p(\log\lambda_{\rm Edd}|M_*, z)$ distributions as in Figure 5.17 (without the uncertain low-$\lambda_{\rm Edd}$ regime), but overlaying the different stellar mass bins (colours) for a given redshift bin (panels). The dashed and dotted vertical lines correspond to $\lambda_{\rm thresh} = 10^{-2}$ and 10^{-1}, respectively. These are the thresholds above which the different cumulative AGN fractions are computed (see text).

to constrain $p(\log\lambda_{\rm Edd}|M_*, z)$ down to values of $\log\lambda_{\rm Edd}$ often an order of magnitude lower than those accessible through direct X-ray detections, owing to the stacking of the low-end tails of these distributions. Dashed red vertical lines indicate the approximate $0.2-2.3\,\mathrm{keV}$ eRASS:4 `DET_LIKE_0` > 10 flux limit of $\sim 2\times10^{-14}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$, converted to an estimate of $\lambda_{\rm Edd}$ using Eq. 2.8 and the median stellar mass and upper redshift bound of each bin. Dot-dashed black vertical lines indicate the 90th percentile of $L_{\rm X,G}/M_*$, considering all sources in a given bin, converted to $\lambda_{\rm Edd}$ with Eq. 2.8 as an estimate for the galactic X-ray contamination. Finally, as a comparison, the $p(\log\lambda_{\rm Edd}|M_*, z)$ curves found by Aird et al. 2012 (for $0.2 < z < 1.0$ and $9.5 \leq \log M_*/M_\odot \leq 12$), extrapolated to the median of each redshift bin, are shown in Figure 5.17 with dark blue dotted lines. Figure 5.18 shows the same $p(\log\lambda_{\rm Edd}|M_*, z)$ distributions from Figure 5.17 divided into three redshift bins (panels), with stellar mass bins overlaid to highlight their differences.

In general, we observe a broad distribution of $p(\log\lambda_{\rm Edd}|M_*, z)$. For high-mass galaxies, the shape of the incidence distribution follows well the power-law scaling with index around -0.65, found by Aird et al. (2012) in the regime $-4 \leq \log\lambda_{\rm Edd} \leq -2$. Whereas in the low-mass regime, particularly for $\log M_*/M_\odot \leq 9.5$, we observe a significant departure from the Aird et al. (2012) relation (see Ch. 5.5.2). We discuss in detail below the observed behaviour of $p(\log\lambda_{\rm Edd}|M_*, z)$ in the high-$\lambda_{\rm Edd}$ regime across the mass scale, as well as the reasons for which the low-$\lambda_{\rm Edd}$ regime is shaded grey in Figure 5.17 (denoting high uncertainty).

Considering first the high-$\lambda_{\rm Edd}$ regime, there is a noticeable break present between $\log\lambda_{\rm Edd} \sim -2$ and -1 across the mass scale. This regime corresponds to AGN nearing the Eddington limit, where the strong radiation pressure acts to regulate black-hole growth through powerful outflows that evacuate gas from the central regions (e.g. Hopkins et al. 2006; Fabian 2012). In fact, Aird et al. (2013) show that such a steep break in the modelled $p(\log \lambda_{\mathrm{Edd}}|M_*, z)$ distribution is necessary to reproduce well the XLF. Furthermore, obscuration cannot explain this break as the obscured AGN fraction is found to decrease with increasing Eddington ratio (e.g. Ricci et al. 2017; Ananna et al. 2022a; Ricci et al. 2022). Igo et al. (2024) explicitly demonstrate that when using only an unobscured selection function for eROSITA, the effects of obscuration appear confined to the low λ_{Edd} regime (see Ch. 3.2.1 and Fig. 3.4 above).

We note that the high-λ_{Edd} behaviour is notoriously difficult to confirm for a variety of reasons. Firstly, sources with such high specific accretion rates are intrinsically rare and can only be efficiently identified in wide-area X-ray surveys, among which eRASS currently provides the largest contiguous coverage of over $13,000$ deg^2. At the same time, deep (multi-wavelength) surveys are also necessary to characterise their host galaxy properties (or ideally their black hole masses directly) in detail. For the latter, one needs an unbiased method to determine stellar masses, in the presence of such strong AGN components at high λ_{Edd}. Previous studies often fully relied on SED-fitting codes such as LePHARE or CIGALE, which can be biased at such high specific accretion rates (Buchner et al. 2024), tending to overestimate stellar masses, and thus underestimating λ_{Edd}. This can artificially truncate the observed λ_{Edd} distributions. In contrast, our approach improves significantly on this aspect by calculating stellar masses using `GRAHSP` (for the subset of sources that require it, recall Ch. 5.2.1) and manages to recover a relatively smooth (but steep) decline in $p(\log \lambda_{\mathrm{Edd}}|M_*, z)$ high-λ_{Edd}.

Secondly, the specific accretion distribution of AGN is not equal the true Eddington ratio distribution function (ERDF) as it crucially depends on the scaling relations adopted to derive $L_{\mathrm{bol}}/L_{\mathrm{Edd}}$ from the observable L_{X}/M_*. Notably, the $M_* - M_{\mathrm{BH}}$ scaling has shown to be highly unconstrained in this low-mass regime (e.g. Greene et al. 2010; Kormendy & Ho 2013; Graham & Scott 2015; Reines & Volonteri 2015; van den Bosch 2016; Martín-Navarro & Mezcua 2018; Zaw et al. 2020; Greene et al. 2020; Pucha et al. 2025). This could have an impact on our interpretations as a low-mass galaxy could have a higher M_{BH} than predicted from our current scaling relation, shifting its true Eddington ratio to a lower value. Future spectroscopic AGN surveys such as 4MOST (de Jong et al. 2019; Merloni et al. 2019) or SPHEREx (Doré et al. 2018; Crill et al. 2020) will enable reliable M_{BH} to be derived for large AGN populations, and better constrain the low-mass regime of the $M_* - M_{\mathrm{BH}}$ relation.

While keeping the above limitations in mind, we observe the high-λ_{Edd} break to shift to lower λ_{Edd} values for higher stellar masses. This means that massive galaxies are less likely to reach high accretion rates than lower mass galaxies, possibly due to such galaxies being more quenched and gas poor, with less cold gas inflows able to fuel the AGN (e.g. Saintonge et al. 2011, 2017; Tacconi et al. 2018). The prominent break in our observed $p(\log \lambda_{\mathrm{Edd}}|M_*, z)$ distribution also points to a characteristic specific accretion rate that is stellar mass-dependent, $\lambda^*_{\mathrm{Edd}}(M_*)$, and not just a stochastic fuelling from the available gas supply (e.g. Hickox et al. 2014), as deduced from early studies finding a power-law trend (e.g. Aird et al. 2012). This is relevant as throughout this Thesis we have referred to the X-ray AGN incidence being mass-invariant *to first order*, but the results of this Chapter highlight important second-order effects, such as the mass-dependent high-λ_{Edd} break.

We defer explicit modelling of the $p(\log\lambda_{\rm Edd}|M_*,z)$ distributions, along with accurate determinations of any possible high-$\lambda_{\rm Edd}$ break, to future work.

In comparison, Ananna et al. (2022b) use the BAT AGN Spectroscopic Survey (BASS) DR2 to compute the ERDF and find a break around a similar value of $\lambda_{\rm Edd} \sim -1.5$. However, they do not find any mass-dependence in the shape of the ERDF. On the other hand, Bongiorno et al. (2016), probing a higher redshift range $0.3 < z < 2.5$ from the XMM-COSMOS survey, find a comparable mass-dependence in the high-$\lambda_{\rm Edd}$ break as our results. They attribute this to the 'AGN downsizing' phenomenon in the AGN luminosity function, whereby the space density of high luminosity AGN peaks at earlier cosmic epochs than that of low luminosity AGN (e.g. Ueda et al. 2003; Hasinger et al. 2005; Aird et al. 2015) and state that this is a consequence of the (weak) mass-dependent evolution of the host-galaxy mass function and the stronger mass-dependent evolution of the specific accretion rate distribution function.

In the low-$\lambda_{\rm Edd}$ regime, there are numerous effects which may affect the robustness of our $p(\log\lambda_{\rm Edd}|M_*,z)$ determination. Most importantly, although we rigorously clean the X-ray *detected* sample (removing 79% of the original detections; see Fig. 5.13), the impact of this spurious contamination (e.g. from wings of nearby bright sources or wrong counterpart associations) among the X-ray *undetected* sample must remain non-negligible for low-mass galaxies. This is most prominently visible in the $0.098 \leq z \leq 0.164$ and $9.5 < \log M_*/M_\odot \leq 10$ bin (yellow curve), where there is a steep rise below $\log\lambda_{\rm Edd} < -2$. As shown by the grey histogram, this happens in a region of parameter space where the observed X-ray detections are dwindling due to the flux limit of the survey (vertical red dashed lines). We note that the high-mass sources do not show this effect, as expected given the low (24%) spurious contamination. We try to quantify where this effect begins to dominate in a statistical manner by performing a shifted aperture analysis: we shift the Declinations of our parent galaxy apertures by $60''$, remove any overlap with real parent galaxies at the shifted positions (a negligible number), and recompute the X-ray aperture photometry in the same way as described in Ch. 5.3.1, keeping the stellar mass and redshift distribution the same. By examining the $p(\log\lambda_{\rm Edd}|M_*,z)$ derived for spurious sources, we are able to estimate a value of $\lambda_{\rm Edd}$ below which the incidence distribution in a given mass and redshift bin is dominated by contamination. We shade out the region below this value in Figure 5.17 and only plot $p(\log\lambda_{\rm Edd}|M_*,z)$ above this value in Figure 5.18 for clarity. Importantly, none of the main conclusions of this Chapter are affected as they deal with the high-$\lambda_{\rm Edd}$ regime.

As the additional contamination from unresolved stellar remnants (XRB) is expected to be important at even lower $\lambda_{\rm Edd}$ (see black dot-dashed vertical lines in Fig. 5.17), we do not discuss this in detail here. We do note, however, that past studies (e.g. Gilfanov 2004; Brorby et al. 2014; Lehmer et al. 2019; Kouroumpatzakis et al. 2020; Kyritsis et al. 2025) have found enhanced X-ray emission for lower-metallicity, lower-mass, higher-SFR galaxies and an increased scatter in the determination of $L_{\rm X,G}$, owed to high variability in the XRB population. Therefore, it may be that Eq. 5.4 is currently underestimating the galactic X-ray emission.

Phenomenologically, a low-$\lambda_{\rm Edd}$ turnover may be expected given that AGN transition from a 'radiative' to 'kinetic' mode of accretion towards lower specific accretion rates (e.g. Merloni & Heinz 2008; Heckman & Best 2014; Hardcastle & Croston 2020b; Harrison & Ramos Almeida 2024). AGN at these low $\lambda_{\rm Edd}$ values would no longer emit dominantly at X-ray wavelengths (as the accretion flow can no longer energetically maintain a hot X-ray–emitting corona), but instead become detectable in the radio bands through their jet kinetic emission (see e.g. Kondapally et al. 2022; Igo et al. 2024; Igo & Merloni 2025). Mathematically, this turnover is also required such that the $p(\log\lambda_{\rm Edd}|M_*, z)$ probability distribution is bounded and integrates to one. However, deeper X-ray and optical data is required to reliably constrain this low-$\lambda_{\rm Edd}$ turnover.

5.4.3 The cumulative AGN fraction

One way to summarise the information contained in Figures 5.17 and 5.18 is to compute the cumulative AGN fraction. We define this in a similar way to Aird et al. (2018), where the cumulative AGN fraction in a given stellar mass and redshift bin is the fraction of galaxies hosting an X-ray AGN with $\lambda_{\rm Edd} \geq \lambda_{\rm thresh}$:

$$f_{\rm AGN}(\lambda_{\rm Edd} \geq \lambda_{\rm thresh}) = \int_{\log\lambda_{\rm thresh}}^{\infty} p(\log\lambda_{\rm Edd}|M_*, z)\, d\log\lambda_{\rm Edd}. \tag{5.13}$$

The cumulative AGN fraction can be interpreted as an 'AGN duty cycle', meaning the fraction of time that a central massive black hole spends in an 'active' state compared to its full lifetime. Uncertainties in this quantity are propagated from the full posterior distributions of $p(\log\lambda_{\rm Edd}|M_*, z)$.

Figure 5.19 shows the cumulative AGN fraction for $\lambda_{\rm Edd} \geq 10^{-2}$ (left) and $\lambda_{\rm Edd} \geq 10^{-1}$ (right) as a function of stellar mass, in the different redshift bins (see legend). At these $\lambda_{\rm thresh}$ values, the incidence distributions are very well constrained by the data and are not significantly impacted by the eRASS:4 flux limit or spurious contamination (see Figure 5.18). Figure 5.19 (left) shows that, for these very low redshifts of $0.03 \leq z \leq 0.164$, the cumulative AGN fraction for $\lambda_{\rm Edd} \geq 10^{-2}$ ranges from $\sim 0.03\%$ to 0.1% for low-mass galaxies in the range of $8.5 \leq \log M_*/M_\odot \leq 10$. Interestingly, we observe $f_{\rm AGN}(\lambda_{\rm Edd} \geq 10^{-2})$ and $f_{\rm AGN}(\lambda_{\rm Edd} \geq 10^{-1})$ to peak around $\log M_*/M_\odot \sim 10 - 10.5$, potentially highlighting differences in the efficiency of fuelling in the high accretion rate regime at both lower and higher masses. The peak seems to align reasonably well with the mass range where the stellar-to-halo mass relation (and thus the efficiency of star formation) is maximal (e.g Behroozi et al. 2010; Girelli et al. 2020). This could potentially highlight a common physical condition, likely related to the cold gas reservoir in the central kpc of the galaxy, regulating both black-hole accretion and star formation (e.g. Hopkins & Quataert 2010; Mullaney et al. 2012; Hickox et al. 2014; Heckman & Best 2014; Harrison 2017; Gaspari et al. 2020). Figure 5.19 (right) also highlights the rare, but non-zero, presence of highly accreting AGN in low-mass galaxies.

It is important to note that by virtue of our soft selection using eRASS:4, we are not sensitive to obscured X-ray sources. According to Ricci et al. (2017), the obscured AGN fraction depends on the Eddington ratio, decreasing dramatically from $\sim 70\%$ at $\lambda_{\rm Edd} \sim$

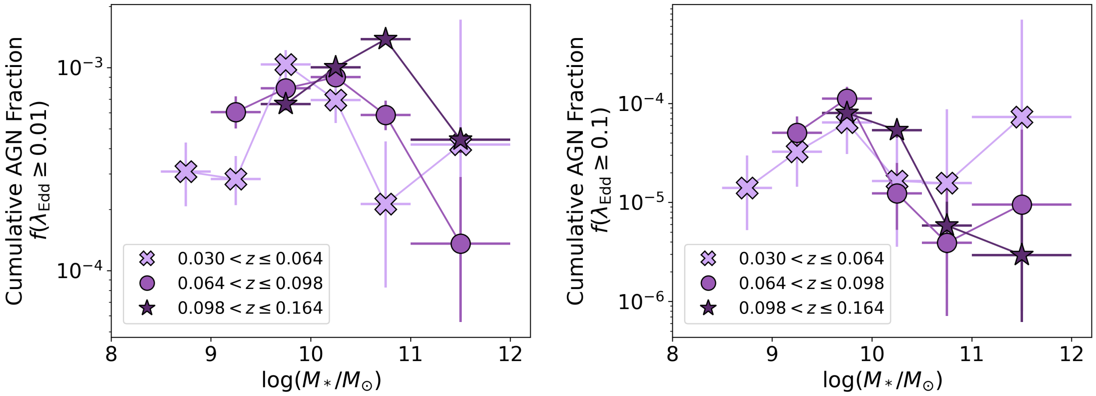


Figure 5.19: Cumulative AGN fraction for $\lambda_{\rm Edd} \geq 10^{-2}$ (left) and $\lambda_{\rm Edd} \geq 10^{-1}$ (right) as a function of stellar mass, in different redshift bins (see legend).

10^{-2} to $\sim 20\%$ at $\lambda_{\rm Edd} \sim 10^{-1}$. If we assume that, to first order, obscuration preferentially removes sources from our parent sample instead of attenuating their luminosity and that $\lambda_{\rm Edd}$ is a proxy for the Eddington ratio, we can estimate a value of the cumulative AGN fraction including both obscured and unobscured sources. This assumption implies that the shape of the $p(\log \lambda_{\rm Edd}|M_*, z)$ at high specific accretion rates would be similar for obscured and unobscured selections, an argument supported (to first order) by recent work of Laloux et al. 2024 (but see also Igo et al. 2024). Thus, by convolving the cumulative AGN fraction calculation as in Fig. 5.19 with the unobscured AGN fraction (1−obscured AGN fraction) as a function of $\lambda_{\rm Edd}$, we estimate an increase of $f_{\rm AGN}(\lambda_{\rm Edd} \geq 10^{-2})$ from $\sim 0.07\%$ to $\sim 0.18\%$ for the low-mass galaxies in the range of $8.5 \leq \log M_*/M_\odot \leq 10$ (taking the median across all mass and redshift bins).

5.5 Discussion

Using our well-defined, complete and rigorously cleaned sample of X-ray AGN in low-and high-mass galaxies in the local universe, we are able to place tight constraints on the $p(\log \lambda_{\rm Edd}|M_*, z)$ distribution and on the cumulative AGN fraction. In this Chapter, we compare our findings to existing literature samples and previous constraints on the X-ray AGN incidence distribution, along with discussing the extent to which local low-mass galaxies can be considered analogues to high-redshift primordial galaxies.

5.5.1 Comparison of X-ray sources with known literature samples

Out of the 874 (192) low-mass (dwarf) X-ray detected sources, only 246 (41) match to SIMBAD within 1″of the optical host galaxy coordinates[16], meaning 628 (151) are newly discovered. Out of the 12, 618 high-mass X-ray detected galaxies, 4, 748 are recorded in SIMBAD. Our work therefore provides the largest X-ray selected low-mass AGN sample to date, including the discovery of over 600 new sources. At the same time, it enlarges our database of X-ray AGN in high-mass galaxies, offering a control sample for comparison studies.

There are also a handful of low-mass galaxies with $\lambda_{\rm Edd} > 100\%$ (see Fig. 5.17). This may be an interesting subset as tidal disruption events (TDEs), where a star is disrupted upon close passage to a massive black hole, are thought to trigger intense accretion events in such low-mass galaxies (see e.g. review by Komossa 2015, and recent work by Lin et al. 2020; Chang et al. 2025). This could offer new ways to find previously dormant black holes in our local Universe.

Moreover, detailed X-ray spectroscopic follow-up of our low-mass X-ray sample may be key to test if black hole accretion mechanisms scale universally across the mass scale or there are some fundamental differences at the low-mass regime. This is motivated by studies at both low- and high-redshift that find differences in the expected X-ray emission from such sources, in comparison to their emission in other wavebands. For example, recent *JWST*-discovered high-redshift black holes, including the 'Little Red Dots', have been found to be X-ray undetected, even after stacking (e.g. Maiolino et al. 2024; Yue et al. 2024). This lack of X-ray emission may be due to extreme neutral hydrogen gas column densities ($N_H > 10^{24}$ cm^{-2}) absorbing the X-ray emission or such sources being intrinsically X-ray weak, possibly due to super-Eddington accretion (e.g. Pacucci & Narayan 2024). Similarly, at low redshift, studies on stacked samples of optical and IR variability-selected, X-ray-undetected massive black holes in low-mass galaxies indicate that the observed X-ray emission originates from galactic processes rather than from a central accreting AGN, potentially alluding to a lack of X-ray corona in these sources (Arcodia et al. 2024). Details of the multi-wavelength follow-up of this sample will be subject of future work.

5.5.2 Comparison of $p(\log \lambda_{\rm Edd}|M_*, z)$ distribution at low-masses to previous works

As mentioned in Chapter 1.9, past work has lacked large statistical samples to probe the X-ray AGN incidence in the low-mass regime. Figure 5.20 directly compares our $p(\log \lambda_{\rm Edd}|M_*, z)$ distributions, with the shaded curves now representing the 90th percentile confidence intervals (and same colour scheme as Fig. 5.17) to the work of Aird et al. 2012; Birchall et al. 2022; Zou et al. 2024. The results of Birchall et al. (2020) are not shown as

---

[16] Using a 5″ match to the eRASS:5 X-ray source coordinates also reveals a similar number of known sources.

they are in agreement with Birchall et al. (2022) and probe a very limited redshift range $z < 0.06$ with only 28 low-mass sources.

Figure 5.20 shows that our work sets the tightest constraints on $p(\log \lambda_{\rm Edd}|M_*, z)$ for $-3 \lesssim \log \lambda_{\rm Edd} \lesssim 0$, at low redshift $0.03 < z < 0.164$, for low-mass galaxies in the range $\log M_*/M_\odot = 8.5 - 10$. Aird et al. (2018) select X-ray AGN using deep *Chandra* observations of the Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey (CANDELS) and UltraVISTA surveys. They constrain $p(\log \lambda_{\rm Edd}|M_*, z)$ to a similar confidence level, but probe a higher redshift range, $0.1 < z < 0.5$ than our sample. In comparison to our lower redshift results, they find an increased X-ray AGN incidence at high $\lambda_{\rm Edd}$, especially towards lower masses ($\log M_*/M_\odot < 9.5$). Future work extending the redshift range of X-ray AGN samples in low-mass galaxies will be vital to understand their evolution as a function of cosmic time. Zou et al. (2024) compiles a sample of X-ray AGN among host-galaxies found in the CANDELS fields, four of the LSST Deep-Drilling Fields (DDFs), and the eFEDS field, also at $0.1 < z < 0.5$. Their results agree with Aird et al. (2018), but only sample down to $9.5 < \log M_*/M_\odot < 10$. Finally, Birchall et al. (2022), who identify AGN in the hard X-ray band with *XMM-Newton* from a parent galaxy sample defined by the MPA–JHU catalogue, recover a simple power-law form for the X-ray AGN incidence at moderate specific accretion rates over a redshift range comparable to ours. Importantly, their distributions agree with our results in that they show a steeper slope than the mass-invariant results of Aird et al. (2012) (blue dotted curves; extrapolated to the median redshift range of each bin), especially towards lower masses. This again highlights second-order mass-dependent effects of $p(\log \lambda_{\rm Edd}|M_*, z)$, which make it difficult to disentangle differences in the accretion processes or occupation fractions of low-mass galaxies.

5.5.3 Towards understanding early black hole seeding: interpretation of cumulative AGN fractions

While Birchall et al. (2020, 2022) constrain the $f_{\rm AGN}(\lambda_{\rm Edd} \geq 10^{-3})$ to $\sim 1\%$ for low-mass galaxies at $z \sim 0.1$ (see also e.g. Aird et al. 2018, who also probe low-mass galaxies at higher redshifts), Figure 5.18 shows that for low-mass galaxies in our eRASS:4-LS10 selected sample, we cannot reliably constrain the cumulative AGN fraction down to $\lambda_{\rm Edd} \geq 10^{-3}$. However, we emphasise that there exists a strong degeneracy between low occupation fractions and high active fractions versus high occupation fractions and low active fractions. We cannot currently distinguish between $\sim 1\%$ of low-mass galaxies hosting a central massive black hole, all of which are accreting, or $\gg 1\%$ hosting a central massive black hole, but only $\sim 1\%$ of which accreting. If one assumes a universal $p(\log \lambda_{\rm Edd}|M_*, z)$ distribution across the mass scale, one could make a statement on the BHOF by looking at the relative normalisations of the curves as a function of mass, anchoring the BHOF at $\sim 100\%$ at the highest masses (see e.g. Burke et al. 2025; Miller et al. 2015). Yet, we show in this Chapter that the X-ray AGN incidence has a more complex mass-, redshift- and $\lambda_{\rm Edd}$-dependent form which cautions against such an assumption. This is further

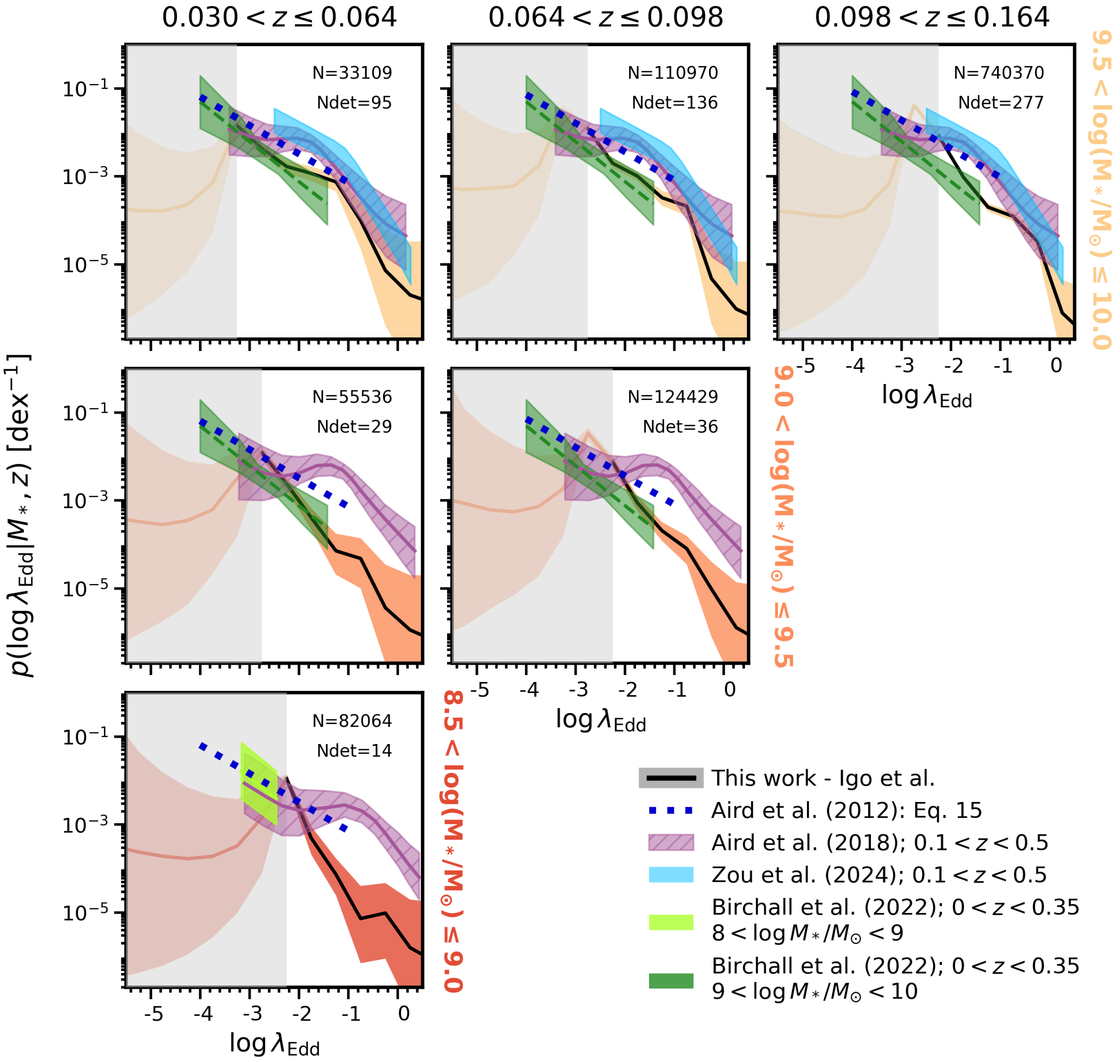


Figure 5.20: Comparison of the X-ray AGN incidence as a function of $\lambda_{\rm Edd}$ to other key works in the low-mass regime: Aird et al. 2018 (purple, hatched), Birchall et al. 2022 (light and dark green for different mass bins, see legend) and Zou et al. 2024 (blue). Differently to Fig. 5.17, the shaded intervals mark 90th percentile confidence intervals to be comparable with those of Aird et al. (2018); Zou et al. (2024). The confidence intervals on the curves from Birchall et al. (2022) are still at the 1σ level.

complicated by the fact that different multi-wavelength selection techniques find different cumulative AGN fractions (see e.g. Wasleske & Baldassare 2024 for a systematic selection of dwarf galaxies across various wavelength regimes or e.g. Menzel et al. 2016 for the case of higher mass systems).

From the simulations perspective, SAMs have repeatedly shown that distinguishing different black hole seeding models is very difficult with current data (e.g. Ricarte & Natarajan 2018; Burke et al. 2023). For example, Chadayammuri et al. (2023) find that, in order to do so, we would need to detect (or model) all AGN with $L_X > 10^{37}$ erg s^{-1} in galaxies of $\log M_*/M_\odot \sim 8 - 10$. This is orders of magnitude below the flux limit of eRASS:4 (see Fig. 5.16) and still out of reach of the deepest X-ray survey fields (e.g. Aird et al. 2018). Even detecting low BHOFs at low-masses is degenerate to slowly accreting heavy seeds or light seeds on fast growth channels (Chadayammuri et al. 2023). Although there are still uncertainties on the simulated BHOF itself that arise from the strong dependence on black hole and galaxy sub-grid prescriptions, as well as the challenges of accounting for obscured AGN in simulations (Haidar et al. 2022; Alonso-Tetilla et al. 2025). We discuss advancements in constraining the BHOF with future work in Ch. 6.

5.5.4 AGN growth and feedback in the low-mass regime: can local low-mass galaxies be considered high-redshift primordial galaxy analogues?

We detect a significant population of X-ray emitting dwarf galaxies ($\log M_*/M_\odot < 9.5$), with $L_\mathrm{X} > 10^{42}$ erg s^{-1} (see Figs. 5.16 and 5.19, right). This means that not only are AGN present in this low-mass regime —a long-standing uncertainty only clarified in the past few years— but they may also be growing and exerting a significant energetic influence on their surroundings. If this is the case, then such local low-mass galaxies cannot be considered analogues to high-redshift primordial galaxies. Although such high specific accretion rates may not necessarily coincide with high Eddington ratios for reasons already discussed above.

Nevertheless, semi-analytical models and hydrodynamical simulations struggle to create highly X-ray luminous AGN powered by low-mass black holes (e.g. Bellovary et al. 2019; Beckmann et al. 2023). This is because their accretion prescriptions (often following the simplistic 'Bondi-Hoyle-Lyttleton' model; Hoyle & Lyttleton 1939; Bondi & Hoyle 1944; Bondi 1952), have a strong dependence on the black hole mass as $\dot{m} \propto M_\mathrm{BH}^2$, where $\dot{m}$ is the physical accretion rate, which render the fuelling of low-mass black holes more difficult. In light of our new findings (see also e.g. Mezcua et al. 2018), such accretion prescriptions may need to be refined and the concept that AGN in low-mass galaxies are pristine remnants of high-redshift black hole seeds should be reconsidered.

Recently, several developments have emerged that weaken the link between the black holes in low-mass galaxies and their original seed masses, thereby complicating attempts to test seeding models using local-Universe analogues (Mezcua et al. 2019). For example, AGN feedback in low-mass galaxies is increasingly thought to play an important role, on par with or potentially even exceeding, stellar feedback from supernovae (e.g. Mezcua et al. 2015; Koudmani et al. 2019, 2021, 2022; Gim & Reines 2024; Rodríguez Morales et al. 2025; Salehirad et al. 2025). Additionally, mergers or other secular instabilities can produce off-centre ‘wandering’ black holes (e.g. Bellovary et al. 2021; Ricarte et al. 2021; Di Matteo et al. 2023; Eróstegui et al. 2025), which are harder to detect (and may even become dormant), potentially leading to an underestimation of the BHOF. Although Wu & Ho (2025) predict that such off-centre wandering black holes could produce micro-lensing effects on quasars detectable in upcoming surveys by LSST (Ivezić et al. 2019).

5.6 Summary

In this work, we compile the largest X-ray selected sample of low-mass ($\log M_*/M_\odot \leq 10$) galaxies in the local Universe using the deep four-pass eROSITA:DE all-sky survey. By combining this sample with a complementary set of high-mass galaxies ($\log M_*/M_\odot > 10$), we identify X-ray–detected AGN across the full mass range and investigate the distribution of their specific accretion rates, $p(\log \lambda_{\rm Edd}|M_*, z)$, through a Bayesian hierarchical inference framework. From this, we estimate the cumulative AGN fraction at varying thresholds in $\lambda_{\rm Edd}$.

Our parent galaxy sample consists of $\sim$ 5.35 million galaxies from LS10 that have good-quality optical photometry, are brighter than a z-band magnitude of 20 mag, and are in the redshift range $0.03 < z < 0.2$. We carefully derive galaxy properties through SED fitting with an AGN component (when needed) and compute X-ray fluxes for all sources using X-ray aperture photometry. We validate X-ray detections, along with their optical host galaxy associations through extensive cleaning procedures. This proves to be an essential step as 79% of the initial X-ray detections in the low-mass galaxy sample are in fact spurious X-ray sources or are the most probable counterpart of a nearby/background high-mass AGN or quasar. In comparison, the high-mass sample suffers much less from contamination, as expected, with 76% of the original X-ray detections being kept.

The $2-10$ keV luminosity range of the X-ray AGN is between $10^{40}-10^{45}$ erg s^{-1}, with several low-mass sources also exceeding $L_{\rm X} \sim 10^{43}$ erg s^{-1}. This observational finding sets an important constraint on the seeding and subsequent growth of black holes in low-mass galaxies in simulations, which currently struggle to create highly luminous X-ray sources (e.g. Bellovary et al. 2019; Beckmann et al. 2023).

Our Bayesian framework takes into account the detected X-ray counts (and background) from all galaxies, and thereby allows us to place tight constrains on the $p(\log \lambda_{\rm Edd}|M_*, z)$ distribution from $\lambda_{\rm Edd} \sim 10^{-4} - 10^{-3}$ all the way to the Eddington limit, $\lambda_{\rm Edd} = 1$. This range is wider than any past study at the same stellar mass and redshift range and allows us to reveal second-order mass-dependent properties of $p(\log \lambda_{\rm Edd}|M_*, z)$. Moreover, we find a steep break in the distribution at high $\lambda_{\rm Edd} \gtrsim 10^{-2} - 10^{-1}$, which is thought to indicate the Eddington-limited, self-regulated black hole growth.

Using the $p(\log \lambda_{\rm Edd}|M_*, z)$ distributions at low redshift ($0.03 \leq z \leq 0.164$), we derive a cumulative AGN fraction as a function of stellar mass for $\lambda_{\rm Edd} \geq 10^{-2}$ that ranges from $\sim$ 0.03% to 0.1% for low-mass galaxies, $8.5 \leq \log M_*/M_\odot \leq 10$. Interestingly, there is also a peak in the cumulative AGN fraction around $\log M_*/M_\odot \sim 10-10.5$, when considering only the highest accretion rate sources, $\lambda_{\rm Edd} \geq 10^{-2}$ and 10^{-1}. This could highlight differences in the efficiency of fuelling in this high accretion rate regime. Additionally, the peak aligns relatively well with the location where the star-formation efficiency as a function of halo mass is maximal, potentially pointing to a common physical condition that is responsible for AGN accretion and star-formation.

Our cumulative AGN fraction results clearly indicate that AGN do exist even in low-mass galaxies, and a non-negligible fraction of them are highly accreting. If AGN feedback is active in these AGN in low mass galaxies, as recent studies are finding, then this could challenge the long-standing view that stellar feedback alone drives low-mass galaxy evolution, revealing AGN as an equally, if not more, critical factor (e.g. Mezcua et al. 2019; Koudmani et al. 2019, 2021, 2022).

Chapter 6

Summary, Conclusions and Outlook

Even though AGN make up only a small fraction of all SMBHs and are orders of magnitude smaller in size and mass compared to their host galaxies, they are an important energetic component to understand galaxy evolution across cosmic time. This was already made clear in early cosmological hydrodynamical simulations that required AGN feedback (e.g. Croton et al. 2006; Bower et al. 2006; Sijacki et al. 2007), but observationally, we are now also able to construct large, complete, volume-limited samples of AGN, thanks to the numerous recent advancements in deep, wide-area (or even all-sky) surveys. This allows us to not only find observational evidence of AGN feedback in action, but characterise its overall behaviour as a function of AGN, host galaxy and temporal properties.

One such deep, all-sky survey is eRASS:4, the four-pass soft X-ray All Sky Survey of the eROSITA telescope. I started the main body of this Thesis in Chapter 2 by exploiting the rich multi-wavelength coverage of the pilot field of eRASS called eFEDS, covering $140\,\mathrm{deg}^2$ of extragalactic sky. In particular, a recent observation by the LOFAR radio observatory at 144 MHz, co-spatial with eFEDS, allowed me to compile a new sample of radio AGN in massive galaxies, with spectroscopic redshifts and host-galaxies characterised thanks to the optical GAMA09 survey. I classify the radio AGN into compact and complex morphologies, finding an abundant population of compact sources (around 2/3 of my mass-complete sample), which often have low radio luminosities. I also find new double-lobed FRII radio galaxies and giant radio galaxies that are included in the publicly released value added LOFAR-eFEDS catalogue produced as part of this work. This may serve as an interesting sample to follow-up with more multi-wavelength observations to probe the accretion mode, jet structure and jet triggering in these powerful sources. I do attempt to characterise the dominant accretion mode for radio- and X-ray detected or non-detected AGN in the LOFAR-eFEDS field. However, I find that I am limited in my discerning capabilities at both low- and high-λ_{Edd} regimes, due to the relatively shallow X-ray flux-limit of eFEDS and small survey volume, respectively.

In Chapter 3, I build on the X-ray and radio AGN from the LOFAR-eFEDS survey to calculate incidence distributions as a function of mass-scaled power indicators. Given the mass-invariant (to first order) triggering and fuelling observed in the X-ray AGN incidence as a function of λ_{Edd}, I was curious to examine if this mass-invariance of accretion translated into a corresponding mass-invariance of jet powering (and associated feedback). However, I found not only a mass-dependence, but also a power and morphology dependence in the way jets are launched and observed in radio AGN. Interestingly, the power and morphology dependence cannot be explained by environmental factors, i.e. higher-power and complex-morphology radio AGN did not preferentially lie in denser environments. Future work could explore this further by using the spectroscopic information from GAMA09 to compute filament catalogues (e.g. Malavasi et al. 2022) or find the associated dark matter halos and explore whether radio AGN, of a given power or morphology, preferentially cluster in a given type of environment (e.g. Magliocchetti 2022; Mingo et al. 2022; Popesso et al. 2024). This could give an indication on the density, temperature and metallicity of the gas in the vicinity of radio AGN and how this contributes to triggering jets.

Moreover, I investigate the radio AGN incidence as a function of host galaxy type and find that, once completeness effects are taken into consideration, quiescent and star-forming galaxies host radio AGN in similar fractions. Therefore, I argue that we should update our view on radio AGN only being found in 'red and dead' galaxies (see also e.g. Kondapally et al. 2022). I emphasize that the main findings in this Thesis are only possible thanks to the careful consideration of mass- and luminosity-completeness, counterpart association and cleaning from contaminants; I caution against drawing conclusions from sample comparisons across different works, or target versus control samples where the aforementioned factors are not taken into account.

Building on my finding that radio morphology plays a key role in understanding jet powering, I explicitly calculate the incidence of radio AGN as a function of $\lambda_{\rm Jet}$, treating compact and complex radio AGN separately. I find that the compact radio AGN incidence follows a steep power-law distribution with small mass-dependence, meaning compact radio AGN dominate the lower jet powers and drop out rapidly at higher jet powers. On the other hand, the complex radio AGN incidence follows a shallower, double power-law distribution with stronger mass-dependence, indicating that complex radio AGN progressively dominate at high M_* and $\lambda_{\rm Jet}$.

Intriguingly, the break in the double power-law distribution occurs around a value of $\lambda_{\rm Jet}$ where the FRII fraction sharply increases. It would be of considerable interest to further investigate the physical mechanisms that shape the observed incidence distributions. I do find a possible physical motivation in explaining the power-law slope for the case of compact AGN, making the simple assumption that the same intrinsic mass accretion distribution is responsible for the energy provided to power the jet. If this assumption holds, then compact sources can be considered to be accreting in a relatively radiatively inefficient mode. However, to gain an overall understanding, also for complex radio morphologies (which by definition come in a variety of shapes and sizes), one would need to couple our existing large-scale cosmological simulations (for the statistics) with our small-scale semi-analytic jet models (e.g. Hardcastle 2018; Hardcastle et al. 2019; Turner et al. 2023) or our sub-grid modelling of jet launching in zoom-in hydrodynamical simulations (for the detailed jet physics and environmental factors; e.g. Talbot et al. 2021; Bourne & Yang 2023). This is innately difficult due to the over 14 orders of magnitude in scales involved and so we are currently limited by using simple scaling relations to convert our observable radio luminosity to a jet power.

Nevertheless, I show in Chapter 4 that even using this simple $Q \propto L_{\mathrm{R}}^{0.68}$ scaling I can accurately reproduce the observed local AGN radio luminosity function by convolving the measured AGN incidence as a function of λ_{Jet} ($\propto Q/M_*$) with the stellar mass function and derive a total radio AGN kinetic luminosity density in line with previous work. In the same Chapter 4, I integrate the radio AGN incidence distributions to calculate the average jet power as a function of stellar mass and radio morphology for the entire massive galaxy population in the local Universe. My main findings are that the kinetic feedback from radio AGN dominates over any plausible inventory of radiatively-driven feedback and that it is the compact radio AGN that dominate this global kinetic energy budget for all but the most massive galaxies. This means that it is not only the powerful 3C 273-like sources (recall Ch. 1) that are doing most of the damage! I advocate for more importance to be given to this compact, low-luminosity radio AGN population, as they are not only abundant (as shown by the large majority of the LOFAR-eFEDS radio AGN being compact), but they are also energetically dominant over complex radio AGN, as I can now quantitatively substantiate. Future radio observations by the Square Kilometre Array (SKA) will provide the necessary sensitivity and resolution to characterise these sources across cosmic time, thereby enabling a more complete understanding of their role in galaxy evolution.

Additionally, we compare the average injected jet energy against the galaxy and larger-scale dark matter halo binding energy and against the total thermal energy of the host gas within the halos. We show that although radio AGN, be it compact or complex, do not have enough energy to fully disrupt the global gas distribution of massive galaxies (which makes sense as we observe such galaxies to exist in our Universe), they provide a significant source of heat that can impact the local thermodynamical balance in even the cores of the most massive clusters. Detailed follow-up work on this statistical radio AGN sample may be key to understand the kinetic energy coupling to the surrounding medium and average non-thermal pressure injected into the intra-cluster medium, which has recently been found by *XRISM* to be lower than expected (Xrism Collaboration et al. 2025). Understanding this coupling is also important in the context of cosmic rays, which are thought to play a crucial role in regulating the thermal balance of the intra-cluster medium and may represent a significant channel through which AGN-driven kinetic energy is dissipated into the surrounding gas (e.g. Ruszkowski & Pfrommer 2023).

Circling back to eRASS:4, in Chapter 5 I capitalise on this unique dataset from the deepest soft X-ray All Sky survey conducted to date to explore AGN accretion across the mass scale, particularly stretching to the low-mass galaxy regime of $\log M_*/M_\odot \leq 10$. Out of the parent galaxy sample of over 5 million optically-selected galaxies from LS10, I find 874 and $12,618$ X-ray emitting low- and high-mass galaxies, respectively, the large majority of which are AGN. This significantly expands the known sample of X-ray AGN in low-mass galaxies, especially considering that more than 600 of them are newly discovered. Interestingly, we also find low-mass galaxies with $2-10$ keV luminosities reaching values above 10^{43} erg s^{-1}, which challenges current black hole growth models in simulations that currently struggle to create such highly luminous low-mass X-ray sources (Bellovary et al. 2019; Beckmann et al. 2023). These new findings may caution against local low-mass galaxies being considered analogues to high-redshift primordial galaxies and may help future work refine the sub-grid prescriptions and develop more physically motivated accretion models, self-consistently bridging different accretion modes and mass regimes (e.g. Koudmani et al. 2024). Detailed X-ray spectroscopy follow-up of this low-mass X-ray sample with *XMM-Newton*, *XRISM*, *Chandra* and/or *NuSTAR* will additionally be key to test the properties of black hole accretion mechanisms. This tightly relates to recent discoveries regarding *JWST* AGN (e.g. Maiolino et al. 2024; Yue et al. 2024; Pacucci & Narayan 2024) and extreme transient phenomena in low-mass galaxies such as tidal disruption events or quasi-periodic eruptions (e.g. Komossa 2015; Liu et al. 2023; Arcodia et al. 2024).

Furthermore, in Chapter 5, I use a Bayesian hierarchical inference framework that takes into account the X-ray information from all parent sample galaxies to constrain the specific accretion rate distribution, $p(\log \lambda_{\rm Edd}|M_*, z)$, across a wide range of $\lambda_{\rm Edd}$, revealing second-order mass-dependent properties. Given the large survey volume, allowing for the detection of rarer high-accretion rate systems, I robustly detect a break at high-$\lambda_{\rm Edd}$, possibly indicating Eddington-limited, self-regulated black hole growth. Future work convolving our derived $p(\log \lambda_{\rm Edd}|M_*, z)$ with the stellar mass function may provide the most detailed synthesis of the (unobscured) X-ray luminosity function (see e.g. Hasinger et al. 2005, for previous results on this topic).

Then, integrating $p(\log \lambda_{\rm Edd}|M_*, z)$ above $\lambda_{\rm Edd} \geq 10^{-2}$ and $\geq 10^{-1}$, I derive the cumulative AGN fraction as a function of stellar mass. I find this cumulative AGN fraction to peak around $\log M_*/M_\odot \sim 10 - 10.5$, potentially highlighting a decrease in the efficiency of AGN fuelling at both very low and very high masses for such moderately- and highly-accreting AGN. The peak also aligns relatively well with the location where the star-formation efficiency as a function of halo mass is maximal. This could point to a common physical condition that is responsible for AGN accretion and star-formation, thereby linking the two processes within a common evolutionary framework. Placing a constraint on the black hole occupation fraction proved difficult with this eRASS:4−LS10 sample, primarily because more-in depth analysis is needed to account for the spurious contamination below $\lambda_{\rm Edd} < 10^{-2}$. However, given the large sample statistics, it could be possible to add a term to the likelihood to self-consistently account for the probability of a given galaxy to be nucleated, along with the probability of that galaxy to be an X-ray emitter given the observed X-ray counts (e.g. Burke et al. 2025).

Future wide-area surveys covering a few 100 deg^2 by *NewAthena* (Nandra et al. 2013; Cruise et al. 2025) will push studies of the incidence of AGN in low-mass galaxies to higher redshifts, which is currently only possible in the deepest *Chandra* fields and with much lower statistics. Meanwhile, LSST (Ivezić et al. 2019) and SPHEREx (Doré et al. 2018; Crill et al. 2020) will open up new channels to detect even larger samples of AGN in low-mass galaxies, through variability and spectroscopic data in the infrared, respectively. Finally, such studies help pave the way for next-generation observatories, such as *LISA* (Amaro-Seoane et al. 2023), which will transform our understanding of black hole growth by directly tracing the mergers of low-mass black hole seeds through their gravitational wave signatures, thereby opening an unprecedented window onto these extreme events.

Bibliography

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

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35

Abel, T., Bryan, G. L., & Norman, M. L. 2002, Science, 295, 93

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

Abramowicz, M. A., Czerny, B., Lasota, J. P., & Szuszkiewicz, E. 1988, ApJ, 332, 646

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

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

Aird, J., Coil, A. L., & Georgakakis, A. 2017, MNRAS, 465, 3390

Aird, J., Coil, A. L., & Georgakakis, A. 2018, MNRAS, 474, 1225

Aird, J., Coil, A. L., & Georgakakis, A. 2019, MNRAS, 484, 4360

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

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

Aird, J., Coil, A. L., Moustakas, J., et al. 2013, ApJ, 775, 41

Aird, J., Nandra, K., Laird, E. S., et al. 2010, MNRAS, 401, 2531

Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, ApJS, 219, 12

Alegre, L., Sabater, J., Best, P., et al. 2022, MNRAS, 516, 4716

Alexander, D. M. & Hickox, R. C. 2012, New Astron. Rev., 56, 93

Allen, S. W., Dunn, R. J. H., Fabian, A. C., Taylor, G. B., & Reynolds, C. S. 2006, MNRAS, 372, 21

Almeida, A., Anderson, S. F., Argudo-Fernández, M., et al. 2023, ApJS, 267, 44

Alonso-Tetilla, A. V., Shankar, F., Fontanot, F., et al. 2025, MNRAS[[arXiv]2509.04553]

Amarantidis, S., Afonso, J., Messias, H., et al. 2019, MNRAS, 485, 2694

Amaro-Seoane, P., Andrews, J., Arca Sedda, M., et al. 2023, Living Reviews in Relativity, 26, 2

Ananna, T. T., Urry, C. M., Ricci, C., et al. 2022a, ApJL, 939, L13

Ananna, T. T., Weigel, A. K., Trakhtenbrot, B., et al. 2022b, ApJS, 261, 9

Antoniou, V. & Zezas, A. 2016, MNRAS, 459, 528

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

Arcodia, R., Merloni, A., Comparat, J., et al. 2024, A&A, 681, A97

Arcodia, R., Merloni, A., Nandra, K., et al. 2021, Nature, 592, 704

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, M., Pratt, G. W., Piffaretti, R., et al. 2010, A&A, 517, A92

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

Aydar, C., Merloni, A., Dwelly, T., et al. 2025, arXiv e-prints, arXiv:2505.03872

Bahar, Y. E., Bulbul, E., Ghirardini, V., et al. 2024, A&A, 691, A188

Bahcall, J. N., Kirhakos, S., Schneider, D. P., et al. 1995, ApJL, 452, L91

Baldassare, V. F., Stone, N. C., Foord, A., Gallo, E., & Ostriker, J. P. 2022, ApJ, 929, 84

Baldi, R. D. 2023, A&A Rev., 31, 3

Baldini, P., Rau, A., Arcodia, R., et al. 2025, A&A, 701, A224

Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, ApJ, 600, 681

Baldry, I. K., Robotham, A. S. G., Hill, D. T., et al. 2010, MNRAS, 404, 86

Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5

Barkus, B., Croston, J. H., Piotrowska, J., et al. 2022, MNRAS, 509, 1

Barthel, P. D. & Arnaud, K. A. 1996, MNRAS, 283, L45

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

Beckmann, R. S., Dubois, Y., Volonteri, M., et al. 2023, MNRAS, 523, 5610

Beckmann, V. & Shrader, C. R. 2012, Active Galactic Nuclei (Wiley-VCH Verlag GmbH)

Begelman, M. C., Blandford, R. D., & Rees, M. J. 1984, Reviews of Modern Physics, 56, 255

Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379

Bellovary, J. M., Cleary, C. E., Munshi, F., et al. 2019, MNRAS, 482, 2913

Bellovary, J. M., Hayoune, S., Chafla, K., et al. 2021, MNRAS, 505, 5129

Bellstedt, S., Robotham, A. S. G., Driver, S. P., et al. 2021, MNRAS, 503, 3309

Bellstedt, S., Robotham, A. S. G., Driver, S. P., et al. 2020, MNRAS, 498, 5581

Bennett, A. S. 1962, , 68, 163

Berlind, A. A., Weinberg, D. H., Benson, A. J., et al. 2003, ApJ, 593, 1

Bernardi, M., Sheth, R. K., Fischer, J. L., et al. 2018, MNRAS, 475, 757

Best, P. N. 2004, MNRAS, 351, 70

Best, P. N. & Heckman, T. M. 2012, MNRAS, 421, 1569

Best, P. N., Kaiser, C. R., Heckman, T. M., & Kauffmann, G. 2006, MNRAS, 368, L67

Best, P. N., Kauffmann, G., Heckman, T. M., et al. 2005, MNRAS, 362, 25

Best, P. N., Ker, L. M., Simpson, C., Rigby, E. E., & Sabater, J. 2014, MNRAS, 445, 955

Best, P. N., Kondapally, R., Williams, W. L., et al. 2023a, MNRAS, 523, 1729

Best, P. N., Kondapally, R., Williams, W. L., et al. 2023b, arXiv e-prints, arXiv:2305.05782

Best, P. N., von der Linden, A., Kauffmann, G., Heckman, T. M., & Kaiser, C. R. 2007, MNRAS, 379, 894

Bicknell, G. V. 1986, ApJ, 300, 591

Birchall, K. L., Watson, M. G., & Aird, J. 2020, MNRAS, 492, 2268

Birchall, K. L., Watson, M. G., Aird, J., & Starling, R. L. C. 2022, MNRAS, 510, 4556

Birchall, K. L., Watson, M. G., Aird, J., & Starling, R. L. C. 2023, MNRAS, 523, 4756

Bîrzan, L., Rafferty, D. A., McNamara, B. R., Wise, M. W., & Nulsen, P. E. J. 2004, ApJ, 607, 800

Blake, C., Amon, A., Childress, M., et al. 2016, MNRAS, 462, 4240

Blandford, R., Meier, D., & Readhead, A. 2019, ARA&A, 57, 467

Blandford, R. D. & Begelman, M. C. 1999, MNRAS, 303, L1

Blandford, R. D. & Payne, D. G. 1982, MNRAS, 199, 883

Blandford, R. D. & Znajek, R. L. 1977, MNRAS, 179, 433

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

Bocquet, S., Saro, A., Dolag, K., & Mohr, J. J. 2016, MNRAS, 456, 2361

Boehringer, H., Voges, W., Fabian, A. C., Edge, A. C., & Neumann, D. M. 1993, MNRAS, 264, L25

Bogdán, Á., Goulding, A. D., Natarajan, P., et al. 2024, Nature Astronomy, 8, 126

Bogensberger, D., Nandra, K., & Buchner, J. 2024, A&A, 687, A21

Böhringer, H., Chon, G., & Collins, C. A. 2014, A&A, 570, A31

Bond, J. R., Arnett, W. D., & Carr, B. J. 1984, ApJ, 280, 825

Bondi, H. 1952, MNRAS, 112, 195

Bondi, H. & Hoyle, F. 1944, MNRAS, 104, 273

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

Booth, C. M. & Schaye, J. 2009, MNRAS, 398, 53

Boquien, M., Boselli, A., Buat, V., et al. 2013, A&A, 554, A14

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

Bourne, M. A. & Yang, H.-Y. K. 2023, Galaxies, 11, 73

Boutsia, K., Grazian, A., Calderone, G., et al. 2020, ApJS, 250, 26

Bower, R. G., Benson, A. J., & Crain, R. A. 2012, MNRAS, 422, 2816

Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645

Bowman, M., Leahy, J. P., & Komissarov, S. S. 1996, MNRAS, 279, 899

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

Brandt, W. N. & Alexander, D. M. 2015, A&A Rev., 23, 1

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151

Brinkmann, W., Laurent-Muehleisen, S. A., Voges, W., et al. 2000, A&A, 356, 445

Bromm, V. & Larson, R. B. 2004, ARA&A, 42, 79

Bromm, V. & Loeb, A. 2003, ApJ, 596, 34

Brorby, M., Kaaret, P., & Prestwich, A. 2014, MNRAS, 441, 2346

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

Brusa, M., Fiore, F., Santini, P., et al. 2009, A&A, 507, 1277

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

Buchner, J. 2021a, The Journal of Open Source Software, 6, 3045

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

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

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

Buchner, J. 2024, A&A, 689, L2

Buchner, J., Boller, T., Bogensberger, D., et al. 2022, A&A, 661, A18

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

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

Buchner, J., Schulze, S., & Bauer, F. E. 2017, MNRAS, 464, 4545

Buchner, J., Starck, H., Salvato, M., et al. 2024, A&A, 692, A161

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

Burke, C. J., Natarajan, P., Baldassare, V. F., & Geha, M. 2025, ApJ, 978, 77

Burke, C. J., Shen, Y., Liu, X., et al. 2023, MNRAS, 518, 1880

Burns, J. O. 1990, AJ, 99, 14

Butler, A., Huynh, M., Kapińska, A., et al. 2019, A&A, 625, A111

Bykov, S. D., Gilfanov, M. R., & Sunyaev, R. A. 2024, MNRAS, 527, 1962

Calistro Rivera, G., Williams, W. L., Hardcastle, M. J., et al. 2017, MNRAS, 469, 3468

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

Cameron, E. 2011, , 28, 128

Cannon, R., Drinkwater, M., Edge, A., et al. 2006, MNRAS, 372, 425

Carilli, C. L., Perley, R. A., & Harris, D. E. 1994, MNRAS, 270, 173

Carpenter, B., Gelman, A., Hoffman, M. D., et al. 2017, Journal of Statistical Software, 76, 1

Cavagnolo, K. W., McNamara, B. R., Nulsen, P. E. J., et al. 2010, ApJ, 720, 1066

Cavaliere, A. & Fusco-Femiano, R. 1976, A&A, 49, 137

Chabrier, G. 2003, PASP, 115, 763

Chadayammuri, U., Bogdán, Á., Ricarte, A., & Natarajan, P. 2023, ApJ, 946, 51

Chang, J. N. Y., Dai, L., Pfister, H., Kar Chowdhury, R., & Natarajan, P. 2025, ApJL, 980, L22

Chartas, G., Kochanek, C. S., Dai, X., Poindexter, S., & Garmire, G. 2009, ApJ, 693, 174

Chen, S.-J., Buchner, J., Liu, T., et al. 2025, A&A, 701, A144

Chen, T. & Guestrin, C. 2016, arXiv e-prints, arXiv:1603.02754

Chilingarian, I. V., Katkov, I. Y., Zolotukhin, I. Y., et al. 2018, ApJ, 863, 1

Ching, J. H. Y., Croom, S. M., Sadler, E. M., et al. 2017, MNRAS, 469, 4584

Cicone, C., Brusa, M., Ramos Almeida, C., et al. 2018, Nature Astronomy, 2, 176

Cicone, C., Maiolino, R., Sturm, E., et al. 2014, A&A, 562, A21

Coil, A. L., Blanton, M. R., Burles, S. M., et al. 2011, ApJ, 741, 8

Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039

Comparat, J., Luo, W., Merloni, A., et al. 2023, A&A, 673, A122

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

Comparat, J., Truong, N., Merloni, A., et al. 2022, A&A, 666, A156

Condon, J. J. 1992, ARA&A, 30, 575

Condon, J. J. & Ransom, S. M. 2016, Essential Radio Astronomy (Princeton University Press)

Cool, R. J., Moustakas, J., Blanton, M. R., et al. 2013, ApJ, 767, 118

Costa, T., Pakmor, R., & Springel, V. 2020, MNRAS, 497, 5229

Costa, T., Sijacki, D., & Haehnelt, M. G. 2014, MNRAS, 444, 2355

Courvoisier, T. J. L. 1998, A&A Rev., 9, 1

Crenshaw, D. M., Kraemer, S. B., & George, I. M. 2003, ARA&A, 41, 117

Crill, B. P., Werner, M., Akeson, R., et al. 2020, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11443, Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave, ed. M. Lystrup & M. D. Perrin, 114430I

Croom, S. M., Richards, G. T., Shanks, T., et al. 2009, MNRAS, 399, 1755

Croom, S. M., Smith, R. J., Boyle, B. J., et al. 2004, MNRAS, 349, 1397

Croston, J. H. & Hardcastle, M. J. 2014, MNRAS, 438, 3310

Croston, J. H., Hardcastle, M. J., Mingo, B., et al. 2019, A&A, 622, A10

Croston, J. H., Ineson, J., & Hardcastle, M. J. 2018, MNRAS, 476, 1614

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11

Croton, D. J., Stevens, A. R. H., Tonini, C., et al. 2016, ApJS, 222, 22

Cruise, M., Guainazzi, M., Aird, J., et al. 2025, Nature Astronomy, 9, 36

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

Dahlen, T., Mobasher, B., Faber, S. M., et al. 2013, ApJ, 775, 93

Dale, D. A., Helou, G., Magdis, G. E., et al. 2014, ApJ, 784, 83

Daly, R. A., Sprinkle, T. B., O'Dea, C. P., Kharb, P., & Baum, S. A. 2012, MNRAS, 423, 2498

Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, MNRAS, 486, 2827

Davis, M., Faber, S. M., Newman, J., et al. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4834, Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II, ed. P. Guhathakurta, 161–172

Davis, S. W. & Tchekhovskoy, A. 2020, ARA&A, 58, 407

de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, The Messenger, 175, 3

de Laplace, P. S. 1796, Exposition du système du monde de Vos, K., Hatch, N. A., & Merrifield, M. R. 2024, MNRAS, 535, 217

De Young, D. S. 1986, ApJ, 307, 62

Delvecchio, I., Daddi, E., Sargent, M. T., et al. 2021, A&A, 647, A123

DESI Collaboration, Abdul-Karim, M., Adame, A. G., et al. 2025a, arXiv e-prints, arXiv:2503.14745

DESI Collaboration, Abdul-Karim, M., Adame, A. G., et al. 2025b, arXiv e-prints, arXiv:2503.14745

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

Di Matteo, T., Ni, Y., Chen, N., et al. 2023, MNRAS, 525, 1479

Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604

Done, C. 2010, arXiv e-prints, arXiv:1008.2287

Done, C., Gierliński, M., & Kubota, A. 2007, A&A Rev., 15, 1

Doré, O., Werner, M. W., Ashby, M. L. N., et al. 2018, arXiv e-prints, arXiv:1805.05489

Drinkwater, M. J., Byrne, Z. J., Blake, C., et al. 2018, MNRAS, 474, 4151

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

Driver, S. P., Bellstedt, S., Robotham, A. S. G., et al. 2022, MNRAS, 513, 439

Dubois, Y., Devriendt, J., Slyz, A., & Teyssier, R. 2012, MNRAS, 420, 2662

Dubois, Y., Peirani, S., Pichon, C., et al. 2016, MNRAS, 463, 3948

Dunlop, J. S., McLure, R. J., Kukula, M. J., et al. 2003, MNRAS, 340, 1095

Dunn, R. J. H. & Fabian, A. C. 2006, MNRAS, 373, 959

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

Dutta, R., Sharma, P., Sarkar, K. C., & Stone, J. M. 2024, ApJ, 973, 148

Eardley, E., Peacock, J. A., McNaught-Roberts, T., et al. 2015, MNRAS, 448, 3665

Eberhard, J.-M., Reines, A. E., Gim, H. B., Darling, J., & Greene, J. E. 2024, arXiv e-prints, arXiv:2411.14535

Eckert, D., Gaspari, M., Gastaldello, F., Le Brun, A. M. C., & O'Sullivan, E. 2021, Universe, 7, 142

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

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

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

English, W., Hardcastle, M. J., & Krause, M. G. H. 2016, MNRAS, 461, 2025

Eróstegui, A., Mezcua, M., Siudek, M., Domínguez Sánchez, H., & Rodríguez Morales, V. 2025, A&A, 699, A330

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

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

Fabbiano, G. 1989, ARA&A, 27, 87

Fabbiano, G. 2006, ARA&A, 44, 323

Fabian, A. C. 1994, ARA&A, 32, 277

Fabian, A. C. 2012, ARA&A, 50, 455

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

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

Falcke, H., Körding, E., & Markoff, S. 2004, A&A, 414, 895

Fan, X., Bañados, E., & Simcoe, R. A. 2023, ARA&A, 61, 373

Fanaroff, B. L. & Riley, J. M. 1974a, MNRAS, 167, 31P

Fanaroff, B. L. & Riley, J. M. 1974b, MNRAS, 167, 31P

Fender, R. P., Belloni, T. M., & Gallo, E. 2004, MNRAS, 355, 1105

Feng, H. & Soria, R. 2011, New Astron. Rev., 55, 166

Ferrarese, L., Ford, H. C., & Jaffe, W. 1996, ApJ, 470, 444

Ferrarese, L. & Merritt, D. 2000, ApJL, 539, L9

Fiore, F., Feruglio, C., Shankar, F., et al. 2017, A&A, 601, A143

Flesch, E. 2013, , 30, e004

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

Frank, J., King, A., & Raine, D. J. 2002, Accretion Power in Astrophysics: Third Edition (Cambridge University Press)

Friedrich, P., Rohé, C., Gaida, R., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9144, Space Telescopes and Instrumentation 2014: Ultraviolet to Gamma Ray, ed. T. Takahashi, J.-W. A. den Herder, & M. Bautz, 91444R

Fujimoto, S., Ouchi, M., Kohno, K., et al. 2025, Nature Astronomy [[arXiv]2402.18543]

Gardner, J. P., Mather, J. C., Abbott, R., et al. 2023, PASP, 135, 068001

Garilli, B., Guzzo, L., Scodeggio, M., et al. 2014, A&A, 562, A23

Garilli, B., Le Fèvre, O., Guzzo, L., et al. 2008, A&A, 486, 683

Garilli, B., McLure, R., Pentericci, L., et al. 2021, A&A, 647, A150

Gaspari, M., McDonald, M., Hamer, S. L., et al. 2018, ApJ, 854, 167

Gaspari, M., Tombesi, F., & Cappi, M. 2020, Nature Astronomy, 4, 10

Gebhardt, K., Bender, R., Bower, G., et al. 2000, ApJL, 539, L13

Gehrels, N. 1986, ApJ, 303, 336

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

Genzel, R., Eisenhauer, F., & Gillessen, S. 2010, Reviews of Modern Physics, 82, 3121

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

Georgakakis, A., Nandra, K., Laird, E. S., Aird, J., & Trichas, M. 2008, MNRAS, 388, 1205

George, I. M. & Fabian, A. C. 1991, MNRAS, 249, 352

Ghez, A. M., Morris, M., Becklin, E. E., Tanner, A., & Kremenek, T. 2000, Nature, 407, 349

Ghez, A. M., Salim, S., Weinberg, N. N., et al. 2008, ApJ, 689, 1044

Ghirardini, V., Bulbul, E., Artis, E., et al. 2024, A&A, 689, A298

Ghirardini, V., Eckert, D., Ettori, S., et al. 2019, A&A, 621, A41

Gianolli, V. E., Bianchi, S., Petrucci, P.-O., et al. 2024, in Memorie della Societa Astronomica Italiana, Vol. 95, 27–31

Gianolli, V. E., Kim, D. E., Bianchi, S., et al. 2023, MNRAS, 523, 4468

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

Gilfanov, M. 2004, MNRAS, 349, 146

Gim, H. B. & Reines, A. E. 2024, ApJ, 963, 103

Girelli, G., Pozzetti, L., Bolzonella, M., et al. 2020, A&A, 634, A135

Godfrey, L. E. H. & Shabala, S. S. 2013, ApJ, 767, 12

Gofford, J., Reeves, J. N., Tombesi, F., et al. 2013, MNRAS, 430, 60

Gordon, C. & Arnaud, K. 2021, PyXspec: Python interface to XSPEC spectral-fitting program, Astrophysics Source Code Library, record ascl:2101.014

Graham, A. W. & Scott, N. 2015, ApJ, 798, 54

Greene, J. E. & Ho, L. C. 2004, ApJ, 610, 722

Greene, J. E. & Ho, L. C. 2007, ApJ, 670, 92

Greene, J. E., Peng, C. Y., Kim, M., et al. 2010, ApJ, 721, 26

Greene, J. E., Strader, J., & Ho, L. C. 2020, ARA&A, 58, 257

Greenstein, J. L. & Schmidt, M. 1964, ApJ, 140, 1

Grimm, H. J., Gilfanov, M., & Sunyaev, R. 2003, MNRAS, 339, 793

Gschwend, J., Rossel, A. C., Ogando, R. L. C., et al. 2018, Astronomy and Computing, 25, 58

Guetzoyan, P., Aird, J., Georgakakis, A., et al. 2025, MNRAS, 536, 79

Gupta, K. K., Ricci, C., Tortosa, A., et al. 2025, ApJ, 990, 86

Gürkan, G., Hardcastle, M. J., Smith, D. J. B., et al. 2018, MNRAS, 475, 3010

Gúrpide, A., Godet, O., Koliopanos, F., Webb, N., & Olive, J. F. 2021, A&A, 649, A104

Guzzo, L., Scodeggio, M., Garilli, B., et al. 2014, A&A, 566, A108

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

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

Habouzit, M., Li, Y., Somerville, R. S., et al. 2021, MNRAS, 503, 1940

Hahn, C., Wilson, M. J., Ruiz-Macias, O., et al. 2023, AJ, 165, 253

Haidar, H., Habouzit, M., Volonteri, M., et al. 2022, MNRAS, 514, 4912

Hardcastle, M. J. 2018, MNRAS, 475, 2768

Hardcastle, M. J. & Croston, J. H. 2020a, New Astron. Rev., 88, 101539

Hardcastle, M. J. & Croston, J. H. 2020b, New Astron. Rev., 88, 101539

Hardcastle, M. J., Evans, D. A., & Croston, J. H. 2007, MNRAS, 376, 1849

Hardcastle, M. J. & Krause, M. G. H. 2013, MNRAS, 430, 174

Hardcastle, M. J., Williams, W. L., Best, P. N., et al. 2019, A&A, 622, A12

Harrison, C. 2014, PhD thesis, Durham University, UK

Harrison, C. M. 2017, Nature Astronomy, 1, 0165

Harrison, C. M., Costa, T., Tadhunter, C. N., et al. 2018, Nature Astronomy, 2, 198

Harrison, C. M. & Ramos Almeida, C. 2024, Galaxies, 12, 17

Hasinger, G., Miyaji, T., & Schmidt, M. 2005, A&A, 441, 417

Hearin, A. P., Zentner, A. R., Ma, Z., & Huterer, D. 2010, ApJ, 720, 1351

Heckman, T. M. & Best, P. N. 2014, ARA&A, 52, 589

Heckman, T. M. & Best, P. N. 2023, Galaxies, 11, 21

Heesen, V., Schulz, S., Brüggen, M., et al. 2023, arXiv e-prints, arXiv:2309.05732

Heinz, S. & Sunyaev, R. A. 2003, MNRAS, 343, L59

HI4PI Collaboration, Ben Bekhti, N., Flöer, L., et al. 2016, A&A, 594, A116

Hickox, R. C. & Alexander, D. M. 2018, ARA&A, 56, 625

Hickox, R. C., Jones, C., Forman, W. R., et al. 2009, ApJ, 696, 891

Hickox, R. C., Mullaney, J. R., Alexander, D. M., et al. 2014, ApJ, 782, 9

Hill, G. J., Lee, H., MacQueen, P. J., et al. 2021, AJ, 162, 298

Hlavacek-Larrondo, J., Li, Y., & Churazov, E. 2022, in Handbook of X-ray and Gamma-ray Astrophysics, ed. C. Bambi & A. Sangangelo, 5

Hönig, S. F. 2019, ApJ, 884, 171

Hopkins, P. F., Grudic, M. Y., Su, K.-Y., et al. 2024, The Open Journal of Astrophysics, 7, 18

Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, ApJS, 163, 1

Hopkins, P. F. & Quataert, E. 2010, MNRAS, 407, 1529

Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, ApJ, 654, 731

Hoyle, F. & Lyttleton, R. A. 1939, Proceedings of the Cambridge Philosophical Society, 35, 405

Hubbard, A. & Blackman, E. G. 2006, MNRAS, 371, 1717

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168

Hudson, D. S., Mittal, R., Reiprich, T. H., et al. 2010, A&A, 513, A37

Ichikawa, K., Yamashita, T., Merloni, A., et al. 2023, A&A, 672, A171

Igo, Z. & Merloni, A. 2025, A&A, 697, A196

Igo, Z., Merloni, A., Hoang, D., et al. 2024, A&A, 686, A43

Igo, Z., Parker, M. L., Matzeu, G. A., et al. 2020, MNRAS, 493, 1088

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236

Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55

Ineson, J., Croston, J. H., Hardcastle, M. J., & Mingo, B. 2017, MNRAS, 467, 1586

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111

Jacob, S. & Pfrommer, C. 2017, MNRAS, 467, 1449

Jahnke, K. & Macciò, A. V. 2011, ApJ, 734, 92

Jansky, K. G. 1933, Nature, 132, 66

Jones, D. H., Read, M. A., Saunders, W., et al. 2009, MNRAS, 399, 683

Jones, D. H., Saunders, W., Colless, M., et al. 2004, MNRAS, 355, 747

Jordan, P. & Wigner, E. 1928, Zeitschrift fur Physik, 47, 631

Käfer, F., Finoguenov, A., Eckert, D., et al. 2019, A&A, 628, A43

Kammoun, E. S., Igo, Z., Miller, J. M., et al. 2023, MNRAS, 522, 5217

Kauffmann, G., Heckman, T. M., & Best, P. N. 2008, MNRAS, 384, 953

Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, MNRAS, 346, 1055

Kaviraj, S., Laigle, C., Kimm, T., et al. 2017, MNRAS, 467, 4739

Kaviraj, S., Martin, G., & Silk, J. 2019, MNRAS, 489, L12

Kawamuro, T., Ricci, C., Imanishi, M., et al. 2022, ApJ, 938, 87

Kellermann, K. I., Condon, J. J., Kimball, A. E., Perley, R. A., & Ivezić, Ž. 2016, ApJ, 831, 168

Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, MNRAS, 372, 961

Khostovan, A. A., Kartaltepe, J. S., Salvato, M., et al. 2025, arXiv e-prints, arXiv:2503.00120

King, A. 2003, ApJL, 596, L27

King, A. 2005, ApJL, 635, L121

King, A. & Pounds, K. 2015, ARA&A, 53, 115

Klindt, L., Alexander, D. M., Rosario, D. J., Lusso, E., & Fotopoulou, S. 2019, MNRAS, 488, 3109

Kluge, M., Comparat, J., Liu, A., et al. 2024, A&A, 688, A210

Kochanek, C. S., Eisenstein, D. J., Cool, R. J., et al. 2012, ApJS, 200, 8

Kollmeier, J. A., Rix, H.-W., Aerts, C., et al. 2025, arXiv e-prints, arXiv:2507.06989

Komossa, S. 2015, Journal of High Energy Astrophysics, 7, 148

Kondapally, R., Best, P. N., Cochrane, R. K., et al. 2022, MNRAS, 513, 3742

Kondapally, R., Best, P. N., Hardcastle, M. J., et al. 2021, A&A, 648, A3

Kondapally, R., Best, P. N., Raouf, M., et al. 2023, MNRAS, 523, 5292

Körding, E. & Falcke, H. 2005, , 76, 80

Körding, E. G., Jester, S., & Fender, R. 2006, MNRAS, 372, 1366

Kormendy, J. & Ho, L. C. 2013, ARA&A, 51, 511

Kormendy, J. & Richstone, D. 1995, ARA&A, 33, 581

Koudmani, S., Henden, N. A., & Sijacki, D. 2021, MNRAS, 503, 3568

Koudmani, S., Sijacki, D., Bourne, M. A., & Smith, M. C. 2019, MNRAS, 484, 2047

Koudmani, S., Sijacki, D., & Smith, M. C. 2022, MNRAS, 516, 2112

Koudmani, S., Somerville, R. S., Sijacki, D., et al. 2024, MNRAS, 532, 60

Kouroumpatzakis, K., Zezas, A., Sell, P., et al. 2020, MNRAS, 494, 5967

Kovlakas, K., Zezas, A., Andrews, J. J., et al. 2021, MNRAS, 506, 1896

Kyritsis, E., Zezas, A., Haberl, F., et al. 2025, A&A, 694, A128

Lacy, M., Baum, S. A., Chandler, C. J., et al. 2020, PASP, 132, 035001

Laha, S., Reynolds, C. S., Reeves, J., et al. 2021, Nature Astronomy, 5, 13

Laloux, B., Georgakakis, A., Alexander, D. M., et al. 2024, MNRAS, 532, 3459

Lang, D. 2014, AJ, 147, 108

Lang, D., Hogg, D. W., & Schlegel, D. J. 2016, AJ, 151, 36

Laor, A. & Behar, E. 2008, MNRAS, 390, 847

Latif, M. A., Schleicher, D. R. G., & Khochfar, S. 2023, ApJ, 945, 137

Latimer, L. J., Reines, A. E., Bogdan, A., & Kraft, R. 2021, ApJL, 922, L40

Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013, A&A, 559, A14

Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, A&A, 439, 845

Lehle, K., Nelson, D., Pillepich, A., Truong, N., & Rohr, E. 2024, A&A, 687, A129

Lehmer, B. D., Basu-Zych, A. R., Mineo, S., et al. 2016, ApJ, 825, 7

Lehmer, B. D., Eufrasio, R. T., Tzanavaris, P., et al. 2019, ApJS, 243, 3

Li, J., Silverman, J. D., Merloni, A., et al. 2023, arXiv e-prints, arXiv:2302.12438

Li, Y.-R., Wang, J.-M., Songsheng, Y.-Y., et al. 2022, ApJ, 927, 58

Lidman, C., Tucker, B. E., Davis, T. M., et al. 2020, MNRAS, 496, 19

Lightman, A. P. & Zdziarski, A. A. 1987, ApJ, 319, 643

Lilly, S. J., Le Fevre, O., Crampton, D., Hammer, F., & Tresse, L. 1995, ApJ, 455, 50

Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70

Lin, D., Strader, J., Romanowsky, A. J., et al. 2020, ApJL, 892, L25

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., Buchner, J., Nandra, K., et al. 2022c, A&A, 661, A5

Liu, T., Buchner, J., Nandra, K., et al. 2022d, A&A, 661, A5

Liu, T., Merloni, A., Comparat, J., et al. 2022e, A&A, 661, A27

Liu, Z., Malyali, A., Krumpe, M., et al. 2023, A&A, 669, A75

Lodato, G. & Natarajan, P. 2006, MNRAS, 371, 1813

Loeb, A. & Rasio, F. A. 1994, ApJ, 432, 52

Lovisari, L., Ettori, S., Gaspari, M., & Giles, P. A. 2021, Universe, 7, 139

Lowell, B., Jacquemin-Ide, J., Tchekhovskoy, A., & Duncan, A. 2024, ApJ, 960, 82

Lyke, B. W., Higley, A. N., McLane, J. N., et al. 2020, ApJS, 250, 8

Lynden-Bell, D. 1969, Nature, 223, 690

Lyu, J., Rieke, G. H., & Shi, Y. 2017, ApJ, 835, 257

Madau, P. & Rees, M. J. 2001, ApJL, 551, L27

Magliocchetti, M. 2022, A&A Rev., 30, 6

Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285

Maiolino, R., Risaliti, G., Signorini, M., et al. 2024, arXiv e-prints, arXiv:2405.00504

Malavasi, N., Langer, M., Aghanim, N., Galárraga-Espinosa, D., & Gouin, C. 2022, A&A, 658, A113

Malyali, A., Liu, Z., Rau, A., et al. 2023, MNRAS, 520, 3549

Malzac, J. & Belmont, R. 2009, MNRAS, 392, 570

Marconi, A. & Hunt, L. K. 2003, ApJL, 589, L21

Marconi, A., Risaliti, G., Gilli, R., et al. 2004, MNRAS, 351, 169

Martín-Navarro, I. & Mezcua, M. 2018, ApJL, 855, L20

Massonneau, W., Volonteri, M., Dubois, Y., & Beckmann, R. S. 2023, A&A, 670, A180

Mateos, S., Alonso-Herrero, A., Carrera, F. J., et al. 2013, MNRAS, 434, 941

Matthews, T. A., Morgan, W. W., & Schmidt, M. 1964, ApJ, 140, 35

McCarthy, I. G., Schaye, J., Ponman, T. J., et al. 2010, MNRAS, 406, 822

McConnell, D., Hale, C. L., Lenc, E., et al. 2020, , 37, e048

McDonald, M., Benson, B. A., Vikhlinin, A., et al. 2014, ApJ, 794, 67

McNamara, B. R. & Nulsen, P. E. J. 2007, ARA&A, 45, 117

McNamara, B. R. & Nulsen, P. E. J. 2012, New Journal of Physics, 14, 055023

McNamara, B. R., Nulsen, P. E. J., Wise, M. W., et al. 2005, Nature, 433, 45

McNamara, B. R., Rafferty, D. A., Bîrzan, L., et al. 2006, ApJ, 648, 164

McNamara, B. R., Russell, H. R., Nulsen, P. E. J., et al. 2014, ApJ, 785, 44

Meenakshi, M., Mukherjee, D., Wagner, A. Y., et al. 2022, MNRAS, 516, 766

Meidinger, N., Andritschke, R., Dennerl, K., et al. 2020, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11444, Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray, ed. J.-W. A. den Herder, S. Nikzad, & K. Nakazawa, 114444O

Meier, D. L. 2002, New Astron. Rev., 46, 247

Menzel, M.-L., Merloni, A., Georgakakis, A., et al. 2016, MNRAS, 457, 110

Merloni, A. 2016, in Lecture Notes in Physics, Berlin Springer Verlag, ed. F. Haardt, V. Gorini, U. Moschella, A. Treves, & M. Colpi, Vol. 905, 101

Merloni, A., Alexander, D. A., Banerji, M., et al. 2019, The Messenger, 175, 42

Merloni, A. & Heinz, S. 2007, MNRAS, 381, 589

Merloni, A. & Heinz, S. 2008, MNRAS, 388, 1011

Merloni, A., Heinz, S., & di Matteo, T. 2003, MNRAS, 345, 1057

Merloni, A., Lamer, G., Liu, T., et al. 2024, A&A, 682, A34

Merloni, A., Predehl, P., Becker, W., et al. 2012, arXiv e-prints, arXiv:1209.3114

Mezcua, M. 2017, International Journal of Modern Physics D, 26, 1730021

Mezcua, M., Civano, F., Marchesi, S., et al. 2018, MNRAS, 478, 2576

Mezcua, M., Roberts, T. P., Lobanov, A. P., & Sutton, A. D. 2015, MNRAS, 448, 1893

Mezcua, M., Suh, H., & Civano, F. 2019, MNRAS, 488, 685

Michell, J. 1784, Philosophical Transactions of the Royal Society of London Series I, 74, 35

Miller, B. P., Gallo, E., Greene, J. E., et al. 2015, ApJ, 799, 98

Mineo, S., Gilfanov, M., & Sunyaev, R. 2012, MNRAS, 426, 1870

Mingo, B., Croston, J. H., Best, P. N., et al. 2022, MNRAS, 511, 3250

Mizumoto, M., Done, C., Tomaru, R., & Edwards, I. 2019, MNRAS, 489, 1152

Mobasher, B., Dahlen, T., Ferguson, H. C., et al. 2015, ApJ, 808, 101

Mohan, N. & Rafferty, D. 2015, PyBDSF: Python Blob Detection and Source Finder, Astrophysics Source Code Library, record ascl:1502.007

Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, ApJS, 225, 27

Mor, R. & Netzer, H. 2012, MNRAS, 420, 526

Morabito, L. K., Sweijen, F., Radcliffe, J. F., et al. 2022, MNRAS, 515, 5758

Morganti, R., Oosterloo, T., Oonk, J. B. R., Frieswijk, W., & Tadhunter, C. 2015, A&A, 580, A1

Morris, M. E., Wilcots, E., Hooper, E., & Heinz, S. 2022, AJ, 163, 280

Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121

Mostert, R. I. J., Duncan, K. J., Alegre, L., et al. 2022, A&A, 668, A28

Mountrichas, G., Masoura, V. A., Xilouris, E. M., et al. 2022, A&A, 661, A108

Moustakas, J., Coil, A. L., Aird, J., et al. 2013, ApJ, 767, 50

Moustakas, J., Lang, D., Dey, A., et al. 2023, ApJS, 269, 3

Mukherjee, D., Bicknell, G. V., Wagner, A. Y., Sutherland, R. S., & Silk, J. 2018, MNRAS, 479, 5544

Mullaney, J. R., Alexander, D. M., Goulding, A. D., & Hickox, R. C. 2011, MNRAS, 414, 1082

Mullaney, J. R., Daddi, E., Béthermin, M., et al. 2012, ApJL, 753, L30

Nandra, K., Barret, D., Barcons, X., et al. 2013, arXiv e-prints, arXiv:1306.2307

Nandra, K. & Pounds, K. A. 1994, MNRAS, 268, 405

Narayan, R. & Yi, I. 1994, ApJL, 428, L13

Narayan, R. & Yi, I. 1995, ApJ, 452, 710

Natarajan, P. 2021, MNRAS, 501, 1413

Natarajan, P., Pacucci, F., Ricarte, A., et al. 2024, ApJL, 960, L1

Nelson, D., Pillepich, A., Springel, V., et al. 2018, MNRAS, 475, 624

Nenkova, M., Sirocky, M. M., Nikutta, R., Ivezić, Ž., & Elitzur, M. 2008, ApJ, 685, 160

Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, ApJS, 208, 5

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJL, 660, L43

Nyland, K., Marvil, J., Wrobel, J. M., Young, L. M., & Zauderer, B. A. 2012, ApJ, 753, 103

Oei, M., Van Weeren, R., Rottgering, H., Stern, D., & Djorgovski, S. 2024, in American Astronomical Society Meeting Abstracts, Vol. 243, American Astronomical Society Meeting Abstracts #243, 425.08

Oppenheimer, J. R. & Snyder, H. 1939, Physical Review, 56, 455

Orosz, J. A., McClintock, J. E., Aufdenberg, J. P., et al. 2011, ApJ, 742, 84

O'Shea, T. M., Heinz, S., Soares-Furtado, M., Igo, Z., & Merloni, A. 2025, arXiv e-prints, arXiv:2510.26881

Osterbrock, D. E. 1989, Astrophysics of gaseous nebulae and active galactic nuclei (University Science Books)

Pacucci, F., Mezcua, M., & Regan, J. A. 2021, ApJ, 920, 134

Pacucci, F. & Narayan, R. 2024, ApJ, 976, 96

Padovani, P., Alexander, D. M., Assef, R. J., et al. 2017, A&A Rev., 25, 2

Panessa, F., Baldi, R. D., Laor, A., et al. 2019, Nature Astronomy, 3, 387

Papaderos, P., Guseva, N. G., Izotov, Y. I., & Fricke, K. J. 2008, A&A, 491, 113

Pardo, K., Goulding, A. D., Greene, J. E., et al. 2016, ApJ, 831, 203

Parkinson, D., Riemer-Sørensen, S., Blake, C., et al. 2012, Phys. Rev. D, 86, 103518

Partmann, C., Naab, T., Lahén, N., et al. 2025, MNRAS, 537, 956

Pasini, T., Brüggen, M., Hoang, D. N., et al. 2022, A&A, 661, A13

Pavlinsky, M., Tkachenko, A., Levin, V., et al. 2021, A&A, 650, A42

Perucho, M. 2019, Galaxies, 7, 70

Peterson, J. R., Kahn, S. M., Paerels, F. B. S., et al. 2004, in The Riddle of Cooling Flows in Galaxies and Clusters of galaxies, ed. T. Reiprich, J. Kempner, & N. Soker, 119

Petter, G. C., Hickox, R. C., Morabito, L. K., & Alexander, D. M. 2024, ApJ, 972, 184

Pierce, J. C. S., Tadhunter, C. N., Gordon, Y., et al. 2022, MNRAS, 510, 1163

Pineau, F. X., Derriere, S., Motch, C., et al. 2017, A&A, 597, A89

Popesso, P., Biviano, A., Bulbul, E., et al. 2023a, arXiv e-prints, arXiv:2302.08405

Popesso, P., Biviano, A., Bulbul, E., et al. 2024, MNRAS, 527, 895

Popesso, P., Concas, A., Cresci, G., et al. 2023b, MNRAS, 519, 1526

Pozdnyakov, L. A., Sobol, I. M., & Syunyaev, R. A. 1976, Soviet Astronomy Letters, 2, 55

Pozzetti, L., Bolzonella, M., Zucca, E., et al. 2010, A&A, 523, A13

Predehl, P., Andritschke, R., Arefiev, V., et al. 2021, A&A, 647, A1

Prescott, M., Mauch, T., Jarvis, M. J., et al. 2016, MNRAS, 457, 730

Prevot, M. L., Lequeux, J., Maurice, E., Prevot, L., & Rocca-Volmerange, B. 1984, A&A, 132, 389

Pucha, R., Juneau, S., Dey, A., et al. 2025, ApJ, 982, 10

Puchwein, E. & Springel, V. 2013, MNRAS, 428, 2966

Ramos Almeida, C., Bessiere, P. S., Tadhunter, C. N., et al. 2012, MNRAS, 419, 687

Ramos Almeida, C., Bischetti, M., García-Burillo, S., et al. 2022, A&A, 658, A155

Ramos Almeida, C. & Ricci, C. 2017, Nature Astronomy, 1, 679

Ramos Almeida, C., Tadhunter, C. N., Inskip, K. J., et al. 2011, MNRAS, 410, 1550

Ramsey, L. W., Adams, M. T., Barnes, T. G., et al. 1998, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 3352, Advanced Technology Optical/IR Telescopes VI, ed. L. M. Stepp, 34–42

Rantala, A. & Naab, T. 2025, MNRAS, 542, L78

Rantala, A., Naab, T., & Lahén, N. 2024, MNRAS, 531, 3770

Raouf, M., Shabala, S. S., Croton, D. J., Khosroshahi, H. G., & Bernyk, M. 2017, MNRAS, 471, 658

Reines, A. E., Condon, J. J., Darling, J., & Greene, J. E. 2020, ApJ, 888, 36

Reines, A. E., Greene, J. E., & Geha, M. 2013, ApJ, 775, 116

Reines, A. E. & Volonteri, M. 2015, ApJ, 813, 82

Reynolds, C. S. 2021, ARA&A, 59, 117

Ricarte, A. & Natarajan, P. 2018, MNRAS, 481, 3278

Ricarte, A., Tremmel, M., Natarajan, P., & Quinn, T. 2021, ApJL, 916, L18

Ricci, C., Ananna, T. T., Temple, M. J., et al. 2022, ApJ, 938, 67

Ricci, C., Ho, L. C., Fabian, A. C., et al. 2018, MNRAS, 480, 1819

Ricci, C., Trakhtenbrot, B., Koss, M. J., et al. 2017, Nature, 549, 488

Richards, G. T., Strauss, M. A., Fan, X., et al. 2006, AJ, 131, 2766

Rihtaršič, G., Biffi, V., Fabjan, D., & Dolag, K. 2024, A&A, 683, A57

Risaliti, G., Elvis, M., Fabbiano, G., et al. 2007, ApJL, 659, L111

Robotham, A. S. G., Bellstedt, S., Lagos, C. d. P., et al. 2020, MNRAS, 495, 905

Robotham, A. S. G., Norberg, P., Driver, S. P., et al. 2011, MNRAS, 416, 2640

Rodríguez Morales, V., Mezcua, M., Domínguez Sánchez, H., et al. 2025, A&A, 697, A235

Ross, R. R., Fabian, A. C., & Young, A. J. 1999, MNRAS, 306, 461

Ruffini, R. & Wheeler, J. A. 1971, Physics Today, 24, 30

Ruszkowski, M. & Pfrommer, C. 2023, A&A Rev., 31, 4

Rybicki, G. B. & Lightman, A. P. 1986, Radiative Processes in Astrophysics (Wiley-VCH)

Sabater, J., Best, P. N., Hardcastle, M. J., et al. 2019, A&A, 622, A17

Sabater, J., Best, P. N., Verdes-Montenegro, L., Leon, S., & Sulentic, J. 2013, in Highlights of Spanish Astrophysics VII, ed. J. C. Guirado, L. M. Lara, V. Quilis, & J. Gorgas, 461–461

Sacchi, A., Bogdán, Á., Chadayammuri, U., & Ricarte, A. 2024, ApJ, 974, 14

Saikia, D. J. 2022, Journal of Astrophysics and Astronomy, 43, 97

Saintonge, A., Catinella, B., Tacconi, L. J., et al. 2017, ApJS, 233, 22

Saintonge, A., Kauffmann, G., Wang, J., et al. 2011, MNRAS, 415, 61

Salehirad, S., Reines, A. E., & Molina, M. 2025, ApJ, 979, 26

Salpeter, E. E. 1964, ApJ, 140, 796

Salvato, M., Buchner, J., Budavári, T., et al. 2018, MNRAS, 473, 4937

Salvato, M., Wolf, J., Dwelly, T., et al. 2022a, A&A, 661, A3

Salvato, M., Wolf, J., Dwelly, T., et al. 2022b, A&A, 661, A3

Salvato, M., Wolf, J., Dwelly, T., et al. 2025, arXiv e-prints, arXiv:2509.02842

Sandage, A. 1965, ApJ, 141, 1560

Sanders, J. S., Bahar, Y. E., Bulbul, E., et al. 2025, A&A, 695, A160

Santini, P., Ferguson, H. C., Fontana, A., et al. 2015, ApJ, 801, 97

Saripalli, L., Hunstead, R. W., Subrahmanyan, R., & Boyce, E. 2005, AJ, 130, 896

Sartori, L. F., Schawinski, K., Treister, E., et al. 2015, MNRAS, 454, 3722

Saxena, A., Salvato, M., Roster, W., et al. 2024, A&A, 690, A365

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Schmidt, M. 1963, Nature, 197, 1040

Schwarzschild, K. 1916, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 189

Seppi, R., Comparat, J., Bulbul, E., et al. 2022, A&A, 665, A78

Seyfert, C. K. 1943, ApJ, 97, 28

Shabala, S. S. 2018, MNRAS, 478, 5074

Shakura, N. I. & Sunyaev, R. A. 1973, A&A, 24, 337

Shapiro, S. L., Lightman, A. P., & Eardley, D. M. 1976, ApJ, 204, 187

Shectman, S. A., Landy, S. D., Oemler, A., et al. 1996, ApJ, 470, 172

Shi, Y., Yu, X., Mao, S., et al. 2021, MNRAS, 507, 2423

Shields, G. A. 1978, Nature, 272, 706

Shimwell, T. W., Hardcastle, M. J., Tasse, C., et al. 2022, A&A, 659, A1

Shimwell, T. W., Röttgering, H. J. A., Best, P. N., et al. 2017, A&A, 598, A104

Shimwell, T. W., Tasse, C., Hardcastle, M. J., et al. 2019, A&A, 622, A1

Sijacki, D., Springel, V., Di Matteo, T., & Hernquist, L. 2007, MNRAS, 380, 877

Sikora, M. & Begelman, M. C. 2013, ApJL, 764, L24

Simmonds, C., Buchner, J., Salvato, M., Hsu, L. T., & Bauer, F. E. 2018, A&A, 618, A66

Sądowski, A., Narayan, R., McKinney, J. C., & Tchekhovskoy, A. 2014, MNRAS, 439, 503

Slyz, A., Devriendt, J., Jarvis, M., Dubois, Y., & Pichon, C. 2015, in The Many Facets of Extragalactic Radio Surveys: Towards New Scientific Challenges, 52

Smith, D. J. B., Best, P. N., Duncan, K. J., et al. 2016, in SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, ed. C. Reylé, J. Richard, L. Cambrésy, M. Deleuil, E. Pécontal, L. Tresse, & I. Vauglin, 271–280

Smith, D. J. B., Haskell, P., Gürkan, G., et al. 2021, A&A, 648, A6

Smolčić, V., Finoguenov, A., Zamorani, G., et al. 2011, MNRAS, 416, L31

Smolčić, V., Novak, M., Delvecchio, I., et al. 2017, A&A, 602, A6

Smolčić, V., Schinnerer, E., Scodeggio, M., et al. 2008, ApJS, 177, 14

Smolčić, V., Zamorani, G., Schinnerer, E., et al. 2009, ApJ, 696, 24

Soltan, A. 1982, MNRAS, 200, 115

Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. 2008, MNRAS, 391, 481

Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15

Springel, V., Di Matteo, T., & Hernquist, L. 2005, ApJL, 620, L79

Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676

Stone, N. C., Küpper, A. H. W., & Ostriker, J. P. 2017, MNRAS, 467, 4180

Storey-Fisher, K., Hogg, D. W., Rix, H.-W., et al. 2024, ApJ, 964, 69

Strüder, L., Briel, U., Dennerl, K., et al. 2001, A&A, 365, L18

Sturm, E., González-Alfonso, E., Veilleux, S., et al. 2011, ApJL, 733, L16

Sun, M. 2009, ApJ, 704, 1586

Sunyaev, R., Arefiev, V., Babyshkin, V., et al. 2021, A&A, 656, A132

Svoboda, J., Guainazzi, M., & Merloni, A. 2017, in AAS/High Energy Astrophysics Division, Vol. 16, AAS/High Energy Astrophysics Division #16, 405.01

Sweijen, F., van Weeren, R. J., Röttgering, H. J. A., et al. 2022, Nature Astronomy, 6, 350

Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2018, ApJ, 853, 179

Tagliacozzo, D., Marinucci, A., Ursini, F., et al. 2023, MNRAS, 525, 4735

Talbot, R. Y., Bourne, M. A., & Sijacki, D. 2021, MNRAS, 504, 3619

Tanaka, Y., Nandra, K., Fabian, A. C., et al. 1995, Nature, 375, 659

Tanner, R. & Weaver, K. A. 2022, AJ, 163, 134

Tasca, L. A. M., Le Fevre, O., Ribeiro, B., et al. 2018, VizieR Online Data Catalog: VIMOS Ultra Deep Survey (VUDS) DR1 (Tasca+, 2017), VizieR On-line Data Catalog: J/A+A/600/A110. Originally published in: 2017A&A...600A.110T

Tasse, C., Shimwell, T., Hardcastle, M. J., et al. 2021, A&A, 648, A1

Tchekhovskoy, A. 2015, in Astrophysics and Space Science Library, Vol. 414, The Formation and Disruption of Black Hole Jets, ed. I. Contopoulos, D. Gabuzda, & N. Kylafis, 45

Thomas, N., Davé, R., Jarvis, M. J., & Anglés-Alcázar, D. 2021, MNRAS, 503, 3492

Thorne, J. E., Robotham, A. S. G., Davies, L. J. M., et al. 2022, MNRAS, 509, 4940

Timmerman, R., van Weeren, R. J., Botteon, A., et al. 2022, A&A, 668, A65

Tombesi, F., Cappi, M., Reeves, J. N., et al. 2010, A&A, 521, A57

Torbaniuk, O., Paolillo, M., D'Abrusco, R., et al. 2024, MNRAS, 527, 12091

Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898

Turner, R. J. & Shabala, S. S. 2015, ApJ, 806, 59

Turner, R. J., Yates-Jones, P. M., Shabala, S. S., Quici, B., & Stewart, G. S. C. 2023, MNRAS, 518, 945

Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, ApJ, 598, 886

Urry, C. M. & Padovani, P. 1995, PASP, 107, 803

van den Bosch, R. C. E. 2016, ApJ, 831, 134 van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28

van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2

Vasudevan, R. V. & Fabian, A. C. 2009, MNRAS, 392, 1124

Vaughan, S., Edelson, R., Warwick, R. S., & Uttley, P. 2003, MNRAS, 345, 1271

Véron-Cetty, M. P. & Véron, P. 2010, A&A, 518, A10

Vikhlinin, A., Kravtsov, A., Forman, W., et al. 2006, ApJ, 640, 691

Vladutescu-Zopp, S., Biffi, V., & Dolag, K. 2025, A&A, 695, A2

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, Nature, 509, 177

Voit, G. M., Donahue, M., Bryan, G. L., & McDonald, M. 2015, Nature, 519, 203

Volonteri, M., Dubois, Y., Pichon, C., & Devriendt, J. 2016, MNRAS, 460, 2979

Volonteri, M., Lodato, G., & Natarajan, P. 2008, MNRAS, 383, 1079

Waddell, S. G. H., Nandra, K., Buchner, J., et al. 2024, A&A, 690, A132

Wagner, A. Y., Bicknell, G. V., & Umemura, M. 2012, ApJ, 757, 136

Walton, D., Mackenzie, A., Gully, H., et al. 2022, in AAS/High Energy Astrophysics Division, Vol. 19, AAS/High Energy Astrophysics Division, 110.86

Wang, L., Tozzi, P., Yu, H., Gaspari, M., & Ettori, S. 2023, A&A, 674, A102

Ward, S. R., Costa, T., Harrison, C. M., & Mainieri, V. 2024, MNRAS, 533, 1733

Ward, S. R., Harrison, C. M., Costa, T., & Mainieri, V. 2022, MNRAS, 514, 2936

Wasleske, E. J. & Baldassare, V. F. 2024, ApJ, 971, 68

Webster, B. L. & Murdin, P. 1972, Nature, 235, 37

Weinberger, R., Springel, V., Hernquist, L., et al. 2017, MNRAS, 465, 3291

Weinberger, R., Springel, V., Pakmor, R., et al. 2018, MNRAS, 479, 4056

Williams, W. L., Hardcastle, M. J., Best, P. N., et al. 2019, A&A, 622, A2

Willmer, C. N. A. 2018, ApJS, 236, 47

Willott, C. J., Rawlings, S., Blundell, K. M., & Lacy, M. 1999, MNRAS, 309, 1017

Wilms, J., Allen, A., & McCray, R. 2000, ApJ, 542, 914

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868

Wu, Z. & Ho, L. C. 2025, ApJ, 985, 197

Wykes, S., Croston, J. H., Hardcastle, M. J., et al. 2013, A&A, 558, A19

Xrism Collaboration, Audard, M., Awaki, H., et al. 2025, ApJL, 982, L5

Yee, H. K. C., Morris, S. L., Lin, H., et al. 2000, ApJS, 129, 475

Yeung, M. C. H., Ponti, G., Freyberg, M. J., et al. 2024, A&A, 690, A399

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

Yuan, F. & Narayan, R. 2014, ARA&A, 52, 529

Yue, M., Eilers, A.-C., Ananna, T. T., et al. 2024, arXiv e-prints, arXiv:2404.13290

Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, ApJ, 554, 803

Zakamska, N. L. & Greene, J. E. 2014, MNRAS, 442, 784

Zaw, I., Rosenthal, M. J., Katkov, I. Y., et al. 2020, ApJ, 897, 111

Zeimann, G. R., Debski, M. H., Schneider, D. P., et al. 2024, ApJ, 966, 14

Zenteno, A., Kluge, M., Kharkrang, R., et al. 2025, arXiv e-prints, arXiv:2503.21066

Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, ApJ, 633, 791

Zhou, R., Ferraro, S., White, M., et al. 2023, , 2023, 097

Zhou, R., Newman, J. A., Mao, Y.-Y., et al. 2021, MNRAS, 501, 3309

Zhuravleva, I., Churazov, E., Schekochihin, A. A., et al. 2014, Nature, 515, 85

Zou, F., Yu, Z., Brandt, W. N., et al. 2024, ApJ, 964, 183

Zou, H., Gao, J., Zhou, X., & Kong, X. 2019, ApJS, 242, 8

Acknowledgements

It's over! It's done! And it wouldn't have been possible without the many wonderful people who have contributed to making these past four years as incredible as they were:

To Dr Andrea Merloni, my supervisor: after every meeting I came out at least 100 times wiser thanks to your boundless knowledge about astrophysics and life, thank you for being a constant for me during this variable, chaotic PhD journey and for creating so many fantastic opportunities for me to progress as a researcher. To Prof Kirpal Nandra: for your great scientific insights and for opening my eyes to the world of managing numerous multi-wavelength ground- and space-based observatories. To Dr Johannes Buchner: for always being available to answer my (many) detailed questions and for teaching me that no problem is too complex for some quick sampling or simulation. To Dr Mara Salvato: for your never-ending support, to me, to the other students in the High Energy Group (HEG) and the whole institute, to be honest! Also for teaching me that often the nitty-gritty work is the most important for scientific success. To the HEG: for all the nice Ringberg retreats, bowlings, Christmas dinners together and for welcoming me to the group.

To my housemates: Ana, for being a little ray of sunshine, for our many episodes and reuniones de baño, I appreciate you a lot; Jelke, for your unmatched boardgame knowledge, masterchef skills and being the protective Viking of the house (even though we sometimes lived on different time zones); Fulvio, for teaching me so much about music and being strong in the face of difficulties (like our struggles in beating the Wii tanks game and playing blindfolded Mario Kart during those many Covid times). And Matti, well for being Matti, you are a unique human, with the kindest of hearts and the ability to bring logic to all emotions, we started and finished together and my time here wouldn't have been the same without you! I also cannot forget the honorary members of the house, the many beautiful people who have come to visit us ... and taken care of our plants (thank you, Andri). It has been one of my greatest sources to happiness over these years to come home after a long day and know that with you all we will always laugh, do something silly, invent a new game, cook something tasty or watch something all huddled together on our couch. I will really miss our little DH, our family, our inside jokes, the squirrels, but I am certain that we will continue our traditions in the future!

To my amazing office mates, Sophia and Soumya (and later Luisa, Ania, Tianying): I have been so lucky to have these inspirational, strong women to share my day to day with, fixing our silly coding or admin bugs, sharing snacks and life chats over a cup of tea. Thank you Sophia and Soumya for your beautiful aura that brightens everyone's day, the delicious lunches/dinners and our crazy conferences together. I am excited for the reunions we will have around the world, hopefully very soon!

To the many HEG and IMPRS friends over the years, especially the Carbonaras, for the wholesome lunches and sunny breaks; the crazy WAW beers with many fooseball and ping pong matches; the hikes, the infinite hours spent at the Englischer Garten and many Helles for sures. One of my fondest memories will remain the Inter-Institute Sports Week that we kick-started with Fabi, Matteo G., Benny, and William that brought over 100 sports enthusiasts together. Also the Girlsday and many repping duties (cakes and parties) with an amazing team (Soumya, Fulvio, Cata, Pietro, Capucine, Andrea and many others). To Luis, for having been connected by the filaments of the universe. To the many friends I have met during the incredible conferences around the world that I was lucky enough to attend: Scott (for learning together that the best science is done with wine and jacuzzis in Umbria); Emmy, Roland, Christian and the LOFAR crew (for escaping jellyfish together in Tropea and actually teaching me about radio astronomy), the Iceland gang with the 10 Euro beers and hilarious card games, Valencia gang playing make-shift spikeball on the beach and many more memories and people that I could never do justice here. To Martin, for being my physics partner in crime from Durham Day 1 (looking forward to our future Igo & Mosny paper!). To my Hungarian buddy Géza, for all the rakott krumpli and lángos nights, pumpkin carving and 'hydration', as well as to my Hungarian friends back home (Odett, Dávid, András és Tamás) and BFF Mari on the other side of the world, for lifelong memories, adventures and laughs. An mein Floorball-Team: für die mega spaßigen Turniere überall in Deutschland, die vielen Schwedenkreuze, die legendären UNO- und Schafkopf-Spiele und den besten Crashkurs in deutschem Slang! To the OG ESAC X-ray crew, especially Michael, Gabi, Norbert, for believing in me since before I even knew I wanted to do a PhD. ¡Nos vemos muy pronto en Madrid!

To my best friends, Emma and Babette: I cannot even begin to explain the gratitude I feel to have had your unwavering support and friendship throughout these past 10 years. Thank you for all the cultural exchanges, the brunches and dinners in Paris (and trying to teach strangers the proper ways to take photos), the deep convos and walks in nature, but most of all for always managing to make me laugh and for being such positive, thoughtful, kind-hearted individuals.

And last, but certainly not least, to my family: Családomnak, Anyának és Apának, amiért hittetek bennem, és feltétel nélküli szeretettel és támogatással álltatok mindig mellettem. Nélkületek ez a doktori disszertáció nem jöhetett volna létre (és persze a sok nasi nélkül sem, amit otthon, a cikkek írása közben biztosítottatok). Köszönöm szeretett Nagyszüleimnek is, akik biztosan fentről figyelnek és szurkolnak nekem a csillagok közül. És a cicámnak, Minounak, aki annyi örömet hozott az életembe az elmúlt négy, sajnos túl rövid év alatt. Szívem legmélyéről szeretlek benneteket.

To all the readers who made it this far, I would like to leave you with a piece of advice and some food for thought: Common sense is not common.

NB: No coffee was consumed in the making of this manuscript (even after being surrounded by Italians).