UNICA

UNIVERSITÀ
DEGLI STUDI
DI CAGLIARI

Ph.D. DEGREE IN
Mathematics and Computer Science

Cycle XXXVII

TITLE OF THE Ph.D. THESIS

Mathematical and Deep Learning Methods in High-Energy Physics: From Temporal Variability in M82 ULXs to Spatial Features in Planck Maps

Scientific Disciplinary Sector(s)

MAT/08

Ph.D. Student:	Hamza El Byad
Supervisor	Giuseppe Rodriguez
Co-Supervisor	Matteo Bachetti

Final exam. Academic Year 2023/2024
Thesis defence session: July 2025

Ringraziamenti

Desidero esprimere la mia più profonda gratitudine ai miei supervisori, il cui sostegno e la cui guida costante sono stati fondamentali durante tutto il mio percorso di dottorato. In particolare, vorrei rivolgere un sincero ringraziamento al Dott. Matteo Bachetti, la cui eccezionale competenza tecnica e il cui continuo incoraggiamento hanno costituito le basi su cui si è sviluppata la mia ricerca. I consigli illuminanti e il supporto costante del Dott. Bachetti mi hanno permesso di superare numerose sfide, esercitando una profonda influenza sulla mia crescita sia personale che professionale.

Sono ugualmente riconoscente al mio supervisore, il Prof. Giuseppe Rodriguez, la cui preziosa guida, il cui costante supporto e la cui grande esperienza sono stati essenziali per affrontare le complessità che caratterizzano il percorso di un dottorato. La dedizione costante e la saggezza del Prof. Rodriguez hanno notevolmente arricchito la mia esperienza accademica e migliorato la qualità della mia ricerca.

Desidero esprimere la mia più profonda gratitudine a tutti coloro che mi hanno supportato durante il percorso di questo dottorato, Maura Pilia, A. Marcos, S. Migliari, Eleonora, V. Piga, Delphine e Silvia Columbu.

Un ringraziamento speciale va all'Osservatorio Astronomico di Cagliari per avermi generosamente concesso l'utilizzo delle proprie strutture e risorse, senza le quali gran parte di questa ricerca non sarebbe stata possibile.

Ringrazio inoltre il Dipartimento di Matematica e Informatica e tutto il personale accademico e amministrativo del Dipartimento per la loro disponibilità e il loro prezioso sostegno durante il mio percorso di studi. Il loro contributo collettivo ha facilitato significativamente lo svolgimento del mio lavoro.

Infine, esprimo il mio più sincero apprezzamento verso la mia famiglia, i miei amici e tutti coloro il cui incoraggiamento, comprensione e supporto emotivo mi hanno permesso di rimanere motivato e concentrato. Senza la loro pazienza e la loro fiducia in me, il completamento di questa tesi non sarebbe stato possibile.

Contents

List of Figures

List of Tables

Introduction and Thesis Structure

Modern Astronomy gathers data from a plethora of instruments from the ground and from space, spanning the entire electromagnetic spectrum and even some non-electromagnetic carriers of information, such as neutrinos and gravitational waves. The quality and quantity of data is rapidly increasing over time, but analysis techniques are often lagging behind in terms of fully taking advantage of this complexity. On the one hand, this is expected: popular analysis techniques receive continuous algorithmic development, making the analysis extremely fast. This is the case of the Fast Fourier Transform, which revolutionized frequency analysis. Today, the power spectral density is still routinely modeled using periodograms calculated through an FFT, despite its very well-known limitations in terms of spectral resolution and rather strict assumptions. Renouncing the FFT means increasing the timescales of analysis by many orders of magnitude. One class of analysis tools that have received considerable development is of course Machine Learning, a portmanteau for many different techniques that span from classic algebraic manipulation to algorithms imitating in some way the functioning of the brain (e.g. neural networks). This thesis addresses the challenge of extracting and characterizing faint signals from the wealth of modern astrophysical data, focusing in particular on two major problems in observational astronomy. On one hand, we investigate the timing properties of cosmic X-ray sources, especially ultraluminous X-ray sources (ULXs) in starburst galaxies, aiming to characterize quasi-periodic variability that reflects the physics of accretion onto compact objects. On the other hand, we develop a machine-learning method to detect and classify sources in wide-field infrared survey maps, where instrumental noise and foregrounds often obscure faint signals. In both cases, the goal is to move beyond traditional techniques—such as naive Fourier analysis of evenly sampled data or simple thresholding of images—to apply advanced statistical cleaning, bootstrap resampling, and neural-network algorithms that leverage the high information content of modern observations. By combining rigorous time-series analysis with contemporary deep learning tools, this work seeks to improve the reliability of signal detection in challenging astrophysical datasets.

This thesis is structured as follows :

1. **Chapter 1 – Astrophysical Background**: Reviews the physics of accreting X-ray binaries and the properties of ULXs. It examines how rapid X-ray variability and QPOs arise in these systems and their implications for accretion physics. The chapter includes a case study of the ULXs in the galaxy M82 to illustrate the observational context and challenges.

2. **Chapter 2 – Methodological Foundations**: Introduces the main data analysis techniques used in this work. Time-series methods (Fourier transforms, peri-

odograms, and treatment of gaps or uneven sampling) are covered, followed by statistical tools such as maximum-likelihood PSD fitting and Monte Carlo/ bootstrap uncertainty estimation. The chapter also presents the machine-learning concepts relevant to this thesis, especially convolutional neural networks and training strategies for source detection.

3. **Chapter 3 – Cleaning and Bootstrapping for QPO Analysis**: Describes the application of data cleaning and bootstrap methods to X-ray timing data. We develop a pipeline that identifies and removes contaminated intervals from X-ray light curves (for example, using Good-Time-Interval filtering) and then uses a parametric bootstrap to assess the significance of candidate QPO signals. This approach is applied to NuSTAR observations of M82 ULXs, demonstrating improved detection of known QPOs and exploration of weak timing features that were previously ambiguous.

4. **Chapter 4 – Neural Networks for Infrared Source Detection**: Covers the development of a deep learning model to find sources in an infrared survey. We design and train a convolutional neural network on a mix of labeled and unlabeled image patches to classify the presence of astrophysical sources. The chapter compares the CNN-based detection performance to traditional methods, showing how the neural network can identify faint or blended sources that simpler algorithms might miss.

5. **Chapter 5 – Regularized Spectral Estimation**: Introduces a novel regularization approach to estimate power spectra from unevenly-sampled time series. We formulate spectral analysis as an inverse problem with a smoothness constraint on the PSD. After deriving the mathematical framework, the method is tested on simulated and real uneven datasets. The results demonstrate that the regularized approach yields higher-resolution PSD estimates and mitigates biases and variance that affect standard periodogram techniques.

Chapter 1

Introduction

1.1 Compact Objects and X-ray Binaries

Astrophysical compact objects are the dense, evolved remnants of stars that have exhausted their nuclear fuel. Primary among these are neutron stars (NSs) and black holes (BHs), which originate from the death of medium- to high-mass stars. Neutron stars were first proposed by Baade & Zwicky (1934) as the collapsed cores of supernova explosions, packing more mass than the Sun into a sphere only $\sim 10 kmacross$. Decades later, this prediction was dramatically confirmed with the discovery of pulsars – rapidly pulsating radio sources – beginning with the landmark detection of a 1.337-second pulsating signal by Jocelyn Bell Burnell in 1967 (Hewish et al., 1968). The recognition that pulsars are rotating neutron stars provided the first direct evidence of NSs as a new class of compact star. Black holes, on the other hand, represent an even more extreme endpoint of stellar evolution. If the collapsing core of a massive star exceeds the neutron star mass limit (around 2–3 $M\odot$), no known force can halt the collapse, and a black hole is formed. The concept of an object so compact that not even light can escape its gravity dates back to ideas by John Michell and Pierre-Simon Laplace in the 18th century, but black holes entered modern science through Einstein's theory of general relativity (Einstein, 1916) and Schwarzschild's solution for curved spacetime around a point mass (Schwarzschild, 1916). Astronomers obtained strong empirical evidence for stellar-mass black holes in the early 1970s with X-ray observations of binary systems like Cygnus X-1, where a $\sim 15 M\odot$ compact object accretes matter from a blue supergiant companion and exhibits properties best explained by a black hole (Bolton, 1972; Webster & Murdin, 1972). Together, neutron stars and black holes (as well as white dwarfs, the compact cores of less massive stars) constitute the family of compact objects that are key players in high-energy astrophysics.

1.1.1 X-ray Binaries

Compact objects become extraordinary X-ray emitters when they reside in binary systems with a stellar companion. In an XRB, a normal star transfers mass onto a dense compact companion, converting gravitational energy into intense X-ray emission (Lewin et al., 1997). Instead of free-falling, the gas forms an accretion disk, where frictional forces heat the material to millions of degrees, causing it to radiate predominantly in X-rays (Shakura & Sunyaev, 1973; Done et al., 2007). This process is highly efficient at

converting mass into energy—far surpassing nuclear fusion—which is why XRBs can rival entire galaxies in X-ray luminosity. Beyond their brightness, X-ray binaries serve as astrophysical laboratories, revealing insights into strong gravity, dense matter, and plasma physics (As depicted in Figure 1.1). Precise pulsar timing in binary systems has enabled tests of general relativity and contributed to the detection of gravitational waves from neutron star mergers (Hulse & Taylor, 1975; Abbott et al., 2016). The turbulent flow of gas within the disk amplifies the radiant output, especially in systems with neutron stars or black holes Frank et al. (2002).

X-ray binaries are commonly classified by the mass of the donor star:

- **Low-Mass X-ray Binaries (LMXBs)**: The companion is a low-mass star ($\lesssim 1M_\odot$). filling its Roche lobe. Many LMXBs are transient systems: they undergo dramatic outbursts separated by long quiescent intervals. During an outburst, the X-ray luminosity can rise by factors of 10^4–10^5 (Tetarenko et al., 2016). Neutron-star LMXBs often exhibit Type I X-ray bursts (thermonuclear flashes on the NS surface), while black-hole LMXBs reach comparable peak luminosities near the Eddington limit (Galloway et al., 2020). Despite their eruptive behaviour, both classes share the same fundamental engine: gravity-powered accretion onto a compact object. Despite differences in their observational features, both types of LMXBs are powered by the same fundamental process: accretion driven by gravity. Their behaviour and evolution differ markedly from high-mass X-ray binaries, as discussed by Done et al. (2007), and this contrast provides valuable context for understanding ultraluminous X-ray sources (ULXs). Once believed to require intermediate-mass black holes, many ULXs are now interpreted as stellar-mass systems accreting at extreme, and possibly super-Eddington, rates (Kaaret et al., 2017a).

- **High-Mass X-ray Binaries (HMXBs)**: The companion is a massive O/B-type star ($\gtrsim 10M_\odot$). Mass transfer often proceeds via the donor's intense stellar wind or Roche-lobe overflow if the star expands (illustrated in Figure 1.2). Figure 1.2 illustrates the Roche-lobe geometry of a binary star system, showing the Lagrange points (L1–L5) and how mass transfer occurs through the inner L1 point. Many HMXBs show persistent X-ray emission due to a strong, steady mass supply. Systems with strongly magnetized neutron stars can display coherent X-ray pulsations (e.g., the pulsar in Vela X-1; see (Kretschmar et al., 2019) .

Another important distinction is whether the compact object is a black hole or neutron star, as this affects the observed X-ray properties. Black hole XRBs often reach higher peak luminosities (up to the Eddington limit of a $\sim$ BH $L_X \sim 10^{39}$ ergs^{-1}) and exhibit characteristic spectral/timing states (see next section). Neutron star XRBs can be equally luminous but also display phenomena absent in BHs, such as Type I X-ray bursts (thermonuclear flashes on the NS surface) and coherent pulsations (if the NS's magnetic field channels the inflow to the magnetic poles). Despite these differences, both classes share the same fundamental engine: gravity-powered accretion onto a compact object. In either case, the binary nature is crucial — the presence of a donor star replenishing the disk distinguishes XRBs from isolated compact objects. The formation and evolution of XRBs are intimately tied to stellar evolution and binary interactions (Tutukov & Yungelson, 1973; Tauris & van den Heuvel, 2006). Notably, the

abundance of X-ray binaries in a galaxy correlates strongly with star formation activity. Galaxies undergoing high rates of recent star formation (producing many massive stars that end their lives as NSs and BHs) tend to host numerous XRBs, especially HMXBs. Indeed, the collective X-ray luminosity of HMXB populations scales roughly linearly with a galaxy's star formation rate. Starburst galaxies – those undergoing short bursts of intense star formation – are therefore expected to be rich in X-ray binaries and related high-energy phenomena.

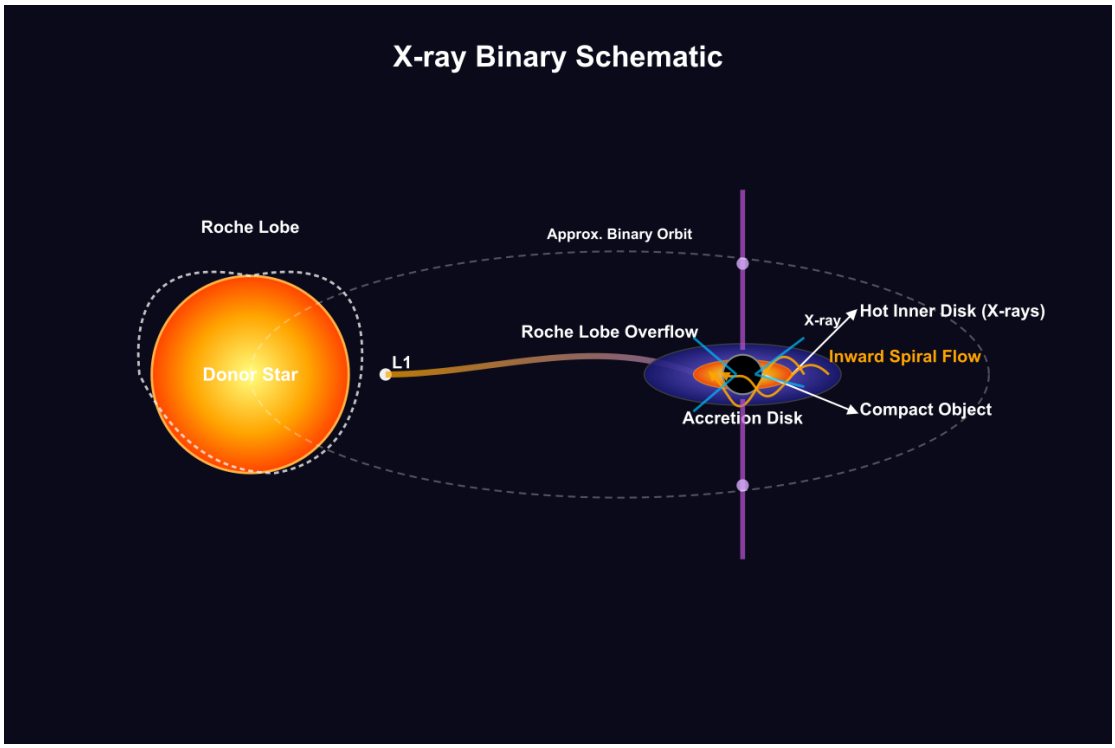


Figure 1.1: Representation of a typical X-ray binary configuration. The companion star (left portion) extends beyond its gravitational boundary, causing material flow toward the dense stellar remnant (right) through boundary overflow. The transferred stellar matter forms a rotating collection around the compact remnant, featuring both an outer cooler region and an intensely heated central area that produces energetic radiation (illustrated by light-blue beams). The central region's extreme temperature generates substantial high-energy output, while the inward movement patterns (orange markers) show how rotational energy diminishes during accretion. Potential high-velocity matter streams may project perpendicular to the main plane (purple elements). A dotted elliptical line indicates the approximate path these two bodies follow in their mutual orbit.

1.2 Accretion Physics

Accretion onto a compact object lies at the heart of X-ray binary emission. Gas spiraling inward releases gravitational potential energy with high efficiency, producing multi-component X-ray spectra. Key components of the emission typically include: (1)

thermal emission from an accretion disk, (2) a high-energy Comptonization component from a hot electron corona, and (3) disk reflection features, often with relativistic effects near the compact object.

Accretion Disc

In soft states, the disk is thin and optically thick, with angular momentum transported by magneto-rotational turbulence rather than the traditional α-viscosity. Recent general-relativistic magnetohydrodynamic (GR-MHD) simulations confirm that, for sub-Eddington accretion rates, the flow collapses into a Keplerian sheet with $H/R \ll 1$, maintaining efficient angular-momentum transport down to the innermost stable circular orbit (ISCO) and placing the inner edge of a $10\,M_\odot$ disc at ~ 90 km—consistent with strong-gravity predictions (Dihingia & Fendt, 2025; Penna et al., 2010; Liska et al., 2019). Long-term monitoring of sources such as LMC X-3 and GX 339-4 supports this picture: multi-temperature black-body fits yield inner radii and temperature profiles $T(r) \propto r^{-3/4}$ that agree with the steady-state disc model, even when irradiation inflates a tenuous atmosphere (Steiner et al., 2010). Radiative MHD simulations show that a thin disc can coexist with a luminous corona, reproducing the observed hard X-ray tail without erasing the thermal hump (Jiang et al., 2014). Recent IXPE polarimetry of Cyg X-1 adds an independent constraint: the observed polarization swing requires relativistic light bending and a beamed corona near the ISCO (Krawczynski et al., 2022). Taken together, these theoretical and observational advances reinforce the classical thin disc as the dominant structure in the soft states of accreting compact objects.

In black hole X-ray binaries (BHXRBs) during their high/soft state, thermal radiation from the accretion disk dominates the soft X-ray band, with characteristic color temperatures around $kT \sim 1$–2 keV for black holes of stellar mass Remillard & McClintock (2006); Done et al. (2007). In NS XRBs, the disk can be even hotter (several keV) since the flow extends down to the neutron star surface, a deeper gravitational potential (Lin et al., 2007; Cackett et al., 2008) . During such "soft states," the disk contributes the majority of the X-ray luminosity. the disk luminosity roughly follows $L \propto T^4$ (as expected for blackbody emission from a constant-area radiator), which suggests the inner disk radius remains fairly stable (near the ISCO) as the luminosity varies Gierliński & Done (2004); Steiner et al. (2010). This behavior supports the thin disk model and allows estimates of the inner disk radius (and thus BH spin or NS radius) by fitting the observed disk spectrum (Miller et al., 2009; McClintock et al., 2014).

At lower accretion rates (the hard spectral state in BHXRBs), the disk cools and may terminate at greater distances from the compact object, contributing primarily in ultraviolet or extremely soft X-ray bands Tomsick et al. (2009); Plant et al. (2015). In these hard states, the disk's contribution to X-ray emission diminishes considerably; instead, other components dominate the spectrum Remillard & McClintock (2006) (discussed in subsequent sections). Nonetheless, even during hard states, sensitive instruments can frequently detect the presence of a faint, cool accretion disk Miller et al. (2006); Reis et al. (2010).).

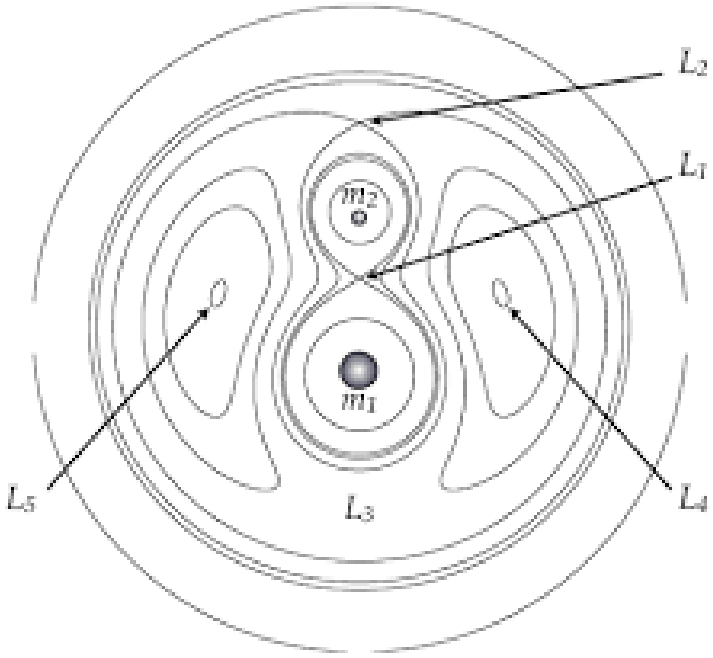


Figure 1.2: Schematic diagram of a Roche lobe in a binary star system. The diagram shows two stars, labeled M1 and M2, with their Roche lobes, which are the regions where material is gravitationally bound to each star. The five Lagrangian points (L1 to L5) are equilibrium points in the rotating frame of the binary system. Only L1 allows direct mass transfer between the stars, as it is the lowest-energy pathway for material to cross from one Roche lobe to the other. This illustration helps explain mass transfer and accretion in binary systems (Frank et al., 2002).

Comptonisation Component

Inverse Compton scattering generates another significant spectral feature in X-ray binaries when disk photons interact with an energetic, diffuse plasma known as the "corona." This distinctive region contains high-temperature electrons (characteristic temperature $kT_e \sim 30$–200 keV) with low to moderate optical thickness $\tau \lesssim 1$) situated near the compact object—either suspended above the disk or embedded within the innermost accretion flow Zdziarski et al. (1996). Soft photons emitted by the disk penetrate this corona and undergo energy augmentation through interactions with these energetic electrons (the Compton process). Individual scattering events boost a photon's energy by approximately

$$\Delta E \approx \frac{4kT_e}{m_e c^2}$$

on average, continuing until the photon escapes or the electron population experiences substantial energy depletion, which leads to a hard X-ray tail (a power-law spectrum) Titarchuk (1994); Coppi (1999).

The Comptonization (or "power-law") component typically dominates the spectrum in low-accretion-rate conditions (the hard state), and its shape encodes information about the corona's optical depth and temperature. For instance, a larger disk truncation (i.e., a hotter, more extended corona) often leads to a harder spectrum Done et al. (2007). NS XRBs also have Comptonized spectra; however, the presence of the NS surface (and boundary layer) adds additional seed photons and can lead to somewhat softer high-energy cutoffs Lin et al. (2007). Fundamentally, the corona and its Comptonization emission provide insight into the geometry of the inner accretion flow, since a smaller, more centrally concentrated disk (in the hard state) likely corresponds to a larger, possibly more spherical corona.

Disc Reflection Component

A fraction of the coronal hard X-ray photons inevitably shines back onto the accretion disk, triggering various interactions: photoelectric absorption (followed by fluorescence), Compton scattering (which can redirect and reduce photon energies), and, at the highest energies, pair production. These phenomena combine to generate a reflection spectrum that accompanies the direct power-law emission Fabian & Ross (2010). Figure 1.3 illustrates the disk–corona geometry in an X-ray binary: soft photons from the accretion disk (red) are Compton-upscattered by a hot corona (yellow), and some hard photons (green) reflect off the disk producing the iron line and Compton hump.

Hallmark features of reflection include fluorescent emission lines (especially the iron Kα line around $6.4-6.9keV$, depending on ionization state) and a broad hump at ∼20–30 keV (the "Compton hump") due to down-scattered photons Magdziarz & Zdziarski (1995). Iron is particularly important: neutral Fe yields a line at 6.4 keV, whereas highly ionized Fe XXVI can produce lines up to ∼6.97 keV Miller (2007). In partially ionized disks, multiple lines blend into a broad Fe K feature García et al. (2013).

Though the reflection emission is weaker than the direct continuum, detecting and modeling features like the Fe Kα line and Compton hump is essential for constraining disk properties García et al. (2015). Figure 1.3 illustrates the geometry of disk

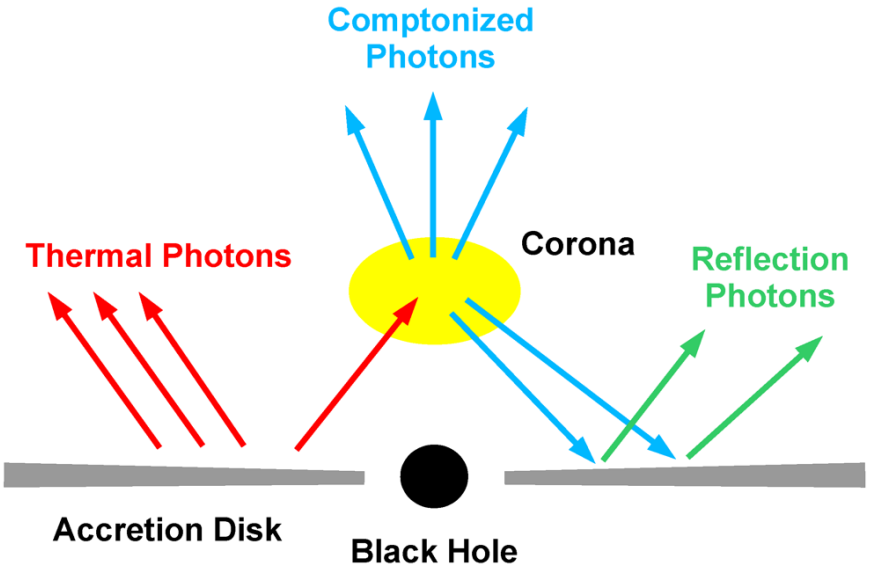


Figure 1.3: Schematic illustration of disc reflection in an X-ray binary. The geometrically thin disc emits soft photons (red arrows). A hot corona (yellow region) Compton upscatters some of these photons to produce a power-law (blue arrows). Some of these hard photons illuminate the disc, leading to reflection features (green arrows), including iron fluorescence lines and the Compton hump.

reflection: the thin disk (red) emits soft X-rays, the hot corona (yellow) Compton-upscatters some photons to hard X-rays (blue), and those hard photons irradiate the disk to produce fluorescence and the Compton hump (green).

Relativistic Effects near the Compact Object

If the reflection originates from the innermost regions of the disk (very close to the BH or NS), relativistic effects markedly shape the observed spectrum. Special relativistic Doppler shifts (from fast orbital motion) and general relativistic gravitational redshift cause the Fe Kα line to appear broad and skewed. The line's red wing can extend down to ∼4–5 keV for a disk around a rapidly spinning BH, and the blue wing to ∼7–8 keV (Reynolds & Nowak, 2003; Miller, 2007).

By fitting relativistically broadened iron lines with disk reflection models (e.g., the `relxill` model that convolves a rest-frame reflection spectrum with relativistic blurring), one can infer the inner disk radius and thus the BH spin parameter a_* (assuming the disk terminates at the ISCO). Many BHXRB observations have measured broad Fe lines and yielded BH spin estimates (McClintock et al., 2014; Reynolds, 2014). NS XRBs also exhibit broad Fe lines, though their profiles can be influenced by the NS surface or boundary layer emission (Miller et al., 2013).

In NS systems, the evidence suggests the inner disk often extends very close to the NS, as indicated by significant relativistic broadening (Miller et al., 2013). Beyond line profiles, general relativistic light-bending can focus more coronal photons onto the disk (enhancing reflection if the corona is compact) (Dauser et al., 2016), and both disk and coronal emission can be modulated by gravitational redshift and Doppler shifts (Bhattacharyya & Strohmayer, 2007). Modern spectral models incorporate these

relativistic effects to interpret high-quality X-ray data (Tomsick et al., 2014).

1.3 Emission Properties of XRBs: Spectra and Variability

X-ray binaries display complex emission features, with rich spectra and pronounced variability across a broad range of timescales. By examining both the spectral components outlined above and the temporal variability, one gains insight into the physical mechanisms in the accretion flows (van der Klis, 2006; Sartorio et al., 2023). These systems can be broadly divided into: Broadly, XRBs can be categorized by their variability patterns and by specific types of timing features they exhibit, into persistent vs. transient sources :

- **Persistent XRBs**, These maintain a relatively steady luminosity over long periods, Many wind-fed HMXBs (e.g. Vela X-1, Cen X-3) are persistent, sustained by a continuous mass supply. They can still show state changes (e.g. transitions between hard and soft states) but lack the years-long quiescent intervals of transients (Neumann et al., 2023).
- **Transient XRBs**, More common among LMXBs, transients alternate between faint (or undetectable) quiescence and dramatic outbursts. During an outburst, luminosity can increase by factors of 10^4–10^5, then decay back to quiescence over months. These eruptions are thought to be triggered by a thermal-viscous disk instability that episodically dumps matter onto the compact object (Hameury et al., 2009). V404 Cyg (a BH LMXB) Rodi et al. (2017) and Campana et al. (2013) are Well-known examples. Transients provide valuable opportunities to observe the full range of XRB behavior (from near-off to near-Eddington states) within a single system (Marcel et al., 2020).

Across both transient and persistent XRBs, the composite X-ray spectrum arises from multiple emission regions. The optically thick, geometrically thin accretion disk, described in early studies by Shakura & Sunyaev (1973), produces thermal emission typically peaking around 1 keV in black hole systems, though it can reach 2–3 keV in neutron star binaries due to the smaller radius and the presence of a stellar surface (Remillard & McClintock, 2006). Superimposed on this thermal component is a high-energy, power-law-like tail arising from inverse Compton scattering in a hot coronal region or the base of a jet (Sunyaev & Trümper, 1979). This Comptonized emission can extend to tens or even hundreds of keV, and in many black hole binaries it dominates the total flux in the so-called “hard state” (Zdziarski & Gierliński, 2004). Hard X-rays from the corona can also irradiate the accretion disk, giving rise to reflection features that include a strong Fe K emission line at 6.4–6.7 keV and a backscattered continuum component (Fabian et al., 1989; Miller, 2002). The detailed shape of the iron line often reveals Doppler and relativistic broadening, which can be used to probe the inner disk radius, black hole spin, and the geometry of the corona (Reynolds & Nowak, 2003).

”In addition to the large-amplitude flux changes already noted, X-ray binaries—especially near or beyond their peak luminosity—exhibit well-defined oscillatory signatures in their X-ray light curves. These include rapid X-ray variability and red noise,

quasi-periodic oscillations (QPOs), coherent pulsations, and features associated with magnetized neutron stars. Each of these phenomena is detailed below."

1.3.1 Rapid X-ray variability and Red Noise

X-ray binary (XRB) light curves display pronounced aperiodic variability that spans several orders of magnitude in frequency. In the Fourier domain, this broadband component appears as red noise, i.e., a power-density spectrum whose amplitude rises toward lower frequencies and is well described by a broken power-law with slopes $\gtrsim -1$ (Heil et al., 2015). The fluctuations are widely interpreted as the imprint of stochastic perturbations in the mass-accretion rate—driven by magneto-rotational turbulence and other disc instabilities—that propagate inward through the accretion flow (Ingram & Klis, 2013).

Characteristic timescales therefore range from milliseconds, set by dynamical processes at the innermost stable circular orbit, up to hundreds of seconds during luminous outbursts. The fractional root-mean-square (rms) amplitude of this red-noise component depends strongly on spectral state: it typically exceeds $\approx$ 20–30% in the hard state but falls below $\approx 5\%$ when the source enters a soft, disc-dominated state (Belloni & Stella, 2014; Motta et al., 2021).

1.3.2 Quasi-Periodic Oscillations

XRBs frequently exhibit distinct quasi-periodic oscillations. Quasi-Periodic Oscillations (QPOs) are peaks in the power spectrum indicating that some variability occurs at a preferred frequency (or narrow range of frequencies), but with a finite coherence (quality factor) rather than being a strictly periodic signal. QPOs were first identified in XRBs in the 1980s (van der Klis et al., 1985) and have since been commonly observed in dozens of systems. They are typically categorized by characteristic frequency ranges and the type of source in which they appear.

Low-frequency QPOs in black-hole LMXBs

Low-frequency QPOs (LF-QPOs) are oscillations with frequencies of order ~ 0.1–$30Hz$, often seen in black hole and neutron star binaries during their intermediate-brightness states. Black hole LMXBs, transients display three types of LF-QPOs (historically labeled A, B, and C) have been distinguished by their properties (Casella et al., 2005; Belloni, 2010).

- **Type-C QPOs** dominate the hard– and hard-intermediate states. Their centroid frequency drifts from $\simeq$ 0.1 to $\sim$ 15 Hz as the source softens, their quality factor is high ($Q \gtrsim 10$), and they ride on top of a strong band-limited noise hump. Relativistic Lense–Thirring precession of a geometrically thick inner flow remains the leading—but still debated—explanation (Ingram et al., 2009; Ingram & Motta, 2019).

- **Type-B** QPOs manifest at centroid frequencies of $\sim$ 4–6 Hz, retain moderate coherence ($Q \approx 5$–7) and fractional rms amplitudes of a few per cent, and they

appear during the brief soft-intermediate intervals that bridge the hard and thermal-dominant regimes. Their arrival is accompanied by a marked suppression of broad-band noise and is often contemporaneous—within hours—with discrete radio flares signalling compact-jet ejections (Motta et al., 2014).

Type-A QPOs, in contrast, emerge only once the spectrum is already dominated by the optically thick disc. They are rarer, weaker (rms $\lesssim 2\%$), and broader ($Q \lesssim 3$), typically clustering around ~ 7–8 Hz and showing almost no harmonics or phase lags. Because sources sometimes evolve smoothly from a Type-B to a low-coherence oscillation as the disc fraction rises, some authors contend that Type-A events may represent the low-contrast tail of the same underlying mode rather than a distinct phenomenon.

Although this A/B/C taxonomy is empirically useful, the underlying physics is almost certainly more continuous than the discrete labels imply. Any model must reproduce both the abrupt B $\leftrightarrow$ C transitions (suggesting geometry changes) and the systematically lower coherence of Type-A oscillations (implying a short-lived or more weakly modulated structure). At present, no single theory satisfies all the constraints. The physical origin of LF-QPOs is still debated; leading models involve geometrical and relativistic effects such as Lense–Thirring precession of the inner disk (Ingram & Motta, 2019) or accretion flow instabilities. For instance, the relativistic precession model posits that Type-C QPOs correspond to the nodal precession of the inner accretion disk, which would naturally produce a modulation at the Lense–Thirring precession frequency (Ingram et al., 2009; Motta et al., 2012). This model can qualitatively explain how QPO frequency correlates with spectral state (through inner disk radius changes). But low-frequency QPOs spanning a range of few mHz–60 Hz have been observed in many classes of accreting binaries, both powered by neutron stars and black holes (van der Klis, 2006).

High-frequency QPOs (HF-QPOs) – oscillations at tens to hundreds of Hz – have been observed in a few high-luminosity X-ray binaries. In some cases, the frequencies are close to the edge range of low-frequency QPOs, for example the 67 Hz QPO in GRS 1915+105 (see Motta & Belloni 2024 for an interpretation as Type-C), or the 40Hz and 80Hz QPOs in XTE J1550–564; (Remillard & McClintock, 2006); the 300-450 Hz pair in GRO J1655-40 (Motta et al., 2014) and as kHz QPOs ($\sim$ 300–1200 Hz) in many neutron star LMXBs (van der Klis, 2006). These HF-QPOs often appear in pairs and in some cases have frequency ratios near 3:2, hinting at an underlying resonance or orbital motion at the inner accretion disk. Because their frequencies approach the Keplerian orbital frequency in the innermost disk, HF-QPOs likely originate very close to the compact object and potentially encode information about the mass and spin of the accretor (Kluźniak & Abramowicz, 2001). However, HF-QPOs are typically weak and rare, making them challenging to study; to date no precise model has fully explained their occurrence.

1.3.3 Coherent Pulsations and Magnetized Neutron Stars

If the compact object is a neutron star endowed with a strong magnetic field—particularly an X-ray or millisecond pulsar—the accretion flow is funnelled onto the magnetic poles,

generating X-ray pulsations at the stellar spin period (Caballero & Wilms, 2012). Classical high-mass X-ray binary (HMXB) pulsars such as Cen X-3, with a spin period $P \approx 4.8$ s, provided the first clear example of this behaviour (Giacconi et al., 1971). Comparable but far faster signals (tens of milliseconds) arise in accreting millisecond X-ray pulsars found in low-mass systems, the prototype being SAX J1808.4-3658 (Patruno & Watts, 2020). Detecting such coherent pulsations confirms the presence of a magnetised neutron star and enables precise tracking of its spin evolution through long-term timing studies (Caballero & Wilms, 2012). In the ultraluminous regime, the discovery of pulsations in a handful of objects—most famously M82 X-2; GC 5907 and NGC 7793—has revealed a distinct population of "ultraluminous pulsars," demonstrating that neutron stars can power some ultraluminous X-ray sources (Bachetti et al., 2014; Israel et al., 2017a,b).

1.4 Ultraluminous X-ray Sources (ULXs)

Among the broad range of X-ray binaries, Ultraluminous X-ray Sources (ULXs) stand out for pushing accretion physics to its limits. Found outside their host galaxy's nucleus, ULXs are point-like X-ray emitters whose brightness surpasses that of typical stellar X-ray binaries (XRBs). ULXs are characterized by X-ray luminosities of approximately $\sim 10^{39}$ erg/s and higher, surpassing the Eddington limit for a typical black hole of mass $\sim 10\ M_\odot$ (Makishima et al., 2000; Colbert & Mushotzky, 1999). Because of these extraordinary luminosities, early studies speculated that ULXs might host intermediate-mass black holes (hundreds to thousands of solar masses. An IMBH accreting at or near its Eddington rate could feasibly achieve ULX-scale luminosities, and some early observations—such as unusually cool thermal spectra—seemed to support this notion (Miller et al., 2004).

However, ULXs also possess traits that run counter to the IMBH picture and instead point toward super-Eddington accretion onto stellar-mass black holes. Many systems display curved or two-component X-ray spectra poorly matched by simple disk models for very massive black holes, showing, for instance, a "soft excess" below ~ 2 keV and a turnover beyond ~ 5 keV (Gladstone et al., 2009). These features are better explained by a radiation-dominated disk or thick outflow in a ~ 10 solar-mass black hole that's accreting above the Eddington threshold. Such ULXs also vary over days to months—timescales more typical of lighter accretors than of a 10^4 solar-mass IMBH, whose changes would occur more slowly. As a result, the consensus emerging from recent multiwavelength studies is that most ULXs are extreme accretors: stellar-mass black holes or neutron stars taking in matter at super-Eddington rates, sometimes aided by geometric beaming of their emission (Kaaret et al., 2017b).

Super-Eddington accretion can drive radiatively powered outflows or funnels within the inflow. The expelled material may focus the X-rays in specific directions, so an observer looking down the outflow channel sees a super-Eddington luminosity—though the emission might not truly exceed the local Eddington limit by as large a factor (King et al., 2001). Even in cases with minimal beaming, objects can exceed classical Eddington limits if, for instance, the disk is thick and the radiation transport diverges from a standard thin-disk scenario, or if the accretor is a neutron star whose strong

magnetic field lowers the effective electron scattering opacity.

Dramatic confirmation of super-Eddington accretion came with neutron-star ULXs. In ULX M82 X-2 (NuSTAR J095551+6940.8), observed X-ray pulsations revealed a spinning neutron star accreting at roughly 100 times its normal Eddington rate (Bachetti et al., 2014).This breakthrough inaugurated a growing category of ultraluminous X-ray pulsars. Several more were soon detected (e.g., in NGC 5907 ULX1 and NGC 7793 P13; (Israel et al., 2017a,b)), all implying neutron-star accretors radiating as much as a few times $10^{40}\,\mathrm{erg\,s^{-1}}$. These findings confirm that some ULXs don't require an IMBH—and that neutron stars can reach extraordinary brightness under intense accretion flows.

The discovery of ultraluminous X-ray pulsars (ULXPs) has fundamentally reshaped theories of super-Eddington accretion in X-ray binaries. Coherent pulsations detected from sources such as M82 X-2, NGC 5907 ULX1, and NGC 7793 P13 reveal that highly magnetized neutron stars can sustain apparent luminosities of $L_X \sim 10^{40}$–10^{41} erg s^{-1}, several hundred times the canonical Eddington limit for a 1.4 $M_\odot$ star (Bachetti et al., 2014; Israel et al., 2017a,b). These extreme outputs are typically attributed to a strong combnation of magnetic pressure, radiation-dominated accretion columns, and geometric beaming, allowing the accretion flow to circumvent traditional luminosity constraints (Kawashima et al., 2016). At the same time, sources like HLX-1 in ESO 243-49 achieve peak luminosities near 10^{42} erg s^{-1}, exhibiting long outburst cycles that are more naturally explained by accretion onto an intermediate-mass black hole (IMBH) (Farrell et al., 2009; Godet et al., 2014).

More broadly, X-ray observations suggest that ultraluminous X-ray sources (ULXs) are not merely scaled-up versions of standard X-ray binaries but instead occupy a distinct *ultraluminous state* (Gladstone et al., 2009; Sutton et al., 2013). This regime often includes a soft thermal component—potentially linked to radiatively driven winds or an inflated inner accretion disk—alongside a harder spectral tail originating from the innermost regions near the compact object. ULXs also exhibit variability patterns consistent with super-Eddington accretion, including dips, stochastic fluctuations, and, in rare cases, quasi-periodic oscillations (QPOs). Notably, M82 X-1 displays a persistent QPO near 0.2 Hz, resembling the low-frequency QPOs seen in Galactic black hole binaries, albeit at much higher luminosities (Strohmayer & Mushotzky, 2003; Pasham et al., 2014a).

By bridging the gap between stellar-mass X-ray binaries and active galactic nuclei, ULXs provide a natural laboratory for testing how compact objects of different masses respond to extreme accretion rates. While many ULXs may represent neutron stars or stellar-mass black holes undergoing super-Eddington accretion, some—such as HLX-1—remain compelling candidates for hosting IMBHs, raising important questions about black hole formation channels and the upper mass limit for stellar remnants (Mezcua, 2017; Pasham et al., 2014b).

1.5 ULXs in Starburst Galaxies

ULXs are observed in a variety of galactic environments, but they show a strong preference for galaxies with active star formation. Starburst galaxies, which undergo short-

lived episodes of exceptionally high star formation rate (often tens or hundreds of times the Milky Way's rate; Weedman 1983), are especially fertile ground for ULXs. In a starburst, the copious production of massive stars leads, after a few million years, to a correspondingly large population of compact remnants (NSs and BHs) and X-ray binaries. It is therefore not surprising that the number and total luminosity of ULXs correlate with the star formation activity of the host galaxy (Swartz et al., 2011; Mapelli et al., 2010). ULXs tend to be absent or very rare in quiescent elliptical galaxies (which lack young massive stars), whereas star-forming spiral and irregular galaxies can host multiple ULXs. For example, the Cartwheel galaxy (a ring starburst) and the Antennae galaxies (an interacting pair) each contain dozens of ULX candidates (Gao et al., 2003; Zezas & Fabbiano, 2002). In these environments, ULXs are thought to be predominantly HMXB systems—either a neutron star or stellar-mass BH accreting from a short-lived massive donor star—hence tightly linked to recent (< 10 Myr) star formation.

Starburst galaxies also shine in X-rays through other mechanisms: they have a diffuse thermal X-ray glow from hot gas (heated by supernovae and stellar winds) and many normal X-ray binaries and supernova remnants contributing to the emission (Griffiths et al., 2000; Persic & Rephaeli, 2002). ULXs, however, often account for a significant fraction of the hard X-ray output of a starburst galaxy, outshining the numerous fainter sources. Their extreme luminosities and often off-nuclear locations make them standout features in high-resolution X-ray images of starbursts.

A prototypical example is the nearby starburst galaxy M82 (NGC 3034). M82, colloquially known as the "Cigar Galaxy", is undergoing a vigorous burst of star formation in its central regions (O'Connell et al., 1995). It was one of the first galaxies recognized to harbor ULXs: early X-ray observations with *Einstein* and *ROSAT* revealed two exceptionally bright X-ray sources in M82, later designated M82 X-1 and M82 X-2. With the advent of *Chandra* and *XMM-Newton*, M82 X-1 was found to reach luminosities of $\sim 10^{40}$ erg s^{-1}, firmly in the ultraluminous regime (Kaaret et al., 2001). M82 X-1 became a prime IMBH candidate after the discovery of a ≈ 54 mHz QPO in its X-ray flux (Strohmayer & Mushotzky, 2003). This low-frequency QPO, if interpreted as a scaled-down analog of the type-C QPOs in stellar-mass BHs, suggested a black hole mass on the order of $10^3\,M_\odot$. Further excitement came when *XMM-Newton* and *Chandra* observations found twin QPO peaks at 3.3 Hz and 5.1 Hz from M82 X-1, in a 3:2 frequency ratio. Such a pair mirrors the 3:2 high-frequency QPO pairs seen in Galactic BH XRBs (like the 300 and 450 Hz QPOs in GRO J1655–40) but scaled down in frequency, again hinting at an intermediate-mass BH. Pasham et al. (2014a) reported these twin QPOs and estimated a black hole mass of $\sim 400\,M_\odot$ for M82 X-1 if the 3:2 QPOs represent orbital motion near the innermost stable orbits. While not yet conclusive, this evidence positions M82 X-1 as one of the most compelling IMBH candidates known.

M82's other ultraluminous source, M82 X-2, delivered a surprise of a different kind. In 2014, *NuSTAR* observations revealed coherent pulsations at 1.37 s from M82 X-2, identifying it as an accreting neutron star rather than a black hole (Bachetti et al., 2014). The pulsar's spin period, combined with its luminosity (approximately $\sim 10^{40}$ erg s^{-1}), implied that M82 X-2 is accreting at ∼100–500 times the Eddington rate for a $1.4\,M_\odot$ neutron star. This discovery was the first confirmation that ULXs can

be powered by neutron stars, overturning the assumption that ultraluminous sources must be black holes. M82 X-2 (also called ULX pulsar M82 X-2) likely has an extremely strong magnetic field (on the order of 10^{13}–10^{14} G) that allows it to funnel accretion and produce such high luminosity without immediately blowing away the inflow. Its pulsations and spin-up rate have provided valuable data on how a neutron star manages hyper-Eddington accretion. The fact that M82, a single starburst galaxy, hosts two ULXs of very different nature—one a candidate IMBH and one a neutron star—underscores the diversity of ultraluminous X-ray sources and the special conditions in starbursts that produce them. In M82's starburst, massive binary evolution can yield either path: a stellar-collapse BH possibly growing via accretion into an IMBH, or a highly magnetized neutron star pushing accretion to record levels.

More broadly, ULXs in starbursts are key objects for understanding the end-products of massive stars in an environment of intense stellar formation. The high supernova rate in a starburst not only seeds many NSs and BHs, but the ensuing stellar feedback (from supernovae and winds) can alter the environment around these XRBs. Starbursts drive galactic superwinds—large-scale outflows of hot gas carrying thermal and kinetic energy into the galaxy's halo (Chevalier & Clegg, 1985; Veilleux et al., 2005). These outflows, observed in M82 and others, can interact with the ULX radiation or the ULX's own smaller-scale winds. While the detailed interplay is complex, one can imagine that a ULX in a dense starburst region might be surrounded by nebulae ionized by its X-rays or by shock-heated bubbles from the collective effects of nearby supernovae. Indeed, some ULXs are associated with optical emission nebulae (Vavilova et al., 2020) , indicating they deposit energy into their surroundings. The spatial distribution of ULXs in starbursts often correlates with stellar clusters or OB associations, supporting their origin in massive binaries. For example, in the Antennae galaxies (NGC 4038/39), ULXs are found near stellar clusters generated in the galaxy collision (Zezas & Fabbiano, 2002). The study of ULXs in different starburst environments—from the dense nuclear starburst of M82 to more extended disks of spirals—helps us learn how factors like metallicity and cluster formation affect the upper end of the X-ray source population. Low metallicity (metal-poor) starbursts might produce more ULXs because lower metal content leads to weaker stellar winds, so massive stars retain more mass and can form heavier BHs (Mapelli et al., 2010; Linden et al., 2010). This is relevant for understanding ULX populations in dwarf starbursts or early-universe star-forming galaxies.

In conclusion, starburst galaxies provide the cosmic laboratories where the most extreme X-ray binaries—ULXs—are forged and observed in large numbers. Their abundant high-mass XRBs and frequent dynamical interactions yield a higher probability of achieving the conditions for ultraluminous X-ray output. Through meticulous observations of ULXs in starbursts like M82, NGC 253, and others, astronomers are piecing together the life cycles of these extraordinary binaries. The temporal variability of ULXs (whether in the form of QPOs, pulsations, or chaotic flaring) carries the imprint of accretion physics under extreme conditions. By analyzing this variability, as well as the spectral signatures, we gain insight into whether a ULX is a neutron star or black hole, how its accretion disk and outflows behave, and how it fits into the broader context of its host galaxy's evolution. These insights not only illuminate the nature of ULXs themselves but also inform models of stellar remnants, binary star evolution,

and feedback in galaxies. The subsequent chapters will build on this foundation, exploring in depth the temporal analysis methods and discoveries related to variability in ULXs (with a focus on M82) and extending the investigation to other high-energy astrophysical data, such as the Planck cosmic microwave background maps, through advanced mathematical and deep learning techniques.

Chapter 2

Mathematical Foundations and Statistical Methods for Time Series Analysis

2.1 Introduction

Time series analysis is a branch of statistics focused on modeling and understanding data that are observed sequentially over time. Broadly, two complementary approaches exist: the *time–domain* approach, which develops models for temporal dependencies (often for forecasting), and the *frequency–domain* approach, which analyses the data in terms of periodic or oscillatory behaviour. Classic texts by Box et al. (2015) and Brockwell & Davis (2016) detail time-domain modelling strategies (e.g. ARIMA processes), whereas this chapter emphasises frequency-domain methods and their associated inference tools, which are especially useful for studying periodicities and variability in physical signals.

In astrophysics—particularly in high-energy X-ray astronomy—frequency-domain techniques have become indispensable. The rapid, aperiodic variability of X-ray sources is often characterised by analysing their power spectra to identify features such as quasi-periodic oscillations (QPOs). By decomposing a light curve into its frequency components, one can identify characteristic time-scales and frequencies of variability that provide insight into the underlying physical processes (e.g. orbital motions or disk instabilities) near compact objects. For instance, frequency-domain analysis has been used to detect QPOs and to constrain the geometry and dynamics of accretion flows around black holes and neutron stars.

We now lay the mathematical groundwork for these analyses, covering fundamental time-series concepts, spectral-analysis methods, statistical inference in the frequency domain (including the Whittle likelihood), and bootstrap techniques for spectral analysis, before discussing their application to X-ray timing and QPO detection.

2.2 Fundamental Time-Series Concepts

2.2.1 Stochastic Processes and Stationarity

A *time series* can be viewed as a realisation of a stochastic process, formally defined as a collection of random variables $\{X(t) \colon t \in T\}$ indexed by time t on a probability space $(\Omega, \mathcal{F}, P)$. In this thesis we deal mainly with discrete-time series

$X_1, X_2, \ldots, X_n$ representing sequential observations (e.g. photon counts in successive time bins). The probabilistic structure of a time series is described by the joint distributions of $\big(X(t_1), \ldots, X(t_k)\big)$ for any set of time points.

An important simplifying property is *stationarity*, which—loosely speaking—means that the statistical properties do not change over time. A process $X(t)$ is *strictly stationary* if, for any collection of times $t_1, \ldots, t_n$ and any time shift h, the joint distribution satisfies

$$F_X\big(X(t_1), \ldots, X(t_n)\big) = F_X\big(X(t_1 + h), \ldots, X(t_n + h)\big),$$

i.e. it is invariant under time translation. Strict stationarity imposes a very strict requirement, so analyses typically assume a milder condition called second-order (weak) stationarity (Brockwell & Davis, 2016). A process is second-order stationary if its mean $\mu = E[X(t)]$ is constant in time and its covariance depends only on the lag between two observations:

$$\mathrm{Cov}\big(X(t), X(t + \tau)\big) = \gamma(\tau),$$

a function of τ alone. Equivalently, the autocovariance function $\gamma(\tau)$ depends only on τ and not on absolute time. For a weakly stationary process, we define the normalized autocorrelation $\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)}$, capturing how observations are correlated at lag τ. The Autocorrelation Function (ACF) provides insight into the memory of the process and is a primary tool for identifying temporal-dependence structures.

Many natural time series, including those from astronomical sources, exhibit stochastic (aperiodic) behaviour that can often be treated as arising from a stationary or approximately stationary process over short time-scales. Stationarity is an essential assumption for the frequency-domain methods discussed later, since the very definition of a power spectrum relies on time-invariant properties. If a series is non-stationary, techniques such as detrending or differencing (in the time domain) or time-localised spectral methods may be required (Priestley, 1981). Throughout this chapter we assume weak stationarity unless stated otherwise.

2.2.2 Examples of Processes and Autocovariance Structure

A broad range of statistical models exist for stationary time series. For example, the classical *autoregressive moving-average* (ARMA) processes introduced by Box et al. (2015) are defined by linear difference equations and can reproduce a variety of autocorrelation structures. In an AR(p) process

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t,$$

the parameters ϕ_i control how past values influence the present, and the noise ε_t is typically white. Such processes are stationary if the roots of the characteristic polynomial lie outside the unit circle. The autocovariance $\gamma(\tau)$ of an ARMA process decays exponentially—or as a mixture of exponentials and damped oscillations—reflecting short-memory behaviour (Brockwell & Davis, 2016). Other processes, such as fractional-differencing models, can produce long-memory behaviour with $\gamma(\tau)$ decaying algebraically.

For a given process the total variance is $\gamma(0) = \mathrm{Var}(X_t)$, and the portion of variance explained by structure at lag τ is $\gamma(\tau)$; the autocorrelation $\rho(\tau)$ gives a normalised

measure. In practice an empirical ACF is computed from the data and compared with theoretical forms or used in model identification. For instance, an exponentially decaying ACF suggests an AR(1)-like process, whereas an oscillatory ACF might indicate an under-damped system or cyclic behaviour. In astronomical time series (light curves), the ACF has been used to identify characteristic time-scales—for example, X-ray light curves of active galactic nuclei (AGN) often show an ACF that decays on time-scales comparable to the viscous time of the accretion disk. A short ACF length implies rapid decorrelation of flux (perhaps small emitting regions or rapid accretion fluctuations), whereas a long correlation time suggests slowly varying processes. Such connections between the ACF and physical mechanisms underscore the importance of careful time-domain characterisation as a prelude to frequency-domain analysis.

2.3 Frequency-Domain Methods

2.3.1 The Fourier Transform and Spectral Representation

Frequency-domain analysis treats the data as a sum of sinusoids. The key mathematical tool for this decomposition is the Fourier transform. For a continuous-time function $x(t)$ that is square-integrable, the Fourier transform is defined by

$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-i2\pi ft}\, dt, \qquad x(t) = \int_{-\infty}^{\infty} X(f)\, e^{i2\pi ft}\, df,$$

where f denotes frequency (cycles per unit time). In practice, our finite discrete time series use the discrete Fourier transform (DFT), typically computed with a fast FFT algorithm.

2.3.2 Spectral Representation Theorem

A fundamental result for stationary processes is the *spectral representation theorem*, which states that any zero–mean, weakly stationary process can be expressed as an integral superposition of orthogonal sinusoidal components with random coefficients (Priestley, 1981). Intuitively, the randomness in the time series can be viewed as the combined contribution of many independent frequency components, motivating analysis of how the variance is distributed across frequency bands.

2.3.3 Power Spectral Density

The distribution of a process's variance over frequency is quantified by the *power–spectral density* (PSD). For a weakly stationary process $X(t)$ with autocovariance function $\gamma(\tau)$, the Wiener–Khinchin theorem defines the PSD via the Fourier transform

$$S_X(f) = \int_{-\infty}^{\infty} \gamma(\tau)\, e^{-i2\pi f\tau}\, d\tau, \tag{2.1}$$

with inverse relation

$$\gamma(\tau) = \int_{-\infty}^{\infty} S_X(f)\, e^{i2\pi f\tau}\, df.$$

Equation (2.1) implies that $S_X(f)\,df$ is the contribution to the variance from an infinitesimal band around frequency f. By Parseval's identity,

$$\int_{-\infty}^{\infty} S_X(f)\,df \;=\; \gamma(0) = \mathrm{Var}\big(X(t)\big).$$

Alternatively, for a series observed over T seconds, let $X_T(f)$ be the Fourier transform of the truncated data. Then

$$S_X(f) \;=\; \lim_{T\to\infty} \frac{1}{T}\,\mathbb{E}\Big[\,|X_T(f)|^2\,\Big],$$

and an analogous definition holds in discrete time. In practice, $S_X(f)$ is unknown and must be estimated from finite data.

Many physical PSDs are smooth continua, but certain phenomena appear as distinct features: a pure sinusoid yields a δ–function spike, whereas broadband stochastic processes often display power–law ("red–noise") behaviour. High-energy X-ray light curves frequently exhibit red noise with superposed narrow QPO peaks.

2.3.4 Periodogram and Basic Spectral Estimation

For a discrete series $x_1, \dots, x_N$ sampled every Δt, the *periodogram*

$$I(f_j) = \frac{\Delta t}{N}\left|\sum_{k=1}^{N} x_k\, e^{-i2\pi f_j t_k}\right|^2, \qquad f_j = \frac{j}{N\Delta t}, \tag{2.2}$$

is the simplest empirical PSD estimator. Although $\mathbb{E}[I(f_j)] = S_X(f_j)$ (away from $f = 0$), the raw periodogram is *inconsistent*: $\mathrm{Var}[I(f_j)] \approx S_X(f_j)^2$ even as $N \to \infty$ (Percival & Walden, 1993). Variance reduction therefore relies on averaging techniques (e.g. Bartlett, Welch) or tapering, each trading resolution for stability (Bartlett, 1950a; Welch, 1967).

Parametric alternatives fit a model PSD (e.g. AR(p) spectra) whose coefficients are estimated from the data (Yule–Walker or maximum likelihood) (Kay & Marple, 1981). When well-specified, parametric PSDs are smoother and less variable than non-parametric estimates, though susceptible to bias if the assumed model is wrong.

2.4 Statistical Inference in the Frequency Domain

For Gaussian, weakly stationary processes, discrete Fourier coefficients at distinct Fourier frequencies are asymptotically independent; the periodogram ordinates satisfy

$$\frac{I(f_j)}{S_X(f_j)} \;\sim\; \chi^2_2,$$

providing an exponential likelihood for $I(f_j)$. Whittle (1953) exploited this to propose the *Whittle likelihood*

$$\ell_{\mathrm{Whittle}}(\boldsymbol{\theta}) = -\sum_{j=1}^{N'}\Big[\,\ln f_{\boldsymbol{\theta}}(f_j) + \tfrac{I(f_j)}{f_{\boldsymbol{\theta}}(f_j)}\Big], \tag{2.3}$$

where $f_{\boldsymbol{\theta}}$ is a model PSD with parameters $\boldsymbol{\theta}$. Maximum-likelihood fitting in the frequency domain is far cheaper than exact time-domain likelihoods and remains asymptotically efficient under standard regularity conditions. When the periodogram has been smoothed (averaged, tapered), its distribution generalises to χ^2_ν with $\nu > 2$ degrees of freedom; the likelihood is modified accordingly.

2.5 Bootstrap Methods for Power–Spectral Analysis

Analytical confidence intervals may be unreliable for short or non-Gaussian series. Resampling provides a flexible alternative:

2.5.1 Block Bootstrap

To preserve temporal dependence, contiguous blocks of length ℓ are resampled (moving, circular, or stationary block variants) (Hall et al., 1995). Choosing ℓ on the order of the longest correlation time balances bias and variance (Shumway et al., 2000).

2.5.2 Frequency-Domain Bootstrap and Phase Randomisation

Because Fourier amplitudes are (approximately) independent, one may randomise phases: multiply each Fourier coefficient by $e^{i\phi_j}$ with $\phi_j \sim \mathrm{Unif}(0, 2\pi)$ and invert the transform (Theiler et al., 1992). Surrogates thus preserve the PSD exactly while removing deterministic phase structure, enabling Monte-Carlo significance tests for narrow features.

2.5.3 Bootstrap Confidence Intervals

Given B bootstrap replicates $\{I^*_b(f_0)\}_{b=1}^{B}$ at some frequency f_0:

- **Normal interval**: $\hat{I}(f_0) \pm z_{1-\alpha/2}\,\hat{\sigma}$, where $\hat{\sigma}$ is the bootstrap standard error.
- **Percentile interval**: $[\, I^*_{(\alpha/2)},\ I^*_{(1-\alpha/2)} \,]$.
- **BCa interval**: Bias-corrected and accelerated percentiles using z_0 and a (Efron, 1987).
- **Studentised (bootstrap-t) interval**: Resample the studentised statistic $t_b = \big(I^*_b(f_0) - \hat{I}(f_0)\big)/\hat{\sigma}^*_b$.

These intervals illustrate the bias–variance trade-offs that underpin modern spectral-analysis practice. Detailed algorithms and further discussion can be found in Efron & LePage (1992).

Each of these methods has advantages and drawbacks. The *normal* interval is simple, but it relies on an unstated normality assumption; the *percentile* and *BCa* intervals are non–parametric and adapt to the bootstrap distribution, yet they are sensitive to the bootstrap sample size B; the *studentised* interval often performs well

in theory, but it requires a stable estimate of the standard error for *every* bootstrap sample.

In spectral problems, raw periodogram ordinates are highly skewed (exponential), so a normal approximation can be poor on the original scale. It is therefore common to work with a *log–periodogram* or with an *averaged* periodogram (which increases the degrees of freedom and reduces skewness) before applying the bootstrap.

The bootstrap is readily extended to more complex tasks, such as simultaneous confidence *bands* for an entire PSD curve or uncertainty estimates for derived quantities (e.g. the location of a spectral peak). Its appeal is that no explicit sampling distribution is required—the distribution is obtained empirically by resampling the data.

2.6 Applications in X-ray Timing Analysis and QPO Detection

High-energy astrophysical sources—accreting black holes and neutron stars—exhibit variability that is often described by a broadband PSD approximated by a power-law in frequency ("red noise"). Detecting a quasi-periodic oscillation (QPO) in such data is inherently a statistical challenge: one must identify a weak, narrow peak against a noisy, steep continuum. The Fourier tools, Whittle-likelihood fitting, and bootstrap or Monte-Carlo techniques developed above are routinely combined to meet this challenge.

2.6.1 Characterising the Noise Continuum

Before claiming a QPO detection, the underlying continuum must be modelled. Typical choices are a simple power law $S(f) \propto f^{-\alpha}$ or a sum of broad Lorentzians (Vaughan, 2013). Because a pure red-noise process produces occasional large periodogram peaks (the "look-elsewhere" effect), significance tests must account for the distribution of the *maximum* noise power.

A widely used approach is to simulate artificial light curves under the null hypothesis of *no* QPO. The algorithm of Timmer (1995) generates Gaussian time series with a prescribed PSD: one draws complex Fourier coefficients with amplitudes $|X(f_j)| = \sqrt{S(f_j)\, N\Delta t/2}$ and random phases, then performs an inverse FFT. By creating, say, $B = 1000$ such curves, computing their periodograms, and recording the largest peak in each, one obtains the null distribution of the maximum power. If the observed peak exceeds the 99.9 th percentile of this distribution, the QPO is detected at 0.1 % false-alarm probability (p-value $= 10^{-3}$). Phase-randomisation surrogates (Theiler et al., 1992) supply a similar test while *exactly* preserving the empirical PSD.

2.6.2 Detection and Parameter Estimation of QPOs

A candidate QPO appears as a localised excess above the continuum. The PSD is then fitted with a model consisting of a continuum plus a Lorentzian peak; using the Whittle likelihood in Eq. (2.3), one obtains maximum-likelihood estimates of the QPO centroid ν_{QPO}, quality factor Q, and amplitude (Barret, 2012). Likelihood-ratio intervals (Wilks' theorem) or curvature-based errors give first-order confidence regions,

but bootstrapping provides a more robust alternative when Gaussian or large-sample assumptions are questionable:

1. Generate bootstrap time series (block bootstrap or phase-randomised surrogates).
2. Fit each bootstrap PSD with the same Lorentzian+continuum model.
3. Use the resulting empirical distribution of $\nu_{\rm QPO}$, Q, and amplitude to form confidence intervals and assess detection significance.

If none of the B bootstrap (or Monte-Carlo) spectra contains a peak as strong as the real data, the QPO is highly significant.

In summary, combining the mathematical foundations of stochastic processes with modern spectral-inference techniques—Whittle likelihood for model fitting, plus bootstrap and simulation for uncertainty quantification—yields a rigorous framework for X-ray timing analysis. Such methods ensure that reported discoveries (e.g. a new QPO) are supported by sound statistical evidence and furnish reliable constraints on the extreme physics occurring near compact objects.

Applications in Astrophysical Time Series Analysis

Bootstrap methods have proven particularly valuable in astrophysical time series analysis, where noise characteristics can be complex and non-Gaussian. Applications include:

Significance Testing of QPOs: By generating bootstrap distributions of peak heights in power spectra, astronomers can assess the statistical significance of potential quasi-periodic oscillations against the null hypothesis of red noise. This approach accounts for the multiple testing problem inherent in searching for peaks across many frequency bins. For example, Timmer (1995) and Uttley et al. (2002) demonstrated the use of bootstrap methods to establish robust significance thresholds for QPO detection in AGN light curves.

Parameter Uncertainty Estimation: We can use bootstrap resampling to estimate uncertainties for model parameters in complex power spectra, such as broken power laws or Lorentzian components. For example, in modeling the power spectrum of an X-ray binary, bootstrapping can yield confidence regions for QPO frequencies, widths, and amplitudes, accounting for correlations between these parameters. Vaughan (2005) applied this approach to constrain power-law slopes and break frequencies in AGN power spectra.

Cross-Spectral Analysis: Bootstrap techniques can quantify uncertainties in phase lags and coherence functions between different energy bands in X-ray observations, providing insights into the causal relationships between different emission regions. This is particularly important for reverberation mapping studies that aim to constrain the geometry of accretion flows. Uttley et al. (2014a) applied bootstrap methods to establish confidence intervals on time lags in X-ray reverberation studies.

Red Noise Simulation: Parametric bootstrap methods can generate synthetic red noise light curves with statistical properties matching observed data, providing a robust basis for significance testing of potential signals against realistic noise backgrounds. The method of Timmer (1995) generates surrogate light curves that match the original power spectrum but with randomized phases, enabling robust hypothesis tests.

Non-stationary Time Series Analysis: Bootstrap methods can be adapted to analyze time-varying power spectra, allowing for the detection of transient oscillations or evolving spectral features in X-ray observations. Belloni et al. (2002) used bootstrapping to characterize the dynamic power spectral properties of black hole binaries during state transitions.

The bootstrap is particularly useful in applications where parametric assumptions about noise distribution are unsuitable, making it an effective tool for analyzing red noise processes with complex spectral shapes. With today's high-performance computing resources, it can be efficiently implemented, allowing for the generation and execution of thousands of bootstrap replications. It provides accurate approximations of variance and bias in estimators and is well-suited to handling the non-Gaussian nature of most spectral estimators. Additionally, it can accommodate complex spectral forms and multiple fitted components simultaneously, offering comprehensive uncertainty quantification for advanced spectral models.

Furthermore, the bootstrap properly accounts for correlations between parameters in multivariate estimation problems, such as multiple QPO parameters used in joint confidence region analyses. It supports various sampling patterns, including gapped or uneven observations, which are common in astronomical time series. Its relative simplicity and flexibility make it an indispensable tool in statistical inference for time series analysis, particularly for complex or non-standard astrophysical data. As computational capabilities continue to advance, bootstrap methods are likely to become increasingly integrated with other sophisticated techniques, such as Bayesian inference and machine learning, for analyzing astronomical time series.

2.6.3 Applications in Astrophysical Data Analysis

X-ray Timing Analysis Applications

The mathematical framework and statistical methods outlined in the previous sections find extensive application in X-ray timing studies, where the PSD of source flux variations provides critical insights into the physical processes operating in extreme environments (Ingram & Motta, 2019). Power spectra of accreting black holes typically are comprised of complex structures including:

- **Broadband Noise Components**, Often modeled as power laws or Lorentzians, these components reflect stochastic variability in the accretion flow (Nowak, 2000). The shape and strength of these components provide information about the geometry and dynamics of the accretion disk and corona.
- **Quasi-Periodic Oscillations (QPOs)** : Manifest as narrow peaks in the power spectrum, these features are associated with characteristic dynamical frequencies

in the inner accretion flow, potentially including effects of General Relativity (Stella & Vietri, 1998; Ingram et al., 2009). Both low-frequency (0.1-30 Hz) and high-frequency (40-450 Hz) QPOs have been observed in black hole binaries, with their frequencies often scaling with the mass of the compact object.

- **Breaks and Cutoffs**: Changes in power-law slope or exponential cutoffs at specific frequencies constrain the geometry and physical scales of the emission region (McHardy et al., 2006; González-Martín & Vaughan, 2012). For example, break frequencies in AGN power spectra have been found to scale inversely with black hole mass, suggesting a universal process in accretion physics across mass scales.

The statistical techniques developed for spectral estimation must be further adapted to account for instrumental effects in X-ray detectors, including dead time, Poisson noise, and background contamination (van der Klis, 1989; Vaughan et al., 2003). These adaptations include:

- **Leahy Normalization:** A standard normalization such that the statistical distribution of the powers of the Poissonian noise follows a chi-squared (χ^2) distribution with 2N degrees of freedom, where N is the number of averaged PSDs. For $N = 1$, the expected value of the PSD powers is 2, as is the variance. The chi-squared distribution, i.e. the knowledge of the statistical behavior of noise powers, provides an objective statistical tool to assess whether a peak in the PSD belongs to the noise or represents a candidate signal.

- **Fractional RMS Normalization**: Expresses variability amplitude as a fraction of the mean count rate, facilitating comparison between different sources or observations.

- **Cross-Spectral Analysis**: Techniques for analyzing coherence and phase lags between different energy bands, providing insights into causal connections and time delays between different emission regions.

Maximum likelihood estimation plays a crucial role in fitting parametric models to observed power spectra, allowing for the quantification of QPO parameters (frequency, width, amplitude) and their uncertainties. Bootstrap methods provide robust estimates of confidence intervals, particularly important when assessing the significance of potential QPOs against complex noise backgrounds. These techniques have been instrumental in advancing our understanding of accretion physics and strong-field gravity. For example:

The discovery and characterization of high-frequency QPOs in black hole binaries have provided observational tests for relativistic precession models (Stella & Vietri, 1998). The observation of similar variability patterns across both stellar-mass and supermassive black holes suggests a scale-invariant nature of accretion processes (McHardy et al., 2006). Energy-dependent time lags detected through cross-spectral analysis have revealed the reverberation of X-ray signals, constraining the geometry of the accretion flow and corona (Uttley et al., 2014b).

The application of these techniques to current and next-generation X-ray missions promises to further refine our understanding of the extreme physics around compact objects.

2.6.4 Advanced Computational Approaches

Neural Networks in Astronomical Time Series Analysis

Convolutional Neural Networks (CNNs) represent a specialized class of deep neural networks designed to efficiently process data that exhibit a grid-like topology, such as images (LeCun et al., 1998). In astronomy, CNNs have gained increasing traction for automated feature extraction in large-scale datasets. Unlike traditional machine learning algorithms, which often require painstaking feature engineering, CNNs learn relevant features directly from raw input data by leveraging convolutional filters and pooling operations to detect and aggregate spatially correlated signals (Krizhevsky et al., 2012). Key components that make CNNs particularly well-suited for astronomical data include:
Convolutional Layers: These layers apply trainable filters (kernels) across an input image or feature map, systematically extracting local patterns such as emission peaks or extended filaments; **Pooling Layers**: Pooling reduces the spatial dimensions of feature maps, aggregating information and providing invariance to small translations or distortions; **Activation and Normalization**: Nonlinear activation functions, such as ReLU (Rectified Linear Unit), and normalization strategies, including Batch Normalization (Ioffe & Szegedy, 2015). improve training stability and accelerate convergence; **End-to-End Learning**: CNNs adapt their internal representations to optimize a task-specific loss function.

Recent advances expanded CNN applications to time series processing in astronomy to cover: **Feature Detection in Power Spectra**: CNNs can automatically identify quasi-periodic oscillations and other features in complex power spectra where traditional methods might struggle with noise or complex backgrounds (Huertas-Company et al., 2018). For example, a CNN trained on simulated power spectra can detect weak QPOs embedded in red noise with higher sensitivity than classical detection methods; **Classification of Variable Sources**: Time-domain surveys generate vast libraries of light curves that require efficient classification. CNNs can distinguish between different types of variable stars, transients, and active galactic nuclei based on their temporal signatures (Dieleman et al., 2015). The Zwicky Transient Facility and other large surveys have begun implementing CNN-based classification pipelines to identify rare or interesting variable sources in real-time; **Anomaly Detection**: By learning the typical characteristics of astronomical time series, CNNs can flag unusual or potentially interesting objects that deviate from expected patterns, enabling focused follow-up observations. This approach has been used to identify peculiar variability patterns in blazar monitoring data and to detect instrumental artifacts in X-ray light curves.

In source detection applications, CNNs can identify both point-like and extended sources by discerning spatial patterns of brightness and morphology. By capitalizing on GPU acceleration, CNN-based pipelines can sift through millions of time series far more efficiently than classical methods (Alger et al., 2018).

Integration of Deep Learning with Statistical Approaches

The integration of traditional statistical methods with deep learning approaches represents a promising direction for astrophysical time series analysis. Several hybrid approaches have emerged:

Neural Bootstrapping applies Bootstrap resampling with neural networks to provide prediction uncertainty estimates for CNN. This produces multiple bootstrap samples of input, feeds each through the network, and examines output distribution to estimate prediction uncertainty (Michelucci & Venturini, 2021).

Likelihood-Free Inference uses simulation-based inference with neural networks as surrogate models for complex likelihood functions that would be intractable with traditional MLE approaches. Techniques such as Approximate Bayesian Computation (ABC) and neural density estimation allow for Bayesian inference without explicitly formulating likelihood functions (Papamakarios et al., 2019; Alsing et al., 2019).

Bayesian Neural Networks incorporate Bayesian principles into neural network training to provide posterior distributions over model parameters rather than point estimates. By treating network weights as probability distributions rather than fixed values, these methods quantify epistemic uncertainty in model predictions (Möller & de Boissière, 2020).

Physics-Informed Neural Networks: Constraining neural network training with known physical laws and relationships, ensuring that predictions respect conservation principles or other theoretical constraints (Sel et al., 2023). This approach combines the flexibility of neural networks with the theoretical rigor of physical models.

These hybrid approaches leverage the flexibility and pattern recognition capabilities of deep learning while maintaining the statistical rigor and uncertainty quantification of traditional methods. As astronomical datasets continue to grow in volume and complexity, such integrated approaches will become increasingly valuable for extracting maximum scientific insight from time series observations. The field of astronomical time series analysis is currently experiencing rapid evolution with the convergence of advanced statistical methods, high-performance computing, and sophisticated machine learning techniques. This synergy promises to unlock new discoveries from both existing data archives and forthcoming large-scale surveys.

Architecture and Working Principles of CNNs

Convolutional Neural Networks (CNNs) have revolutionized data analysis across numerous domains, including astronomical time series analysis. The architecture of a CNN is specifically designed to exploit the spatial or temporal structure present in the input data, making it ideal for analyzing signals that exhibit local patterns and hierarchical features (LeCun et al., 2015).

Fundamental components

1. **Input layer.** Raw or pre-processed light-curves (normalised, detrended, or converted to a time–frequency map).

2. **Convolutional layers.** For a discrete signal $x[n]$ and kernel $w[k]$ of length K, the

convolution at position n is

$$y[n] = \sum_{k=0}^{K-1} w[k]\, x[n+k], \tag{2.4}$$

where early layers capture simple local variations and deeper layers combine them into higher-level motifs (e.g. burst envelopes or quasi-periodic modulations).

3. **Activation functions.** Non-linearities such as $\mathrm{ReLU}(z) = \max(0, z)$, Leaky-ReLU or ELU promote hierarchical feature learning.

4. **Pooling layers.** Max, average, or global pooling summarise local information, give translation invariance, and reduce parameter count.

5. **Batch normalisation** (Ioffe & Szegedy, 2015). For a mini-batch $\{x_i\}$,

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}, \quad y_i = \gamma \hat{x}_i + \beta, \tag{2.5}$$

accelerating convergence and acting as a mild regulariser.

6. **Dropout** (Srivastava et al., 2014). Randomly zeroing a fraction p of activations during training reduces co-adaptation and approximates an ensemble of subnetworks.

7. **Fully-connected & output layers.** Dense layers aggregate the learned features and produce class probabilities (soft-max) or regression outputs.

Specialised architectures for time-series

- **1D CNNs**: One-dimensional convolutions slide only along the time axis, making them suitable for raw time series data such as light curves. The filters learn to detect temporal patterns at different scales (Hon et al., 2018).

- **2D CNNs**: When time series are transformed into time-frequency representations (e.g., spectrograms or scalograms), 2D CNNs can exploit both temporal and frequency patterns. This approach has been particularly effective for detecting quasi-periodic oscillations in power spectra (Bouchard et al., 2019).

- **Recurrent Convolutional Networks**: These hybrid architectures combine CNNs with recurrent neural networks (RNNs) such as Long Short-Term Memory (LSTM) or Gated Recurrent Units (GRU). The CNN component extracts local features, while the RNN component captures long-range dependencies in the time series (Brunel et al., 2019).

- **Temporal Convolutional Networks (TCNs)**: These use dilated causal convolutions to achieve a large receptive field while maintaining computational efficiency. TCNs have shown superior performance to RNNs on many time series tasks while being easier to parallelize during training (Bai et al., 2018).

- **Attention-augmented CNNs**: Incorporating attention mechanisms allows these networks to focus on the most relevant parts of the input sequence when making predictions, improving performance on complex time series tasks (Cheng et al., 2016).
- Variational Autoencoders (VAEs): VAEs map time series into a compact, probabilistic latent space that preserves its statistical properties, making them well-suited for anomaly detection, feature extraction and probabilistic forecasting. Augmenting VAEs with CNNs sharpens their ability to learn complex temporal patterns—especially in irregular or noisy sequences Desai et al. (2021)

Training and optimisation

Loss functions such as mean-squared error for regression or cross-entropy for classification,

$$\mathcal{L}_{\mathrm{CE}} = -\sum_{c} y_c \log \hat{y}_c,$$

are minimised via optimisers like Adam or SGD with momentum. Hyper-parameters (depth, kernel width, learning rate, batch size, etc.) are tuned by cross-validation or Bayesian optimisation, while L1/L2 penalties, early stopping, and data augmentation mitigate over-fitting.

Hyperparameter Tuning

The performance of a CNN depends critically on hyperparameters such as network depth (number of layers), number of filters per layer, filter size Pooling window size, stride learning rate, optimization algorithm, batch size, amd regularization strength. These hyperparameters are typically optimized through techniques such as grid search, random search, or Bayesian optimization, often using cross-validation to estimate performance on unseen data.

Interpretability and Applications to Astronomical Time Series

Tools for interpreting deep models—most notably feature-visualisation maps, Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al., 2017), and integrated gradients (Sundararajan et al., 2017)—reveal which time segments drive a convolutional neural network's decisions, thereby providing physical insight into the learned representations. Armed with this transparency, astronomers have deployed CNNs across a wide spectrum of time-series problems: detecting transients such as stellar flares, fast radio bursts, and gravitational-wave events in noisy light curves (George & Huerta, 2018); classifying variable stars by light-curve morphology (Naul et al., 2018); recovering exoplanet transits in the presence of complex stellar variability (Shallue & Vanderburg, 2018); discovering pulsars and fast radio bursts in radio dynamic spectra (Connor & van Leeuwen, 2018); identifying quasi-periodic oscillations in X-ray power spectra under strong red-noise backgrounds (Bouchard et al., 2019); and flagging anomalous patterns that may signal new astrophysical phenomena or instrumental artefacts (Pruzhinskaya et al., 2019). In Chapter 4, we will investigate the performance of CNNs in binary classification tasks using Planck data.

Challenges and Limitations

Despite the impressive gains that convolutional neural networks (CNNs) have delivered in astronomical signal classification, several substantive challenges still constrain their routine deployment on time–series data. Foremost among these is the requirement for large, vetted training sets: comprehensive light-curve libraries remain scarce for most rare or fast phenomena—such as fast-radio bursts, unusual super-nova subclasses, or quasi-periodic oscillations—so networks risk over-fitting or learning survey-specific artefacts rather than genuine variability patterns. Even where sizeable catalogues exist, severe class imbalance is typical, with "background'' instances dwarfing bona-fide events by orders of magnitude; practitioners therefore rely on cost-sensitive losses, synthetic minority re-sampling, or focal-loss variants (Lin et al., 2017) to prevent the decision boundary collapsing around the majority class.

Interpretability constitutes a second major hurdle. Gradient-based saliency techniques—e.g. Grad-CAM (Selvaraju et al., 2017) or integrated gradients (Sundararajan et al., 2017)—highlight which temporal segments drive the network's output, yet translating these heat maps into statements about accretion-rate fluctuations or magnetospheric instabilities is rarely straightforward. A related issue is that conventional CNNs return only point estimates. Recent work on Monte-Carlo dropout (Gal & Ghahramani, 2016), deep ensembles (Lakshminarayanan et al., 2017), and full Bayesian neural networks (Neal, 2012) now supplies epistemic and aleatoric uncertainties, but propagating those error budgets into subsequent population-level inferences remains an open problem.

Domain shift further complicates matters. Astronomical light curves are irregularly sampled, heteroscedastic, and often contaminated by survey-specific systematics; hence models pre-trained on natural-image corpora seldom transfer cleanly. Self-supervised contrastive pre-training (Moschou et al., 2023a) and other representation-learning strategies can mitigate data scarcity, but each new facility still tends to require bespoke architectures or extensive fine-tuning, undermining one of deep learning's practical attractions. Finally, although GPU advances and distributed optimisation have eased computational constraints, exhaustive hyper-parameter searches—especially the posterior sampling demanded by Bayesian CNNs—can still be prohibitive for the petabyte-scale streams anticipated from surveys such as *LSST* (Ivezić et al., 2019) and the *SKA* (Braun et al., 2015).

Active research is tackling many of these limitations. Physics-informed neural networks embed conservation laws directly into their loss functions, reducing parameter space and improving interpretability (Raissi et al., 2019; Moschou et al., 2023b). Uncertainty-aware losses encourage calibrated predictive intervals, and the combination of transfer-learning with domain adaptation is beginning to close the gap between laboratory benchmarks and on-sky performance. As these techniques mature, CNNs will complement—rather than replace—the classical spectral-analysis and stochastic-process frameworks developed earlier in this thesis, broadening the methodological repertoire for probing variable high-energy sources.

Chapter 3

A NuSTAR study of quasi-periodic oscillations from the ultraluminous X-ray sources in M82

3.1 Discovery and M82 History

M82, or "Cigar Galaxy", is a starburst galaxy (exhibiting an extraordinarily high rate of star formation) approximately 12 million light-years away in the Ursa Major constellation. As a nearby edge-on galaxy, M82 has a luminosity five times that of the entire Milky Way (Sanders et al., 2003). M82 was discovered in 1774 by Johann Elert Bode and later catalogued in 1781 by Charles Messier (Holmberg, 1958). Initially classified as an irregular galaxy due to having a disrupted morphology, M82 now exists in a disrupted disk galaxy form that strongly interacts with its giant neighbour Messier 81 (M81) (de Grijs et al., 2001). The gravitational interaction with M81 has triggered intense star formation and has sculpted M82's unusual structure, making it one of the best-studied starburst galaxies in the local universe (Mayya et al., 2006).

3.1.1 The Impact of Starburst Activity on the High-Energy Environment

M82's starburst activity is a direct result of its historical gravitational interaction with M81, which is presumed to have triggered a starburst event between 200 and 600 million years ago (Mayya et al., 2006). The resulting supernova explosions and energetic stellar winds have driven large-scale outflows, culminating in a galactic superwind that can be detected across multiple wavelengths, including optical, infrared, radio, and X-ray observations (Strickland & Heckman, 2009). The interplay between starbursts, supernovae, and black-hole accretion creates a complex, energetic environment in M82. Soft, diffuse X-ray emission permeates the galaxy, arising from hot gas heated by these stellar processes and superwind shocks. Meanwhile, small, point-like hard X-ray sources in M82—mainly the ULXs—stand out against this spread-out background, highlighting sites of extreme accretion activity amid the starburst-driven environment.

3.1.2 Early Observations and Multi-Wavelength Studies

Early optical observations revealed M82 to be a highly perturbed system [[has a complex structure]. Subsequent multi-wavelength studies across radio, infrared, and X-ray

bands established M82 as a prototypical starburst galaxy. Infrared observations by the *IRAS* satellite and later by *Spitzer* confirmed an exceptionally high rate of star formation concentrated in the galaxy's core. In X-rays, *Einstein Observatory* and later missions detected unusually bright point sources in M82, hinting at the presence of ultraluminous X-ray sources (ULXs). Over time, coordinated observations have linked the galaxy's intense starburst to its X-ray output: young supernova remnants and ULXs contribute significantly to the hard X-ray emission, while hot plasma from superwinds produces extended soft X-ray glow. These early studies set the stage for focused investigations of M82's compact X-ray sources.

3.1.3 X-ray Binary Systems in M82

In X-ray binaries, the electromagnetic radiation emitted by accreting matter can reach extreme luminosities. M82 hosts several ULXs, which are exceptional X-ray binaries exhibiting luminosities above the Eddington limit for typical stellar-mass black holes (i.e., $L_X \gtrsim 10^{39}$ erg s^{-1}). The galaxy's starburst central region provides a fertile ground for the formation of such systems, including neutron stars and black holes accreting at high rates. Below we summarize the key known ULXs of M82 and their relevance to understanding accretion physics in this galaxy.

M82 X-1 : Evidence for an Intermediate-Mass Black Hole

M82 X-1 represents one of the most promising candidates for an intermediate-mass black hole (IMBH). Timing analysis of its X-ray emission has provided major evidence in favor of this hypothesis (Strohmayer & Mushotzky, 2003).first reported a QPO at 54 mHz associated with M82 X-1 using XMM-Newton data. Building on this research, (Mucciarelli et al., 2006).reported QPOs spanning between 50–166 mHz using combined XMM-Newton and RossiXTE archival observations. The most compelling evidence for M82 X-1 harboring an IMBH came from (Pasham et al., 2015). who detected stable twin-peak QPOs with a 3:2 frequency ratio (3.32 $\pm$ 0.06 Hz and 5.07 $\pm$ 0.06 Hz). This frequency pattern, which is characteristic of high-frequency QPOs in stellar-mass black hole binaries but scaled down to lower frequencies—indicates a black hole of approximately 400 $M_\odot$. This places M82 X-1 firmly in the intermediate-mass range, representing a critical "missing link'' between stellar-mass and supermassive black holes.

M82 X-2: Challenging ULX Paradigms

Prior to 2014, the astronomical community largely assumed that ULXs were accreting black holes. This view was challenged when Bachetti et al. (2014) unveiled the nature of M82 X-2, the second-brightest X-ray source in M82, as an X-ray pulsar. M82 X-2 was found to be a neutron star accreting matter at super-Eddington rates, with coherent pulsations implying the presence of a strong magnetic field. This discovery demonstrated that neutron stars can power ULX-level luminosities, overturning the paradigm that ULXs necessarily host black holes. The detection of pulsations from M82 X-2 also provided key insights: the measured pulse period (of 1.37 seconds and a pulse fraction $\sin 20\%$) and its derivatives allowed constraints on the neutron star's

magnetic field and spin-up rate due to accretion torque, offering a rare opportunity to study magneto-hydrodynamic processes in an extreme accretion regime.

M82 as a Key Laboratory for Astrophysics

M82 continues to be a valuable target in astrophysics across a wide wavelength regime. With its unparalleled combination of energetic starbursts, a rich population of compact objects, and superwinds extending to large scales, it provides an ideal environment for studying the interplay between stellar evolution, black hole growth, and mechanisms of galactic feedback (Zhang et al., 2019). Ongoing and future investigations utilizing next-generation observatories—including the *James Webb Space Telescope (JWST)*, the Advanced Telescope for High Energy Astrophysics (ATHENA), and the Square Kilometre Array (SKA)—promise to further illuminate M82's population of black holes, the physics of ULXs, and the dynamics of starburst-driven outflows (Padovani et al., 2017). In particular, M82's ULXs serve as natural laboratories for extreme accretion: continuing to monitor and model these sources will help bridge our understanding between stellar-mass X-ray binaries and more massive accretors.

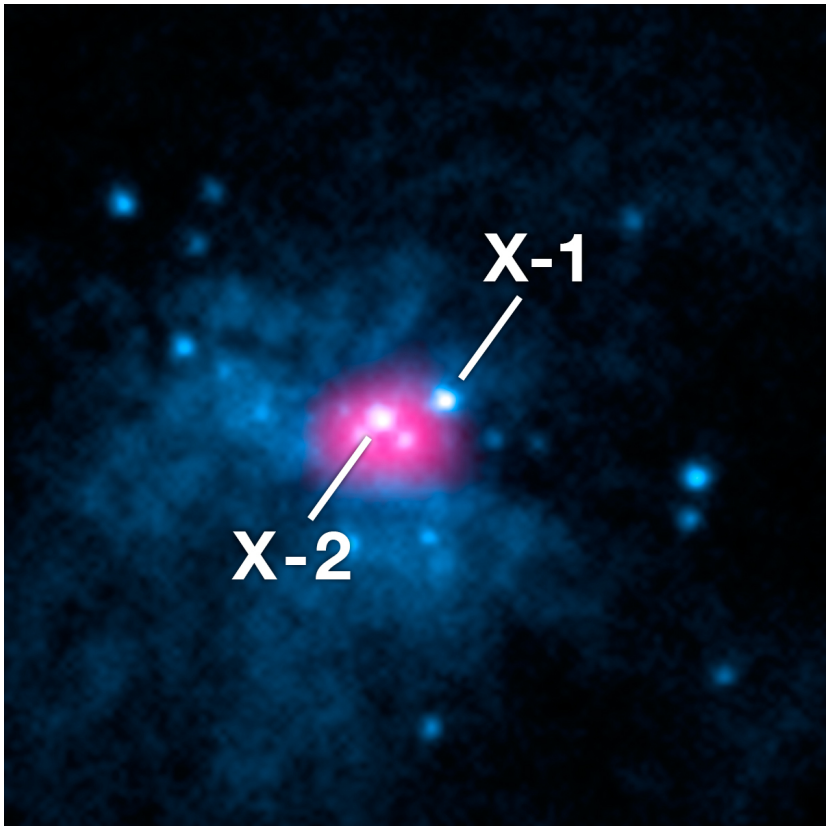


Figure 3.1: A composite high-energy X-ray image of the core of M82, combining *NuSTAR* data (shown in pink) with lower-energy X-rays from *Chandra*. The two brightest compact X-ray sources, M82 X-1 and M82 X-2, are indicated. ULXs are regions that shine intensely in X-rays, and in M82 these lie in the starburst core.

3.1.4 Quasi-Periodic Oscillations: Theoretical Background

One of the most intriguing phenomena in the accretion processes of X-ray binaries is the presence of quasi-periodic oscillations (QPOs). These oscillations are manifested as relatively narrow peaks in the X-ray power spectrum of accreting neutron stars and black holes (Ingram & Motta, 2019; Van der Klis, 1989). QPOs are typically characterized by a Lorentzian profile with three key parameters: the centroid frequency, the width (inversely related to the coherence time of the oscillation), and the strength (the integrated power, proportional to the variance contributed by the QPO) (Motta et al., 2017).

The study of QPOs provides crucial insights into the fundamental physics of compact objects, including their masses, radii, and spin rates (Kluźniak, 2006). Low-frequency QPOs (LFQPOs), with frequencies up to tens of Hz, have been observed in the majority of black hole X-ray binaries (e.g., up to ∼30 Hz in black hole binaries and ∼60 Hz in neutron star binaries (Belloni et al., 2005; van der Klis, 2006). High-frequency QPOs (HFQPOs, sometimes called kHz QPOs in neutron star systems) occur at frequencies between 100 and 500 Hz in BHs to over 1kHz in NSBs with weaker magnetic fields (Motta et al., 2017). These variations highlight differences in the accretion environments and compact object properties of black holes versus neutron stars, as reflected in their distinct variability patterns. By studying QPOs, researchers gain insight into the inner accretion disk dynamics and strong-gravity effects near the compact object (Remillard & McClintock, 2006; Belloni & Stella, 2012; Zhang et al., 2006). For example, certain QPO frequency ratios or evolutions may indicate specific physical mechanisms such as relativistic precession or resonance in the inner disk (Ingram et al., 2009; Motta et al., 2016). Thus, QPOs serve as important diagnostics of the conditions and processes in the immediate vicinity of compact objects.

A significant challenge in studying X-ray variability in M82 is attributing specific QPO signals to their source ULXs and separating the contributions of M82 X-1 and M82 X-2. This difficulty arises because most X-ray telescopes (including *NuSTAR*) have insufficient spatial resolution to distinctly resolve the two ULXs in the crowded starburst core. As a result, any variability detected in integrated X-ray lightcurves of M82 could originate from either one of these sources (or a combination of both), complicating the interpretation of timing features.

In this work, we analyze observational data acquired with NASA's *NuSTAR* () satellite spanning from 2014 to 2024, to investigate the variability of the X-ray flux in the M82 galaxy. to assess consistency of previous works and to identify new insights.

3.2 Data reduction

We used data from all available ustar observations of the M82 galaxy between 2014 and 2024, reduced with the standard pipeline (`nupipeline`) from the High Energy Astrophysics Science Archive Research Center (HEASARC), and selected photons from a region of 70″ around the position of the source. Our data consist of individual events that are timestamps of when a photon reached the detector and their associated properties, such as the photon energy. Each timestamp was corrected from local time to the barycenter of the Solar System using the `barycorr` FTOOL, using the ICRS

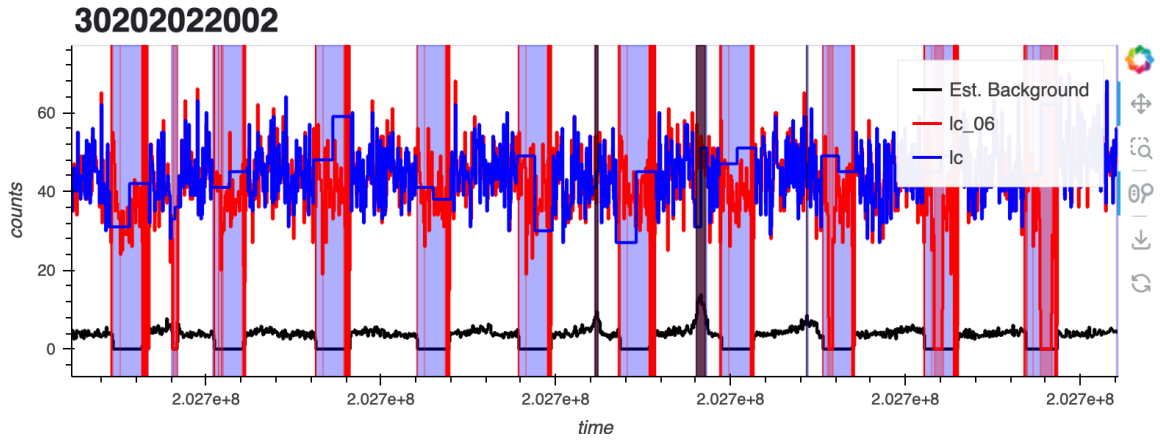


Figure 3.2: Example of the GTI cleaning process. We plot a rescaled background light curve and the source light curve, and eliminate intervals where the background light curve is above 10 % of the mean level. Vertical black bands indicate these bad intervals, while blue bands indicate the standard bad time intervals due to occultation or poor star tracker coverage. Red data points are taken in intervals with poor star tracker coverage. We will not consider them when modeling the PDS, but we will use them to improve the detectability of unseen features.

coordinates of M82 X-2 09:55:51.040 +69:40:45.49[1] (Kaaret et al., 2006) and using the DE430 JPL ephemeris[2].

For the power spectral fitting, we only used events from good time intervals and with full star tracker visibility (mode-01 data). For a limited part of the analysis, we made use of photons from intervals with limited star tracker coverage (the SCIENCE_SC mode, or mode-06, data). We used the tool `nusplitsc` to split the mode-06 data in interval with single star tracker combinations, we moved the extraction region for events in each sub-interval to adapt to the position of the PSF centroid, we extracted the events from the source region, and finally we merged the events together with the mode-01 data. Given the circumpolar position of the source and the close-to-equatorial orbit of *NuSTAR*, this allowed to almost double the number of photons available for sensitive QPO searches, at the cost of increasing the red noise level. Therefore, we only used these data to increase the sensitivity to the search of new features at high frequencies where the red noise was negligible.

We ran the analysis using Stingray (Huppenkothen et al., 2019), a Python library built to perform time series analysis, providing implementations of the most advanced spectral timing techniques available in the literature.

Raw data do not provide the photon energy directly. However, in *NuSTAR* the energy channel number (PI) and the central energy E_p of the channel are related by the simple formula $E_p(\mathrm{keV}) = 1.62 + 0.04\,\mathrm{PI}$. This is accounted for automatically when loading data in Stingray.

[1] Data were barycentered in order to allow the study of aperiodic variability from M82 X-1 and pulsations from M82 X-2. A 5″ mismatch is irrelevant for the study of slow variability, but it might be detectable in precise pulsar timing.

[2] https://naif.jpl.nasa.gov/pub/naif/generic_kernels/spk/planets/aareadme_de430-de431.txt

We plotted light curves of the source region of all observations, and background light curves containing all photons further than 100″ from the source. This criterion is different from the one usually employed for spectral analysis, where a background region is chosen in a large circular region of the FOV devoid of sources. In our case, these background light curves were meant to catch flares in the background, corresponding to anything from particles hitting the telescope or other increased environmental background (e.g. approaching the SAA), and we needed to gather all the photons we could at a reasonable distance from the source region. We rescaled the background light curve by multiplying by the ratio of pixels inside the source and background regions. Whenever the rescaled background light curve had flares reaching more than $\sim 10\%$ of the source mean flux, we rejected that time interval and excluded it from the good time intervals (GTI). See Fig. 3.2 for an example.

3.3 Periodogram production and filtering

Our source light curve is produced by sampling the photons from the source region falling into equispaced time intervals with resolution $t_{\rm samp}$. Alongside the source light curve, we create synthetic visibility light curves with the same sampling interval that are equal to the mean counts per bin of the source light curve during GTIs and 0 outside. We take a periodogram of this synthetic light curve, that shows strong features corresponding to the missing data while going rapidly to zero at high frequencies, because it does not contain noise. In this periodogram, we can set a threshold, and single out the strongest peaks, to be used for notch filtering. The idea is that we will eliminate these frequencies from the final periodogram, and then make a geometrical rebinning (as it is commonly done) that will average the remaining nearby bins, maintaining the statistical properties of the periodogram.

However, if we take the source light curve at face value, using 0 outside GTIs, we can see that if one applies the notch filter above straight away, a large number of powers do not follow the expected distribution (Figure 3.3, middle panels). This is because some noise is still leaking at frequencies near the blacklisted ones.

But we can also fill the bad time intervals of the light curve not with zeros, but with the mean of the data. This produces a much smaller effect of the visibility windows on the final periodogram, creating an almost flat distribution of power, and the notch filter will just be an additional cautionary measure (Figure 3.3, bottom panels, black data). In a real-life situation, when there is some long-term source variability, filling band intervals with the mean will not be as clean as this example, so the notch filtering will be useful. There is an additional measure to take, however. The periodogram is a measure of the variance of the data, and it is calculated from an FFT which contains a division by the total number of data points. However, the filled data points do not contribute to the variance, and this means that the calculated power will be lower than the expected value. To reinstate the correct normalization, we need to multiply this periodogram by $n_{\rm tot}/n_{\rm gti}$, where $n_{\rm tot}$ is the total number of bins and $n_{\rm gti}$ the number of bins in GTIs. (Figure 3.3, bottom panels, magenta data).

To verify that the method does not alter the response of the periodogram in a frequency-dependent way, we performed the following test: we generated synthetic

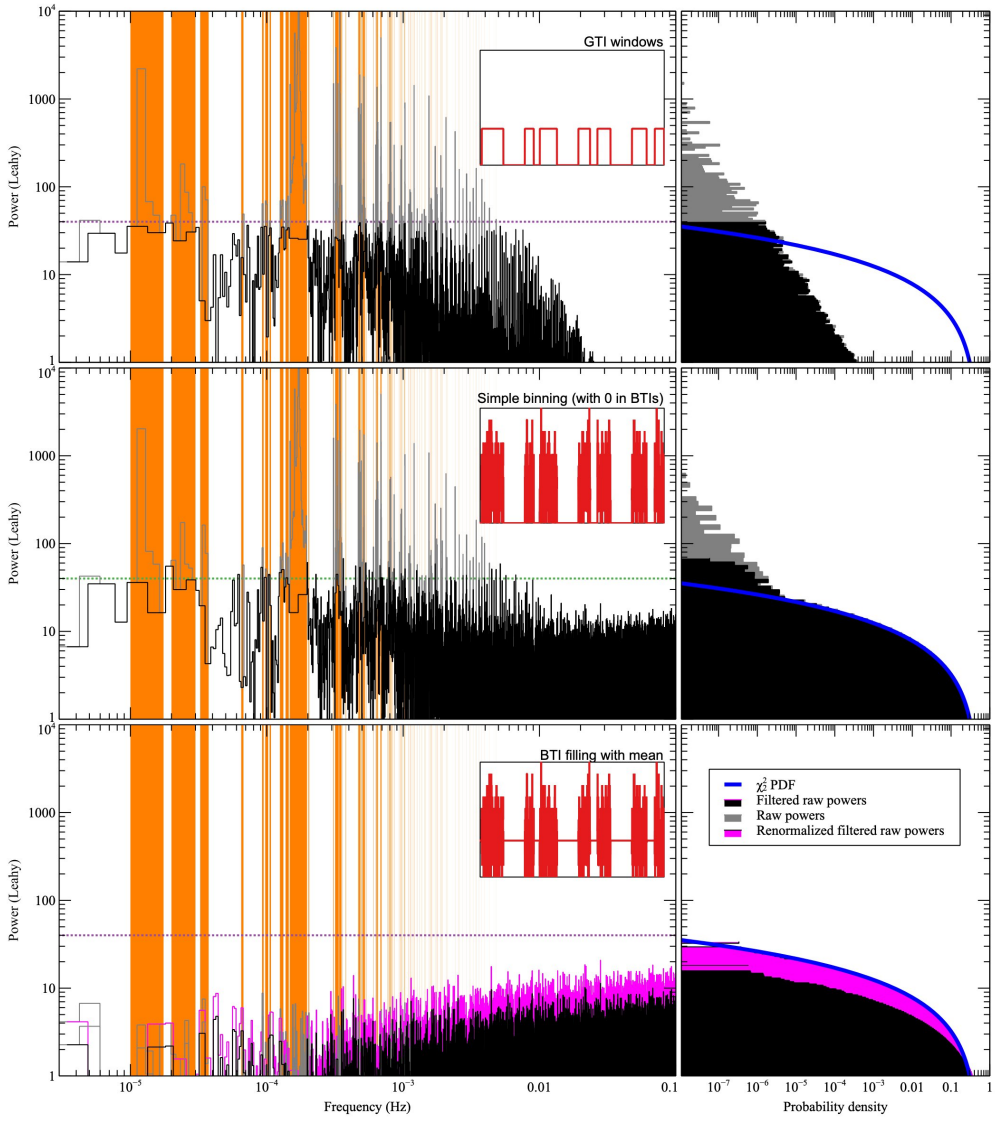


Figure 3.3: Procedure to filter periodograms from the effect of visibility windows described in Section 3.3, using simulated Poissonian data (so, no source variability) with the same mean count rate and GTIs of ObsID 80002092006. (Top) Periodogram of the visibility light curve, showing the features corresponding to the missing data. We set a threshold and select a number of bad frequency intervals (orange) to be black listed. (Middle) Periodogram of the binned light curve with 0 outside GTIs: it contains many of the same features, with similar powers, plus the expected white noise from the data. After notch filtering, some powers clearly exceed the expected χ^2_2 distribution. (Bottom) Periodogram of the binned light curve with the mean counts per bin used as a filler outside GTIs instead of 0. Most of the features disappear from the periodogram even before notch filtering, and the powers follow the correct distribution, but with the wrong normalization, which is corrected as described in the text

light curves with a sinusoidal modulation at 10% fractional amplitude, using the same GTIs and mean flux as Obsid 80002092006. The light curve had no Poisson noise, only the smooth sinusoidal modulation. Outside GTIs, we filled the time intervals with the mean value of the flux as we did for the data. We repeated the experiment for 1000 frequencies distributed log-uniformly over the range $10^{-4} - 1$ Hz. We always used frequencies extracted from the grid of the periodogram in order to avoid the expected sinc^2 response degradation when moving away from the centre of the frequency bin (Van der Klis, 1989). We applied the same notch filters as the real data, and we measured the decrease of rms (hereafter, the damping factor) in the remaining frequencies compared to an uninterrupted pulsation over the whole observation. The decrease of rms was the same at all frequencies and exactly what was expected from the fact of having missing data (i.e., outside GTIs there is no variability, and the total variance decreases). Only very close to the notch-filtered frequencies, we noticed some "wiggles" in the damping factor, by 10% at most.

3.4 Timing analysis

3.4.1 Statistical properties of the periodogram

Periodogram-based methods are commonly used as a parametric approximation of the power spectrum. Given the Discrete Fourier Transform (DFT) components a_i of a light curve x_n $(n = 0, \ldots, N)$ (see Eq. 3.1)

$$a_i = \frac{1}{N} \sum_{n=0}^{N-1} x_n e^{-j2\pi ni/N}, \quad (i = -N/2, \ldots, N/2 - 1) \tag{3.1}$$

the periodogram is defined as $P_i = |a_i|^2$.

In X-ray Astronomy, it is common to use the normalization from Leahy et al. (1983), where the periodogram defined above is multiplied by a factor $2/N_{\mathrm{ph}}$, where N_{ph} is the number of photons in the light curve. With this normalization, the powers of a periodogram of pure white noise follow a χ^2_2 distribution, allowing for an easy identification of outliers. A common procedure to limit the noise of the periodogram is indeed based on power averaging, either of W nearby bins from the same periodogram, or M periodograms from different segments of the data (the so-called Bartlett periodogram, from Bartlett 1950b), or using both methods. It is easy to demonstrate that the effect of averaging MW noise powers leads to a normalized χ^2_{2MW}/MW distribution, which resembles more and more a Gaussian distribution with width $\sigma = 2/\sqrt{MW}$ (Van der Klis, 1989) as the number of averaged powers increases.

The Bartlett periodogram has the additional major advantage of being applicable to observations containing missing data, for example due to Earth occultation, South-Atlantic Anomaly passes, high background, and so on. In X-ray observations, these "bad intervals" are usually eliminated from the observations during the data reduction procedure, and good observing conditions are encoded in a Good Time Interval (GTI) list in the same FITS files of the data. The Bartlett periodogram can then be chosen so that one or more intervals of duration t_{seg} fit inside the typical length of GTIs.

The powers I_j of a periodogram containing signal are distributed following a χ^2_{2MW} around the real spectrum S_j (Barret & Vaughan, 2012).

$$I(f_j) = \frac{S(f_j)}{2MW} X \tag{3.2}$$

where X follows a χ^2_{2MW} distribution.

From the properties of the χ^2_{2M} distribution, the log-likelihood formula in this general case can be derived as follows:

$$\log \mathcal{L} = \frac{-\nu}{2} \sum_{j=1}^{N-1} \left\{ \frac{I_j}{S_j} + \ln S_j + \left(\frac{2}{\nu} - 1 \right) \ln I_j + c(\nu) \right\} \tag{3.3}$$

where $\nu = 2MW$ is the number of degrees of freedom and $c(\nu)$ is a constant for fixed ν. Since minimization algorithms are more common than maximization ones, maximum-likelihood fitting procedures usually consist of minimizing the quantity $-2 \log \mathcal{L}$.

However, the Bartlett periodogram calculated from a typical Fast Fourier Transform (Cooley & Tukey, 1965) has a major limitation in terms of frequency resolution, limited to $\Delta\nu = 1/t_{\text{seg}}$. This also means that we cannot investigate frequencies lower than $\Delta\nu$. Some techniques to improve the frequency resolution, such as interbinning or Fourier interpolation (Ransom et al., 2002), can be used, but at the cost of altering the statistical properties of the data and making the fitting and interpretation of results less robust.

On the other hand, a single periodogram of the whole time series has a much better frequency resolution and sensitivity to low frequencies (as now the resolution is $\Delta\nu = 1/t_{\text{obs}}$), but it is affected by the missing data and contains a large number of (typically) low-frequency peaks that correspond to the variability introduced by the visibility windows. If the bad time intervals (BTIs) are very small (e.g. less than 1% of the data), it is customary to add some white noise to fill-up the intervals and have a final periodogram with all the desired statistical properties, at the expense of some minor loss of sensitivity. However, in our observations, the bad intervals are comparable in length with the good ones, which would imply simulating about half the data, which is unacceptable for our purposes.

For this work, we devised a treatment for the periodogram that limits the effects of windowing while maintaining most of the statistical properties of the periodogram (as detailed in section 3.3).

3.4.2 Model construction

The periodograms of X-ray binaries show a variety of variable phenomena, and can be conveniently modeled through a composition of Lorentzian components (Belloni et al., 2002). For these functions, we use the definition:

$$P(\nu) = \frac{A_0 \left(\frac{w}{2}\right)^2}{(\nu - \nu_0)^2 + \left(\frac{w}{2}\right)^2} \tag{3.4}$$

where ν_0 is the centroid frequency or the frequency at the peak of the signal, w represents the full width at half maximum (FWHM), and A_0 is the amplitude of the signal.

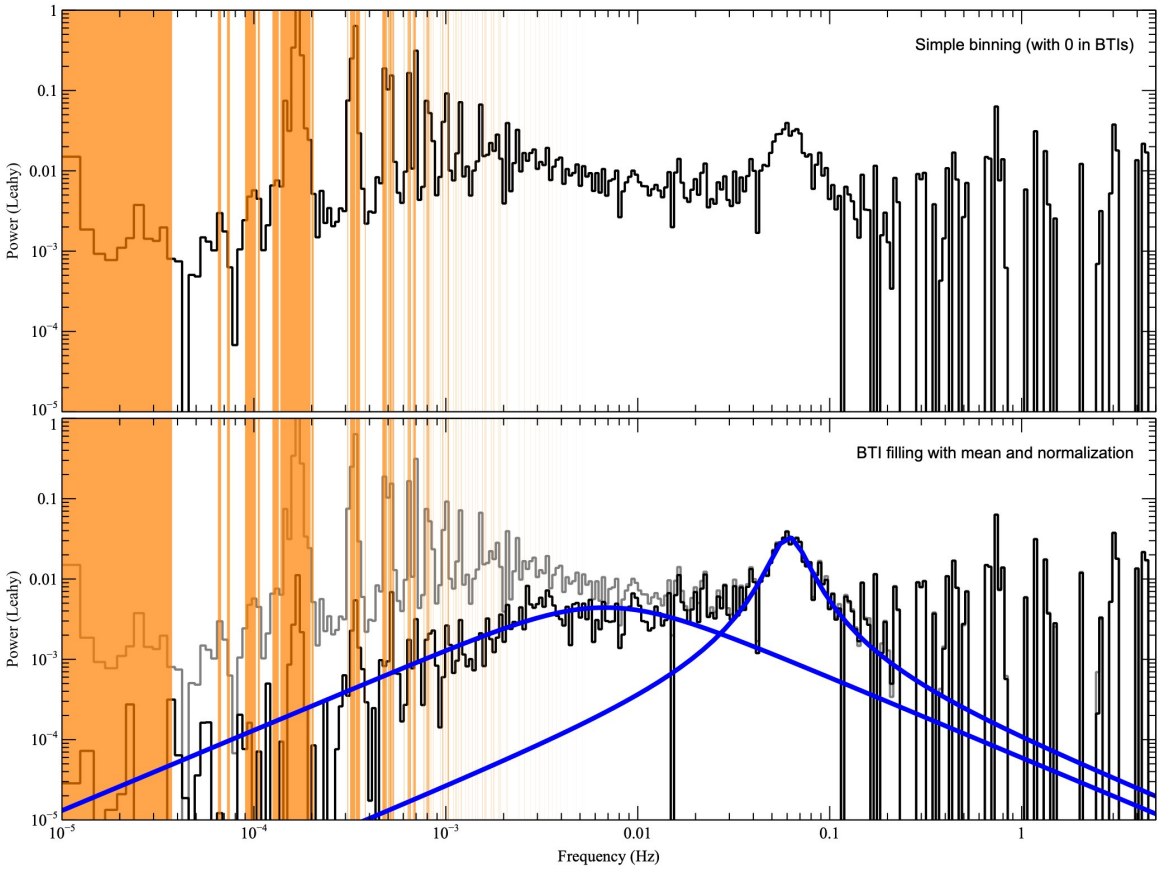


Figure 3.4: Example analysis using obsid 80002092006. Data are cleaned as described in Section 3.3. Orange bands show the notch-filtered frequencies. The top panel shows the periodogram of the light curve when BTIs contain zeroes, and the lower panel shows the result when BTIs are filled with the mean counts per bin. It is clear that notch filtering is still needed in this non-ideal case. The two blue curves are the two best-fit Lorentzians for the red noise component and the QPO.

The use of Lorentzian components is mostly phenomenological given the symmetry properties of these functions, even though it is originally rooted in the fact that Lorentzians are the Fourier transforms of exponentially decaying oscillations, a common phenomenon in nature.

These Lorentzian components can be characterized through three main quantities. First of all, the **characteristic frequency** ($\nu_{\max}$):

$$\nu_{\max} = \sqrt{\nu_0^2 + \left(\frac{w}{2}\right)^2},$$

that represents the maximum of the Lorentzian in a νP vs ν plot and the frequency at which the Lorentzian contains the most power per logarithmic frequency interval (Belloni et al., 2002). This quantity is very close to ν_0 for coherent QPOs, while it departs considerably from it for broadband red-noise components. It is particularly useful when considering the evolution of timing features, as it is common to observe a broadband noise component evolve into a QPO. The characteristic frequency, in this case, makes a smooth transition that would not be as clear when using the central frequency (e.g. Motta et al., 2016). The second important quantity is the **Quality factor (Q)**: Defined as $Q = \frac{\nu_0}{\text{HWHM}} = \frac{2\nu_0}{w}$, it measures the signal's coherence[3]. Commonly, Lorentzian components with $Q > Q_{lim}$ are classified as QPOs, while those with $Q < Q_{lim}$ are considered broadband peaked noise. Q_{lim} used in literature is variable. In this work we will use $Q_{lim} = 2$ as the boundary between a QPO and a broadband component. However, it will become clear that the feature we identify as QPO can sometimes have a low coherence. Finally, we can define the **(fractional or absolute) root-mean-squared (rms) amplitude**: A measure of the signal's strength, which depends on the source flux. It is proportional to the square root of the integrated power contributed by the QPO to the periodogram. Using eq. 3.4, the rms amplitude can be calculated as the square root of the integral of $P(\nu)$ normalized in the desired rms units (fractional à la Belloni & Hasinger 1990 or absolute, in counts per second). Since power is only calculated at positive frequencies, assuming $P(\nu)$ was fit in Leahy normalization, the rms amplitude can be calculated as:

$$\text{rms} = \sqrt{F \int_0^\infty P(\nu) d\nu} = \sqrt{A_0 F \left(\pi/2 - \tan^{-1} \frac{-\nu_0}{w/2}\right)} \tag{3.5}$$

where F is the conversion factor between the Leahy normalization and the desired rms units.

3.4.3 Inference

Our periodogram modeling consists of two main steps: a maximum likelihood estimation (MLE) of the best fitting model, following Barret & Vaughan (2012), and a parametric bootstrap technique to evaluate uncertainties. The MLE is done with the `stingray.modeling` package (Huppenkothen et al., 2019). The procedure consists

[3]Many works in the literature divide by the full width at half maximum, so their values of Q would be half ours.

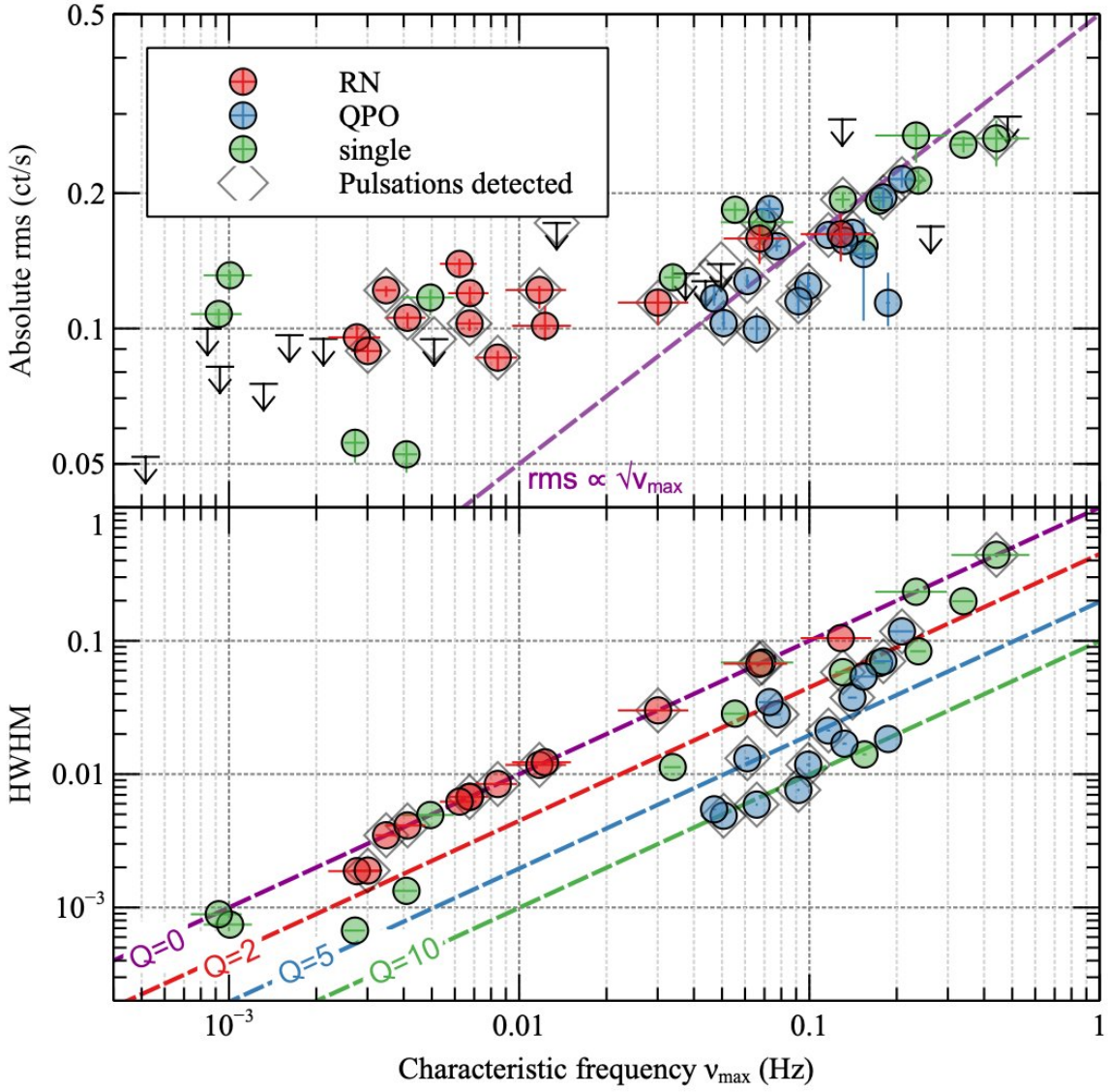


Figure 3.5: (Top) rms versus frequency for the power spectral features fit in Section 3.1 and (Bottom) HWHM versus frequency for the same features. There does not seem to be a systematic pattern in the appearance of these timing features and pulsations. We calculated 1-σ uncertainties and 3-σ, upper limits through the bootstrap procedure in Section 3.4.3. The identification of the two features is often tricky if only one of them is present in the data, but simple criteria to distinguish them seem to emerge from this visualization, with the QPO, regardless of its Q factor in a given observation, following a rms $\propto \nu^{1/2}$ law and generally having a frequency above 0.02 Hz.

of minimizing the negative loglikelihood in Equation 3.3, using a model power spectrum. We employ an optimization algorithm that supports bounds, like the limited-memory Broyden–Fletcher–Goldfarb–Shanno scheme (L-BFGS; Byrd et al. 1995) to maximize the likelihood function, thereby obtaining parameter estimates. We fit the periodograms with one or two Lorentzians. When in doubt about the most appropriate number of components, we use the Akaike Information Criterion (Akaike, 1974) to determine whether a model is best described by one or two Lorentzians. For the Lorentzian components, we use an Astropy model (`Lorentz1D`) with the same parameters as Equation 3.4, and we fit an additional constant with starting value 2, the expected white noise level in the Leahy normalization. For the initial values of the Lorentzians, we use an interactive interface that allows to create a reasonable starting model, but then we do not set boundaries to the parameters other than being positive definite, and the FWHM being more than 0.001 Hz.

Once we obtain a best-fit model, we use a parametric bootstrap to evaluate the uncertainties, which involves the following steps:

1. Generate random powers: Randomly simulate powers from the best-fit model, ensuring that they follow a distribution of $\chi^2_{2M}/2M$, where M is the number of averaged powers in each bin of the original periodogram. These simulated powers should scatter around the best-fit model.

2. Fit a model to the random powers, starting from random parameters distributed within 10% of the real parameters.

3. Bootstrap procedure: Repeat steps 1–2 1,000 times, recording the fit parameters in every iteration (a bootstrap distribution is obtained for each parameter)

4. Parameter estimates: use the appropriate percentiles from the bootstrap results of the parameters to compute relevant statistics such as means, two-sided standard errors, and confidence intervals around estimates. These provide information on the uncertainty associated with the parameter estimates based on the simulated powers. To take into account cases where the fit swaps the two Lorentzian components, we always order them by central frequency ν_0 and, if they both go to zero, by characteristic frequency $\nu_{\max}$. To reject outliers, we calculate the median and reject points at more than 5 median absolute deviations from the median.

Table 1 presents the complete set of best-fit parameters derived from our modeling of the PDS described above (see also in the Appendix 5.4 a table that prensent the parameter inferences of the model). Figure 3.5 shows the variation of the total absolute rms of the features with frequency. The points with upper limits refer to models where the AIC criterion suggested an advantage in adding a model component, but the bootstrap procedure returned a 3-σ confidence interval including zero. We use total rms and not the more customary fractional rms because the total X-ray flux of M82 is the combination of many X-ray sources, and in particular of both M82 X-1 and M82 X-2. Therefore, the fractional rms of the features will change randomly based on which source is more luminous during each observation, while the absolute rms retains a physical meaning in terms of total variable luminosity, in counts/s.

From this visualization, two distinct features are clearly identified. One at frequencies below ~ 0.02 Hz whose rms does not depend on frequency, and one above, with a clear correlation between rms and frequency, that can confidently be identified with the QPO from M82 X-1 studied by Strohmayer & Mushotzky (2003); Dewangan et al. (2006); Mucciarelli et al. (2006) (see below for an additional test of the association). The increase of absolute rms with frequency follows an approximate $\nu^{1/2}$ law, which might imply that the frequency is increasing with increasing flux of M82 X-1 (assuming a constant *fractional* amplitude).

3.4.4 Energy dependence

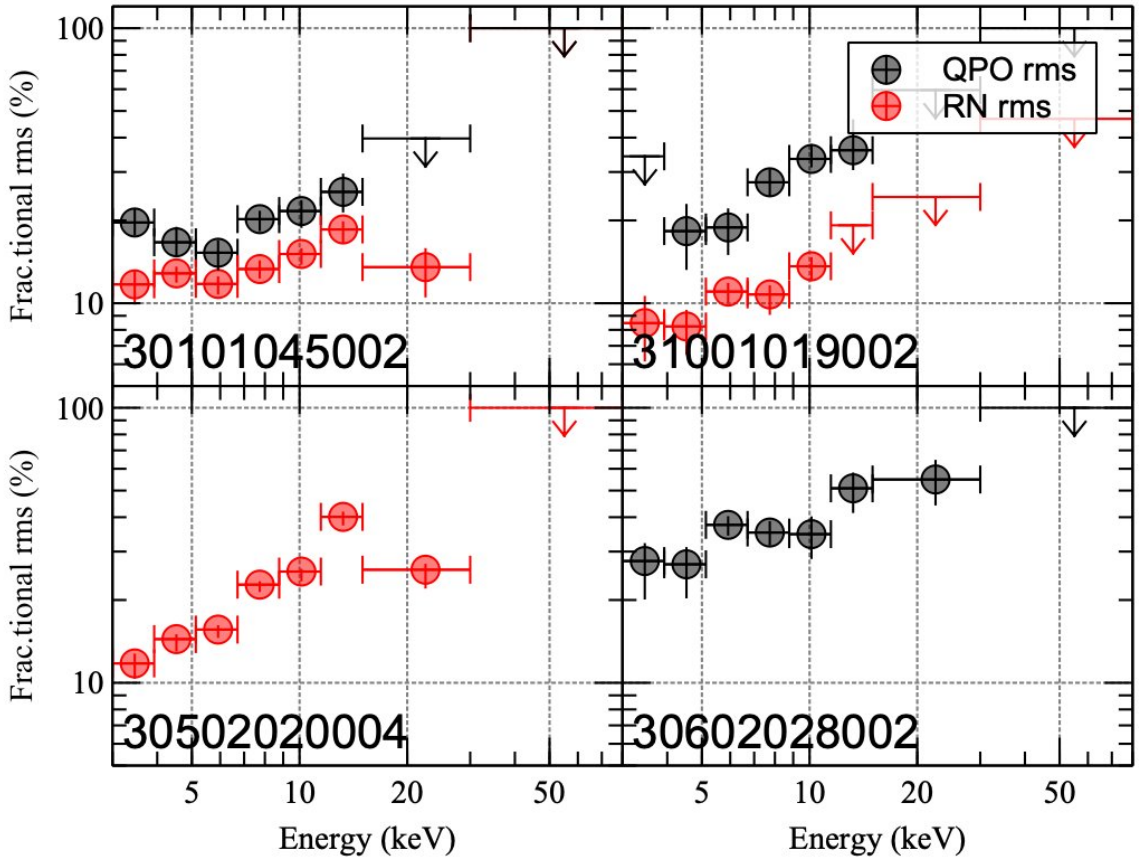


Figure 3.6: Fractional rms of the red noise and the QPO in different obsids. Horizontal bars with caps represent exact energy ranges, vertical bars are 1-σ uncertainties. All variability is more significant at higher energies

It is interesting to investigate how the features evolve with energy. We divide the 3-80 keV energy band into 8 intervals whose width follows approximately a geometric sequence but with larger intervals at higher energies to account for the very small number of high-energy counts. We re-fit the model in each energy band; however, in most observations, the fit is not robust enough to leave all parameters free. Since there is no evidence of changes in the shape of the Lorentzian components with energy, only their normalization, and the Poisson noise level does not depart significantly from 2,

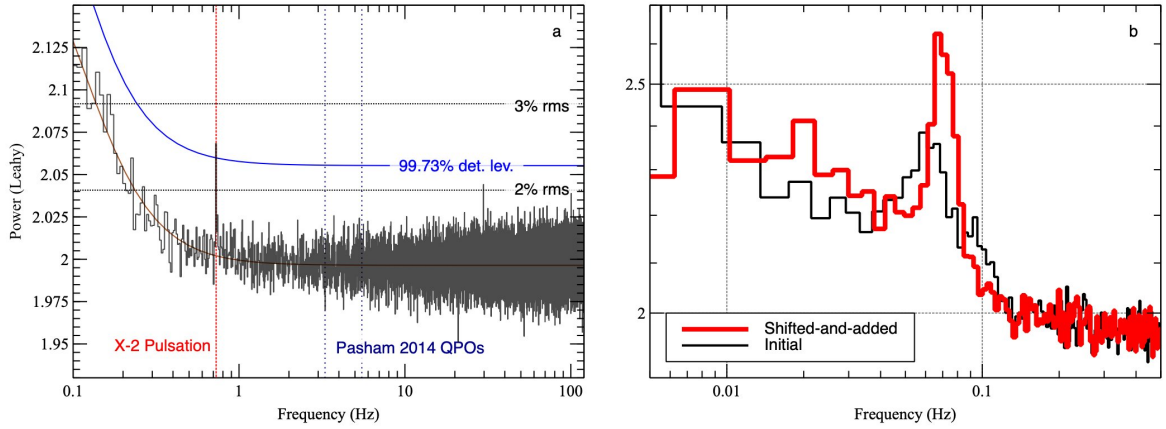


Figure 3.7: **(a)** Averaged PDS of the full ∼5 Ms *NuSTAR* data set on M82. The 3–79 keV light curve, sampled at $\Delta t = 0.1$ s, was divided into ∼5 000 non-overlapping segments of 1024 s duration. Each segment's Leahy-normalised periodogram was computed and the set was averaged to form the final PDS using the Bartlett estimator. This procedure is consistently applied to all PDS shown in this paper. The solid blue curve marks the 99.73% (3σ) single-trial detection level, computed from the cumulative χ^2 distribution with $2M$ degrees of freedom (where M is the number of averaged segments) and scaled by the best-fitting broadband noise model (see Section 3.1). The strong feature at ∼0.73 Hz corresponds to the 1.37 s pulsation of M82 X-2. The absence of a peak near 30 Hz sets an upper limit of $\lesssim 2\%$ rms on any quasi-coherent signal at that frequency. Horizontal dotted lines mark 2% and 3% rms amplitudes. **(b)** Result of the shift-and-add technique applied to the ∼50 mHz QPO. Individual 1024 s PDS were aligned on the QPO centroid measured in each segment and co-added (red histogram). The resulting peak has a quality factor $Q \approx 8$. No significant harmonic is detected at twice the frequency. The black histogram shows an individual PDS for comparison.

we fix all parameters but the amplitudes of the one/two Lorentzians to their best-fit values from the total-flux analysis. Through repeating the bootstrap procedure, we get sensible values for the amplitudes and their uncertainties, and we can calculate the rms at different energies. Since we are proceeding on an observation-by-observation basis, we calculate the *fractional* rms, as we are only interested in the relative variation of rms with energy. The results are shown in Figure 3.6, and clearly show that the rms increases with energy for all components.

3.4.5 Search for other QPOs and harmonics

Pasham et al. (2014a) reported the detection of 3- and 5-Hz QPOs from M82, most probably from M82 X-1, using *RXTE* data. We now know that there are ∼15 ULXs between the galaxies M82, M81 and the satellite Holmberg IX (Walton et al., 2022), all within the 1^o field of view of the *RXTE*/PCA instrument used to detect the QPOs. Despite M82 X-1 being the brightest at its maximum, all bright ULXs in the field, including M82 X-1, M82 X-2, Holmberg IX X-1, are known to be transient. This leaves open the possibility that M82 X-1 is not the source of the twin QPOs, and generally implies that the 3–5% rms limit measured by Pasham et al. (2014a) is underestimated.

Reproducing the *RXTE* twin-QPO result with *NuSTAR* is challenging for several

reasons. Although the focused optics greatly reduce background compared with *RXTE*, the *NuSTAR* point-spread function still encloses M82 X-1, M82 X-2, and at least one additional ULX ("X-3"), and its effective area is an order of magnitude smaller. With these caveats in mind, we averaged *all* M82 observations discussed above. Because the PDS is relatively clean above 1 Hz, we included even intervals with poor star-tracker coverage (Section 3.2); the circumpolar location of M82 almost doubled the usable exposure time, improving the detection sensitivity by $\sim\sqrt{2}$.

Pulsation check. The 1.37 s pulsation of M82 X-2 is recovered at $\simeq 0.73$ Hz *without applying barycentric, orbital, or $\dot{P}$ corrections*. The Doppler smearing across each 1024-s segment amounts to < 0.3 cycle, so the peak remains well above the 3σ detection threshold even in the uncorrected PDS. As shown in Figure 3.7, no significant feature appears at frequencies above the pulsation, and we constrain any quasi-coherent signal near 30 Hz to $\lesssim 2\%$ rms.

Additionally, we looked for harmonics of the 50-mHz QPO, in order to compare it to other known classes of QPOs such as Type-C from BH LMXBs. We used the shift-and-add technique (Barret et al., 2005) on the longest observing span available, the series of ObsIDs 80002092002..11 from 2014. We split the observation in 128 s intervals, calculated the periodogram in each, then averaged 25 such intervals to gain in signal to noise, creating a series of periodograms, each from 3.2 ks of non-contiguous data. These were sufficient to track the evolution of the QPO frequency $\nu_{\rm QPO}$ during the observations. We calculated the average QPO frequency $\bar{\nu}_{\rm QPO}$ and then shifted each periodogram by an amount corresponding to the $\nu_{\rm QPO} - \bar{\nu}_{\rm QPO}$, obtaining the average shape of the QPO. This technique was used successfully to characterize kHz QPOs in NS LMXBs, and in some cases even to discover the upper kHz QPO (Barret et al., 2005). We also used a variation of this technique: since we were not looking for a frequency following a parallel track but for a harmonic, we shifted the periodogram around $2\bar{\nu}_{\rm QPO}$, by a double amount. We did not find evidence for a harmonic of the QPO.

Symmary Tables

Obsid	MJD	Exposure (ks)	Obs. conditions	Phenomenology
80002092002	56680.5	67.4		QPO, pulsations
80002092004	56682.8	92.3	Bkg Flaring	QPO, pulsations
80002092006	56685.5	321.1	Bkg Flaring	QPO, pulsations
80002092007	56692.2	319.3	Bkg Flaring	QPO, pulsations
80002092008	56698.8	35.2	Bkg Flaring	QPO, pulsations
80002092009	56699.5	119.6	Bkg Flaring	QPO, pulsations
80002092011	56719.7	114.5	Bkg Flaring	Flaring, QPO, pulsations
50002019002	57037.9	32.8	Bkg Flaring	
50002019004	57041.8	168.2	Bkg Flaring	QPO, Flaring
90101005002	57193.6	38.8	Bkg Flaring	Flaring
80202020002	57413.8	37.9		Flaring
80202020004	57441.7	32.4	Bkg Flaring	
80202020006	57483.4	31.8	Bkg Flaring	Flaring
80202020008	57502.8	41.7	Bkg Flaring	Flaring
30101045002	57493.3	195.5		QPOs, pulsations
80202020008	57502.8	41.7	Bkg Flaring	
30202022002	57542.9	40.3	Bkg Flaring	
30202022004	57570.7	48.2		
30202022008	57599.0	43.7	Bkg Flaring	Flaring
30202022010	57619.4	44.1	Bkg Flaring	
90201037002	57641.5	82.5		Pulsations
90202038002	57668.8	46.0	Bkg Flaring	QPO
90202038004	57722.6	45.0	Bkg Flaring	QPO
30502020002	58691.9	93.0		QPO
30502020004	58701.9	91.5		Flaring
30502021002	58918.1	86.3		QPO, pulsations
30502021004	58929.2	79.8	Bkg Flaring	QPO
30502022002	59000.8	90.9		QPO, Fast Flaring
30502022004	59012.6	99.2	Bkg Flaring	QPO, Fast flaring
30602028002	59215.2	69.0	Bkg Flaring	QPO
30602028004	59226.9	70.8	Bkg Flaring	
30602027002	59311.9	73.6		QPO, pulsations
30602027004	59325.3	71.6	Bkg Flaring	QPO, pulsations
30702012002	59504.0	128.3		QPO, pulsations
30702012004	59674.3	124.4		QPO
30901038002	60110.6	128.1	Bkg Flaring	Flaring
90901332002	60263.9	71.9		
90901333002	60275.3	53.7	Bkg Flaring	QPO
31001019002	60657.4	133.9	Bkg Flaring	QPO, pulsations

Table 3.1: Summary of observing conditions and source phenomenology for all observations in this work.

obsid	mjd	$\mathrm{rms_{rn}}$	$\nu_{\mathrm{max,rn}}$	$\mathrm{rms_{qpo}}$	$\nu_{0,\mathrm{qpo}}$	w_{qpo}	Q_{qpo}
80002092002	56680.5	0.130(18)	12(4)	0.095(19)	43(16)	6.7(15)	12(8)
80002092004	56682.8	0.122(11)	11.7(28)	0.103(11)	50.4(6)	9.8(17)	10.3(20)
80002092006	56685.5	0.103(4)	6.7(7)	0.127(4)	59.7(7)	26.3(22)	4.5(4)
80002092007	56692.3	0.086(6)	8.4(14)	0.152(5)	71.9(16)	56(5)	2.56(29)
80002092008	56698.8	0.172(19)	68(19)				
80002092009	56699.5	0.114(12)	30(8)	0.124(8)	98.3(12)	23(4)	8.4(15)
80002092011	56719.7	0.118(15)	49(17)	0.099(9)	65.6(5)	11.8(17)	11.1(17)
50002019002	57037.9			0.130(8)	31.9(16)	22(4)	2.8(7)
50002019004	57041.8	0.101(11)	12.3(28)	0.114(9)	36(14)	14.1(20)	5.2(27)
90101005002	57193.6	0.063(32)	0.8(5)				
80202020002	57413.9	0.079(18)	1.6(6)				
80202020004	57441.7	0.269(35)	230(60)				
80202020006	57483.4			0.184(9)	47.6(30)	56(9)	1.7(4)
30101045002	57493.3	0.122(4)	3.48(29)	0.161(6)	114.3(14)	42(4)	5.4(6)
80202020008	57502.8			0.153(12)	154.1(22)	28(6)	10.0(26)
30202022004	57570.7	0.069(30)	2.1(12)				
30202022006	57597.9	0.08(8)	1.4(7)				
30202022008	57599.0	0.055(24)	0.9(7)				
30202022010	57619.4	0.055(20)	1.3(6)				
90201037002	57641.5	0.265(35)	440(130)				
90202038002	57668.8			0.23(4)	30(100)	960(320)	0.07(23)
90202038004	57722.6			0.213(13)	222(10)	166(31)	2.7(6)
30502020002	58692.0			0.256(11)	275(16)	400(50)	1.39(27)
30502020004	58701.9	0.135(9)	0.98(20)				
30502021002	58918.2			0.194(8)	116(5)	116(15)	2.01(34)
30502021004	58929.2	0.162(21)	128(35)	0.114(19)	185.7(25)	36(10)	10.1(29)
30502022002	59000.8	0.120(8)	6.8(11)	0.117(8)	46.6(6)	10.9(19)	8.6(16)
30502022004	59012.6	0.096(7)	2.8(6)	0.184(9)	64(4)	69(11)	1.8(4)
30602028002	59215.2			0.192(8)	159(6)	136(18)	2.3(4)
30602028004	59226.9	0.038(14)	0.52(25)				
30602027002	59311.9	0.074(11)	5.1(23)	0.196(12)	165(8)	140(27)	2.4(6)
30602027004	59325.3	0.159(19)	67(17)	0.115(12)	91.3(8)	15.2(33)	11.0(27)
30702012002	59504.0	0.106(6)	4.1(6)	0.163(9)	135(4)	75(11)	3.6(6)
30702012004	59674.3	0.13(5)	130(50)	0.15(4)	143(9)	110(40)	2.7(12)
30901038002	60110.6	0.108(7)	0.92(18)				
90901332002	60263.9	0.117(10)	4.9(10)				
90901333002	60275.3	0.139(8)	6.2(9)	0.157(10)	131.0(19)	33(5)	7.7(13)
31001019002	60657.4	0.089(5)	3.0(4)	0.215(11)	171(10)	240(40)	1.46(31)

Table 3.2: Best-fit values and uncertainties for the red noise and QPO components obtained in Section *bootstrap*

3.5 Discussion

ObsID	Instrument	Date MJD	Exposure ks	Simultaneous NuSTAR ID
16580	ACIS-S	56691.8	47.5	80002092007
17578	ACIS-S	57038.6	10.1	50002019002
16023	ACIS-S	57042.0	10.1	50002019004
18064	ACIS-I	57483.7	25.1	80202020006
18068	ACIS-I	57502.8	25.1	80202020008
18070	ACIS-I	57669.0	25.1	90202038002
18072	ACIS-I	57723.4	25.6	90202038004
26664	ACIS-S	60274.5	40.1	90901333002

Table 3.3: Quasi-simultaneous *Chandra* and *NuSTAR* observations used for the QPO identification.

Our analysis is based on a multi-Lorentzian fit of the periodogram, as is often done in similar studies from accreting sources (Belloni & Hasinger, 1990). One difference is that we cannot rely on fractional variability when we study the evolution of the QPO, as we cannot have a clean view on M82 X-1 using *NuSTAR* due to the presence of M82 X-2, and the source flux is unknown. Hereafter, we will use fractional rms only to compare the strength of the QPO at different energy bands in a given observation, while we will use the absolute rms (in counts per second) when discussing the evolution over time.

This source confusion, on a related note, also hinders us from doing detailed modeling of the spectral break and its relation with the QPO frequency (à la Atapin et al. 2019), because the red noise component is produced by both sources.

The identification of the QPO with M82 X-1 is tricky by itself. Most detections come from missions that do not resolve the two ULXs (Strohmayer & Mushotzky, 2003), such as *XMM-Newton* and *RXTE*, although with a good degree of confidence given the higher flux of M82 X-1 and the change of the QPO strength when carefully selecting data closer to M82 X-1 in *XMM-Newton* (Feng & Kaaret, 2007). We looked for observations having quasi-simultaneous *Chandra* observations showing a low state from M82 X-2 and the QPO in the data, and searched the *Chandra* data themselves for QPO detections. Most on-axis *Chandra* observations of M82 are often plagued by pileup and the sensitivity to any variability was low. We found that the QPO was present in *NuSTAR* ObsID 90202038004, with *Chandra* ObsID 18072 showing a low state of M82 X-2, and, for the first time, a tentative detection in off-axis *Chandra* ObsIDs 17578 and 18064, this time using events firmly associated with M82 X-1, at a frequency compatible with the detection from the simultaneous *NuSTAR* ObsIDs (Figs. 3.8 and 3.9).

The evolution of the $20 - 300$ mHz QPO from M82 X-1 observed in our 10-year *NuSTAR* campaign shows behavior compatible with what was previously observed for this source (Mucciarelli et al., 2006; Feng & Kaaret, 2007; Atapin et al., 2019), extending the range of observed frequencies. The low background and hard response

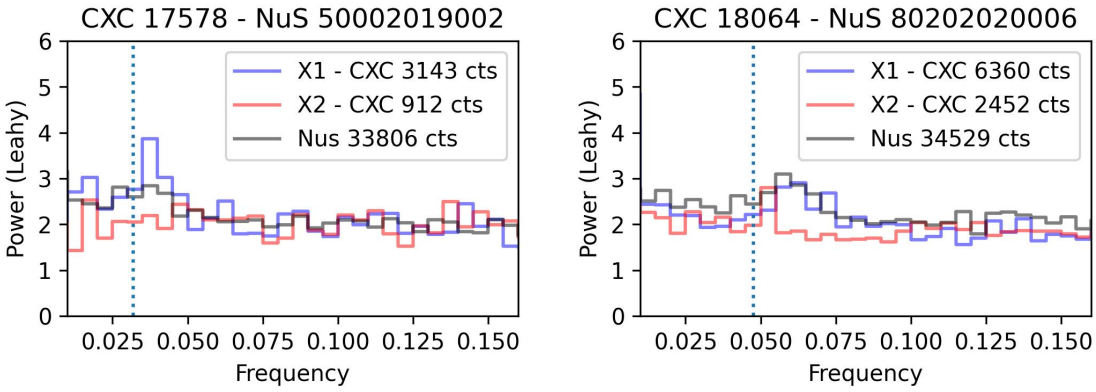


Figure 3.8: Simultaneous detections of the QPO in *NuSTAR* and *Chandra* data, showing a clear association with M82 X-1

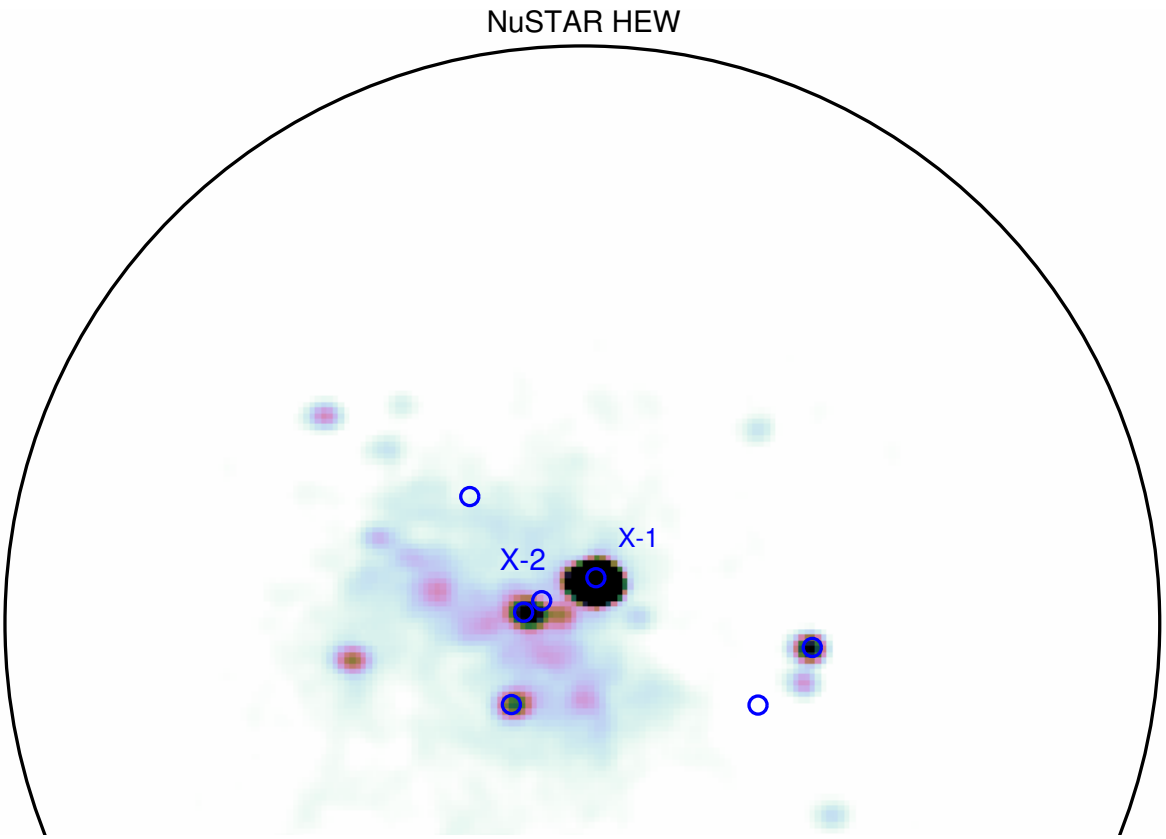


Figure 3.9: Image of *Chandra* ObsID 18072, simultaneous to *NuSTAR* ObsID 90202038004. Blue circles indicate the 7 ULXs in the catalogue by Liu & Bregman (2005). The QPO is significantly detected in *NuSTAR* data and not in *Chandra* data of any source, but the *Chandra* image shows that the emission is dominated by M82 X-1, while M82 X-2 is undetected.

of *NuSTAR*, together with our power spectral cleaning procedure, allows for a better modeling of the red noise component of the power spectrum and a systematic analysis of the evolution of the QPO using almost 3 Ms of exposure.

Low-frequency QPOs are observed in all classes of accreting sources (Wijnands & van der Klis, 1999). In particular, NS and BH binaries have a number of low-frequency features spanning the frequency range ∼0.001–50 Hz. Some evolution of the QPO frequency over time is a hallmark of almost all classes of QPOs (van der Klis, 2006; Motta et al., 2016), and in particular, all classes of low-frequency QPOs. Typically, the absolute rms is linearly correlated with flux, with the fractional rms being stable or slightly decreasing. Also the QPO frequency is often seen correlating with the flux on short time scales, while the correlation is broken on long time scales (the so-called "parallel tracks" van der Klis, 2006). At higher energies, the characteristic frequency does not change significantly, while their fractional rms generally increases.

One class of QPOs that can naturally be compared with ours is Type-C QPOs from BH LMXBs, the most common oscillatory pattern in this class of sources. They start to appear in the low-hard state and their frequency generally increases as the luminosity (and probably the mass accretion rate) increases, going through the intermediate states that lead to the high-soft state and sometimes the so-called ultraluminous state. These intermediate states are also associated with the presence of transient jet ejections. Type-C QPOs are generally also accompanied by broad red noise (or flat-top) components, usually also modeled with Lorentzians whose characteristic frequencies (Eq. 3.4) evolve in parallel with that of the QPO, and one or more harmonics. Models for these QPOs often involve Lense-Thirring precession around a rotating BH, and put their frequency in relation with other oscillatory components such as broadband noise or high-frequency QPOs. These models use different approaches from single-particle motion (the original relativistic precession model, or RPM; Stella & Vietri 1999), to precessing rings (e.g. Psaltis & Norman, 2000), to entire regions of the disk that precess like a solid body (Ingram & Done 2011, see also Ingram & Motta 2019 for a review).

Interpreting M82 X-1's QPO as a Type-C QPO, one can also be tempted to go one step further: all time scales around a gravitating body scale with mass, and notably this includes orbital frequencies. The fact that this QPO are about an order of magnitude slower than typical Type-C QPOs leads to a mass estimate of an order of magnitude above stellar-mass BHs, in the regime of small intermediate-mass BHs. This kind of scaling is often attempted, using different variability components, and in the case of M82 X-1 this has often led to claims of IMBH origin for this source (e.g. Strohmayer & Mushotzky, 2003; Pasham et al., 2014a).

However, similar claims were made for other sources. One such source is M82 X-2, for which Feng et al. (2010) estimated a mass of 12,000–43,000 $M_\odot$ by rescaling the *Chandra*-detected mHz QPOs to Type-C QPOs. However, this famously turned out to be a pulsar. It is generally difficult to compare QPO phenomena from different objects, and there is a wealth of QPO phenomena on many time scales in stellar mass compact objects, including at even lower frequencies. One example is the mHz QPO observed in low-mass NS and BH X-ray binaries (e.g. Revnivtsev et al., 2001; Xiao et al., 2024).

Moreover, we note that the spectrum of M82 X-1 does not depart significantly from the bulk of ULXs (Brightman et al., 2020), that are increasingly identified as super-Eddington accreting stellar-mass objects. In some cases the detection of pulsations

that univoquely identify them as NSs, but in general, their spectral shapes do not resemble the bulk of sub-Eddington accreting objects, and frequently show signatures of the strong winds expected from super-Eddington accretion (Pinto & Walton, 2023). In particular, NGC 5907 X-1 has a comparable flux to M82 X-1 despite being a NS (Israel et al., 2017; Fürst et al., 2023).

In addition, the decrease of the coherence of the M82 X-1 QPO with frequency observed here (Fig. 3.5) is puzzling, and unlike what is observed in Type-C QPOs where the coherence tends to generally increase with frequency, at least up to the onset of the soft state where these oscillations disappear. Admittedly, this might be an observational bias: the integration time required to detect the QPO might just be longer than the variation time for the frequency. We need thousands of oscillations of the QPO before being able to detect it, and if the frequency is changing rapidly this would artificially increase the measured width, and so its quality factor. A clear example of this are ObsID 80002092006 and 80002092007, where using the shift-and-add technique clearly improves the quality factor (Section 3.4.5). In an effort to determine if this played a role, we used the shift-and-add technique in all observations with a strong QPO and at least ∼90 ks exposure. The technique always improved the Q factor throughout the frequency range, but without significant advantages for the higher frequencies.

Interestingly, bright accreting pulsars also show low-frequency QPOs, with a phenomenology and characteristic frequencies similar to what we observe here. For these sources, low-frequency QPOs are often hypothesized to arise from the Keplerian frequency at some important radius, or from a beat between the Keplerian frequency at the truncation radius and the spin frequency (beat frequency model, Alpar & Shaham 1985). Manikantan et al. (2024) provides a table of the energy-dependent QPO parameters for a number of sources over multiple observations. V0332+53 is the only one with sufficient observations to see the evolution of the quality factor over a wide range of QPO frequencies, and Q is interestingly seen to decrease between 10 and 2 with increasing frequency, as we observe for M82 X-1. As a general rule, we argue against using QPOs alone to infer the mass of accreting objects. The QPOs studied in this work are compatible with phenomenology observed in accreting sources of different kinds, including NSs.

The non-detection of the twin 3–5 Hz QPOs reported by Pasham et al. (2014a) can be explained in various ways. First of all, the oscillations might be transient, and have disappeared over time: the result by Pasham integrated many years of data. It is also possible that the filtering of "flaring" observations by Pasham et al. (2014a), which reduced considerably the time intervals analyzed to obtain their average periodogram, was more aggressive than ours, and that we need more observations without flaring or high variability to detect those QPOs with *NuSTAR*. Another possibility might even be that the source of QPOs is *not* in M82 after all: the field of view of *RXTE* is one degree, which includes three ULX host galaxies (M82, M81, Holmberg IX) with at least 9 ULXs (Liu & Bregman, 2005).

3.6 Conclusions

We made an extensive, 10-year study of the QPOs from M82 X-1 using *NuSTAR*. Using archival *Chandra* data, we made a robust identification of the 50-300 Hz QPO with M82 X-1, which confirms previous evidence. Thanks to the sensitive response at and above 10 keV, we were able to detect the QPO over ∼70% of the 3 Ms of existing M82 observations, characterizing its behavior. We used a novel approach to cleaning the periodogram in order to better fit the low-frequency component and get reliable fit parameters for the QPO as well. The QPO tends to decrease its coherence as its frequency increases, but it is not clear whether this is due to fast variations of the frequency that we are not able to follow due to the long exposure required by the detection at these count rates. We note that LFQPOs in this range of frequencies and with similar behavior are observed in many accreting systems, including NSs, so that any inference on the mass of the compact object based on the frequency of the QPOs should be taken with a grain of salt.

3.7 Alternative periodogram analysis: Bartlett + Lomb-Scargle

Another widely used tool for detecting and characterizing periodic signals is the Lomb-Scargle periodogram (Lomb, 1976; Scargle, 1982). which has proven to be more effective in detecting periodic patterns even when observations are unevenly spaced, providing a reliable solution for analyzing time series data in astronomy and diverse scientific fields as well. The Lomb-Scargle allows for covering a wide range of low frequencies in comparison to the Bartlett periodogram, allowing for a more comprehensive analysis of low-frequency signals.

The only downside of this periodogram is that its powers are not guaranteed to be uncorrelated, and the assumption of χ^2_2-distributed powers is not as solid as for the periodogram. For a comprehensive view on modern algorithms to compute the Lomb-Scargle periodogram and their limitations, see VanderPlas (2018).

Nonetheless, the Lomb-Scargle periodogram is very useful for the analysis of unevenly sampled data, including light curves with missing data. Our original approach to the analysis of the datasets in this paper was a hybrid approach, using the Bartlett periodogram and the Lomb-Scargle periodogram at the same time. We write it here because it could be an interesting inspiration for future works and, in any case, it represents an alternative (even if not completely independent) approach to the analysis presented in this paper. We calculated the Lomb-Scargle periodogram simply discarding light curve bins outside GTIs, and avoiding oversampling (i.e., using the same spectral resolution of the FFT), which limited the correlations between powers. Our sample time was 0.1, giving 5 Hz as the Nyquist frequency, but we found that the periodogram departed significantly from the expected noise level of 2 when reaching about half the Nyquist frequency. Otherwise, the Lomb-Scargle and the Bartlett periodograms have very good overlap in the common frequency ranges, and we decided to eliminate the frequencies above 1 Hz from the Lomb-Scargle periodogram for extra caution. We checked that the assumption of χ^2_2 distributed powers is justified for our

case by plotting the distribution of the Lomb-Scargle powers of simulated and real data similarly to 3.3. Moreover, even though in theory it should not have been needed, we blacklisted the frequencies corresponding to the orbital occultations similarly to the description in section 3.3. The raw periodograms have very different frequency resolutions, but we rebinned both periodograms with geometrically increasing frequency bin sizes, and defined a threshold frequency $\nu_{\rm thr}$ where the frequency resolution of the rebinned Lomb-Scargle periodogram reached the frequency resolution of the Bartlett periodogram. From that point on, we used a hybrid periodogram containing the Lomb-Scargle powers below $\nu_{\rm thr}$ and the Bartlett ones above. The periodogram was characterized by the power values and the number of averaged powers, either from rebinning or – in the case of the Bartlett one – averaging of multiple periodograms.

We proceeded to fit the multi-Lorentzian model of Section 3.4.3, with the same methods, to this hybrid periodogram. Indeed, the results using this alternative method are compatible with those obtained for the single filtered periodogram in Section 3.4.3.

Chapter 4

Deep Learning Pipeline for Compact Source Detection in Planck Maps

4.1 Introduction and Context

Mapping the cosmic microwave background (CMB) requires separation of the CMB signal from foreground emissions and compact sources. Missions such as WMAP and Planck have conducted full-sky surveys across microwave to submillimeter wavelengths, providing extensive catalogs of embedded compact sources. WMAP's nine-year analysis identified numerous extragalactic sources in the microwave sky Bennett et al. (2013), and the Planck mission produced nine all-sky maps (30 to 857 GHz) containing not only the CMB signal but also a complex mixture of foreground emissions: diffuse Galactic radiation (especially thermal dust at 857 GHz), the cosmic infrared background (CIB) from unresolved distant galaxies, and various compact sources (Collaboration, 2014; Bonavera et al., 2021). The Planck Catalogues of Compact Sources (PCCS) document a vast array of compact objects that contribute to contamination in CMB maps Ade et al. (2014, 2016).

Planck released two PCCS versions: the first, derived from its nominal 15-month mission covering frequencies from 30 to 857 GHz, and a subsequent expanded version incorporating data from the entire mission Ade et al. (2016). These catalogs list tens of thousands of Galactic and extragalactic compact sources, including cold clumps, dusty star-forming galaxies, and blazars, which complicate the extraction of the CMB signal Ade et al. (2014, 2016). Detecting fainter sources, particularly at the highest frequency (857 GHz), presents a significant challenge due to strong emission from Galactic dust. These sources are often embedded in bright and highly non-uniform Galactic emission, which can drown out or mimic their signals, especially in crowded sky regions. Developing methods to reliably identify these faint sources and measure their flux (brightness) is crucial for catalog completeness and for foreground separation.

The Mexican Hat Wavelet 2 (MHW2) filter has been widely used for the HFI channels, including the construction of the first Planck Catalogue of Compact Sources (PCCS) (Ade et al., 2014). Wavelet filters enhance structures of the same size as the instrument's point-spread function while suppressing large-scale background variations. Applying the MHW2 filter converts a sky map into a wavelet space where compact objects stand out with higher signal-to-noise. However, a wavelet assumes a particular shape for both the source and background, so if the actual background is highly structured—such as filamentary cirrus clouds of dust—the filter might not perfectly

remove it, leading to false detections. Moreover, the MHW2 algorithm can introduce slight ringing artifacts around real sources due to its oscillating kernel (Akrami et al., 2020a), requiring careful handling to avoid misidentifications.

In the second Planck catalogue (PCCS2), the collaboration adopted a dual-pipeline strategy (Ade et al., 2016):

1. **MHW2** (updated implementation) for the six HFI frequencies (100–857 GHz);
2. **Bayesian** PowellSnakes[1]

Both pipelines generate single-frequency detection lists which are later merged and validated. While wavelet filtering efficiently boosts point sources, it still assumes a fixed analytic kernel; PwS, in contrast, models the local background and instrument noise jointly with the source profile, improving reliability at low signal-to-noise. Nevertheless, classical filters (whether MHW2 or PwS) operate per frequency map and cannot exploit spectral coherence across bands (Akrami et al., 2020b).

To improve reliability, the Planck collaboration introduced the Bayesian Extraction and Estimation Package (BeeP) (Akrami et al., 2020a). BeeP applies a multi-frequency Bayesian approach: it simultaneously analyzes Planck maps at 857 GHz, 545 GHz, and 353 GHz (the highest HFI channels) along with an external 3000 GHz IRIS map to identify sources that exhibit consistent spectral signatures. BeeP assumes that most 857 GHz sources emit like warm dust, producing detectable emission in adjacent bands. Instead of simply filtering out large-scale emission, BeeP explicitly models the level and properties of the surrounding diffuse emission, incorporating it into the detection process. This helps reduce confusion noise—the risk of mistaking a clump of dust for a compact source. The algorithm returns not only detections but also fitted parameters for each source, providing flux measurements at each frequency (after background subtraction) and fitting a modified blackbody curve to characterize the source's spectral energy distribution. It also assigns a reliability metric to each detection, indicating the probability that the source is genuine given the surrounding cirrus noise. While the Bayesian approach results in a more physically informed catalog, it remains computationally intensive and depends on an assumed spectral model (dust emission), which may not capture all types of sources.

In recent years, deep learning techniques have emerged as powerful alternatives for source detection in astronomy. Unlike fixed analytical filters, learning-based methods can automatically capture intricate signal patterns. Baron (2019) provides a review of the expanding role of machine learning in astrophysics, highlighting that deep neural networks can often outperform traditional algorithms by learning from large datasets.

[fix Citation] Convolutional Neural Networks (CNNs), in particular, have shown promise for detecting and classifying astrophysical sources due to their ability to recognize spatial features (Casas & et al., 2022). For example, Bonjean (2020)applied a U-Net CNN to Planck multi-frequency maps to detect galaxy clusters via the Sunyaev–Zel'dovich (SZ) effect.The CNN recovered known Planck clusters and identified thousands of new SZ cluster candidates, demonstrating the effectiveness of deep learning in extracting faint signals that classical methods might miss Bonjean (2020). Beyond clusters, CNNs have been used for classifying point sources in

[1]A multi-frequency, Bayesian-inference algorithm; see Carvalho et al. (2009) (PwS) for the three LFI channels (30–70 GHz) and for polarized-flux estimates.

surveys; attention mechanisms and advanced architectures have been integrated to boost performance.Bhavanam et al. (2024), for instance, enhanced a CNN-based classifier (MargNet) with attention layers and Vision Transformer components, yielding improved accuracy in distinguishing stars, quasars, and compact galaxies. This illustrates the benefit of allowing the network to focus on relevant features via attention, an idea that can be extended to point-source detection in confused fields. Meanwhile, unsupervised and generative models are being explored to assist CNN-based detection. Gagliano & Villar (2023) employed a variational autoencoder (VAE) for anomaly detection in galaxy data, showing that VAEs can learn the underlying distribution of astronomical images and flag unusual sources. Such generative models can be repurposed to learn generic sky backgrounds, helping to isolate point-source signals. Finally, semi-supervised learning techniques have begun to reduce the need for large labeled datasets in astronomy. Sohn et al. (2020) introduced FixMatch, a semi-supervised framework that combines consistency regularization and pseudo-labeling, which has been highly successful in general image classification. Astronomical applications are now adopting similar ideas – for example, Slijepcevic et al. (2022) applied semi-supervised and self-supervised learning to radio galaxy morphology classification, greatly reducing the required training labels. These advances suggest that a hybrid approach (combining unsupervised pre-training, semi-supervised learning, and supervised CNN detection) could significantly improve compact source detection in CMB maps. Our work builds on this prior art, bringing together a CNN with attention mechanisms, a VAE for feature learning, and a semi-supervised training scheme to push detection limits in the challenging 857 GHz Planck band.

The next step was MultiPoSeIDoN (Casas & et al., 2022), a multi-frequency CNN designed to leverage information across several Planck channels simultaneously. MultiPoSeIDoN uses a U-Net based architecture to perform image segmentation, essentially labeling pixels that contain a source across multiple map layers. Two versions were explored: one assuming sources have a flat spectrum across bands, and another allowing for a realistic spectral shape (so it looks at three frequency maps and learns the typical frequency-dependent signature of sources). When evaluated on realistic simulations at frequencies 143, 217, and 353 GHz, MultiPoSeIDoN showed impressive gains over traditional filtering. For instance, using a detection threshold corresponding to 90% completeness (meaning 90% of true sources above a certain flux are detected), the neural network was able to detect sources roughly as faint as 60–80 • mJy, whereas the wavelet-based filters required sources to be brighter (on the order of $\sim 100 mJy$ or more) to reach the same completeness(Casas & et al., 2022). In all cases, the CNN produced far fewer spurious detections (false positives) than the wavelet method at comparable detection levels. Furthermore, the fluxes of sources recovered by the neural network were closer to the true values – the network learned to output a more accurate flux estimate for each source, whereas the wavelet filter tended to underestimate fluxes of faint objects or struggled in the presence of strong background noise (Casas & et al., 2022). These results highlight that deep learning can adapt to complex background patterns and intrinsically combine multi-scale information, giving it an edge in both finding and characterizing dim sources in Planck maps.

Building on this progress, our work focuses on developing a deep learning framework for detecting compact sources specifically in the Planck HFI 857 GHz maps. This is

a particularly challenging case due to the bright Galactic emission at 857 GHz, but it is also scientifically rich, as this band contains many dusty galaxies and cold clumps. Our approach follows the general strategy of training a convolutional neural network on simulated sky "patches" so that it learns to identify point sources and estimate their fluxes. We describe below the key components of our method: the preparation of training data, the network architecture (which includes attention mechanisms and a Transformer module), the training procedure with a hybrid loss function, and the evaluation methodology.

4.2 Methodology

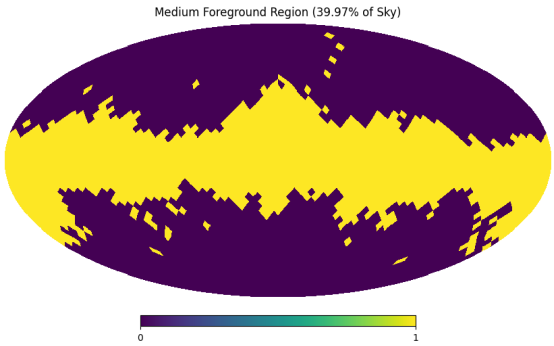
Our experiments are based on submillimeter astronomical sky images containing both faint compact sources and background noise. The dataset consists of large sky maps from a submillimeter survey, which we subdivide into smaller fixed-size image patches for analysis. Each patch is labeled as either "source" (containing at least one compact source) or "no-source" (containing no detectable source). The labeled set was constructed using an existing source catalog from the survey : patches centered on cataloged source positions are marked as positive examples, while patches extracted from blank regions (with no known sources) serve as negative examples. To avoid ambiguity, positive patches are chosen such that the source lies within the central region of the patch (minimizing cases where a source is near the edge), and negative patches are taken from regions well away from any known source. This results in a curated collection of labeled patches reflecting the two classes of interest.

Because genuine compact sources are relatively rare in wide-field sky images, the class distribution in the raw data is highly imbalanced – there are many more empty sky patches than source-containing patches. We address this by stratified sampling when splitting the data into training, validation, and test sets. In particular, we ensure that each subset (train/validation/test) contains a representative proportion of source vs. no-source patches, so that performance can be fairly evaluated on both classes. The final training set comprises 52166 of labeled patches, 50% of them are patches with sources listed with high reliability in the BeeP catalog, and the other half of the set is randomly sampled at a distance $\geq 1^\circ$ away from any source of the PCCS2 and its extension PCCS2E catalogs. The test set is kept aside with a balanced mix of positives and negatives for unbiased evaluation.

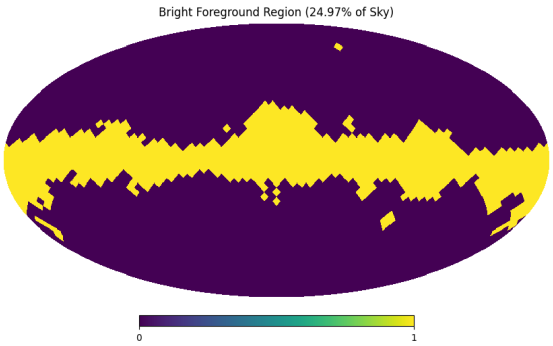
In addition to the labeled data, we leverage a large pool of unlabeled image patches drawn from the same survey. These unlabeled patches are taken from regions of the sky images that have not been exhaustively labeled (they may contain undetected sources or just noise ; we treat them as having no ground-truth labels). The unlabeled dataset greatly outnumbers the labeled set, which is typical in astronomy – only a small fraction of objects are confirmed or annotated, while vast areas of the sky remain unlabeled. By using this unlabeled pool in a semi-supervised learning framework, we aim to improve the model's sensitivity to faint sources without requiring additional manual annotations.

Each patch in our dataset is a grayscale image (intensity map) at submillimeter wavelength with dimensions of (64×64) pixels (with a field of viewo 1.8° covering

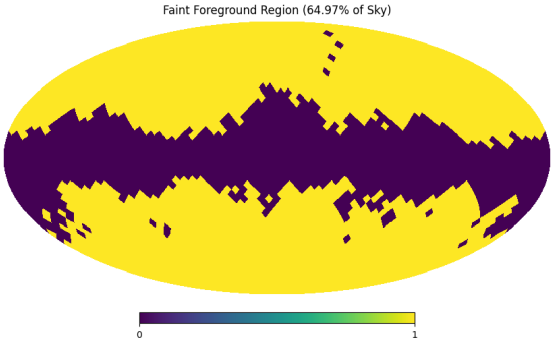
a region of sky sufficiently large to contain a compact source and surrounding background). The telescope's point spread function causes compact sources to appear as roughly point-like blobs a few pixels wide. Pixel values represent flux density and typically follow a noisy background distribution with occasional peaks for sources. Given the nature of the instrumentation and sky background, noise in these images is approximately Gaussian with spatially varying standard deviation (due to instrumental and atmospheric effects). This context motivates some of our pre-processing steps, as described next.



(a) 25% : Faint foreground emission



(b) 40% : Medium foreground emission



(c) 65% : bright foreground emission

Figure 4.1: Mask regions by defined by three foreground emission levels

4.3 Pre-processing

We apply a series of pre-processing steps to normalize and augment the data before it is fed into the neural network. These steps ensure that the input patches are on a consistent scale and improve the model's ability to generalize. The key pre-processing operations are described in the following paragraphs.

Patch extraction and alignment. Patches are extracted using a Gnomonic Projection[2] Patches These patches are centered such that any potential, if present, is roughly in the

[2]A gnomonic projection is used to project a spherical surface onto a plane; the center of projection is the sphere's center, and the surface can be projected onto any plane not passing for the center, typically tangential to the surface.

middle of the patch. Centering the sources in positive patches is not strictly required for classification, but it provides a consistent reference frame for the model and avoids issues where a source might be cut off at a patch boundary.

Background subtraction and normalization. Prior to training, each patch undergoes Gaussian blurring (kernel 3×3, $\sigma = 1$) to suppress high-frequency noise. The blurred patch is then standardized by subtracting its mean pixel value and dividing by its standard deviation, enhancing robustness to background variations. Labeled patches are divided into training (70 %), validation (15 %), and test (15 %) sets via *stratified sampling* (**?**), ensuring that positive and negative classes are proportionally represented in every subset.

Region Masks. (Figure 4.1)To cope with different foreground levels across the Planck sky, we slice the full-sky map into three partly overlapping brightness zones. Roughly 25% of the sky falls in the "faint" zone, about 40% in the "medium, and 65% in the "bright". The boundaries are drawn with a mask built from the Planck 857 Ghz foreground components—dust, free-free, and synchrotron—and depend only on the local integrated intensity of those components. The faint region therefore picks out the high-latitude, low-emission patches, while the bright region hugs the Galactic plane where dust light blazes most strongly. Splitting the data this way matters because background brightness strongly influences how easily a network can spot compact sources. Training a separate model on each brightness tier lets the network focus on the noise characteristics that are unique to that tier, instead of trying to learn a single compromise that works everywhere. In practice, this regional training strategy yields more dependable detections and keeps the false-positive rate in check even in the glare of the Galactic plane.

Pre-training with Variational Auto-Encoder (VAE). A Variational Auto-Encoder (VAE) is employed for pre-training to learn a compressed representation of the sky patches. The VAE architecture consists of: **Encoder** : Two convolutional blocks (each with two 3x3 convolutions, batch normalization, and ReLU activation) followed by max-pooling, reducing the spatial dimensions. The bottleneck outputs the mean and log-variance of a 64-dimensional latent space. **Decoder** : Two transposed convolutional blocks to reconstruct the input patch, mirroring the encoder structure. The VAE is designed to capture the underlying distribution of the patches, providing a robust initialization for subsequent steps.

Data augmentation. To increase the effective size of the training dataset and make the model more robust, we employ data augmentation on the fly during training. Simple geometric transformations that preserve the semantics of the image are applied to patches. For labeled training patches, we randomly flip images horizontally and vertically, and also rotate by 90, 180, or 270 degrees. These transformations are appropriate because astronomical sources at these wavelengths have no inherent orientation, so a source rotated or flipped is still a valid example of the same class. We also apply small random translations (shifts by a few pixels) to simulate slight mis-centering, and

occasional random Gaussian noise added to mimic instrumental noise variation. These augmentations help the model not to overfit to the exact positions or noise realization of sources in the training set. For the unlabeled patches, augmentation plays a central role in the semi-supervised strategy : we will use a distinction between "weak" and "strong" augmentations as part of the FixMatch algorithm. In brief, weak augmentations are a subset of the above operations (e.g. flips and minor shifts), whereas strong augmentations include more aggressive transformations (such as larger rotations, intensity jittering, or Cutout mask augmentation). All augmented versions of patches are generated on-the-fly each epoch, ensuring the model sees a new variation of each patch each time, which improves generalization.

After these pre-processing steps, the labeled and unlabeled patches are ready for input into the model. The normalization ensures that the network's weights can learn on a consistent intensity scale, and the augmentations expose the model to a wide range of plausible variations, which is particularly important given the limited number of labeled examples. Next, we describe the architecture of the deep learning model used for classification of these patches.

4.4 Model Architecture

Our model 4.2 follows a U-Net style convolutional encoder-decoder architecture with the addition of attention gates to enhance feature focus. The U-Net architecture, originally introduced by Ronneberger et al. (2015) for biomedical image segmentation, is well-suited to our task because it combines multi-scale feature extraction with precise localization through skip connections. In essence, the network has an encoder (downsampling path) that learns a hierarchy of features from the input patch, and a decoder (upsampling path) that reconstructs a spatial output using those features, with shortcuts connecting encoder and decoder layers of the same scale. This design allows the model to capture the context of an entire patch (important for distinguishing noise fluctuations from real sources) while preserving fine-grained spatial information (important for pinpointing the presence of a small source in the patch).

4.4.1 Training Strategy

Training the model for reliable source detection is challenging given the limited availability od ground truth data. More that 40 % of entries in the PCCS2 catalog are suspected to be spurious. To cope with this issue, we employ a semi-supervised learning strategy to allow the model to learn from abundant unlabeled patches alongside the labeled ones. The less reliable sources will be treated as unlabeled data, enabling the model to be fine-tuned for improved performance and adaptability.

Supervised training

We perform standard supervised learning on the labeled dataset. The model's predictions for the patches are compared to the known ground-truth labels (1 for source, 0 for no-source) using a binary cross-entropy (BCE) loss. Binary cross-entropy is appropriate for a single-probability Ramos et al. (2018). Labeled patches are augmented with the

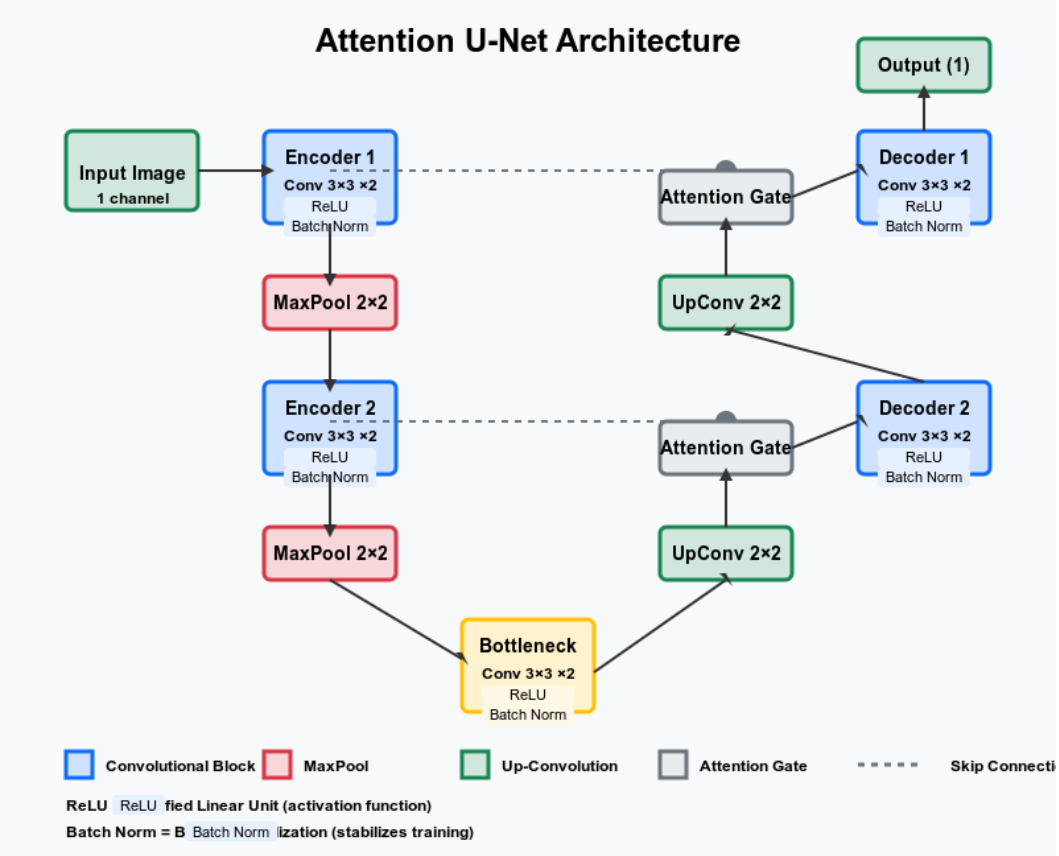


Figure 4.2: Attention U-Net architecture highlighting the flow through encoders, bottleneck, attention gates, and decoders for enhanced image processing tasks.

"weak" transformations (random flips/rotations as described in pre-processing) during training to add variation, but their true labels remain the supervisory signal.

Semi-Supervised Training (FixMatch)

We leverage a large pool of unlabeled data with FixMatch strategy Sohn et al. (2020). which combines pseudo-labeling and consistency regularization. The idea is to generate a pseudo-label for an unlabeled example using the model's own prediction (when confident), and then enforce that the model's prediction remains consistent under a more distorted view of that same example. For each unlabeled patch in a training batch, we proceed as follows :

- **Weak Augmentation and Prediction** : We first, apply a weak augmentation to the unlabeled patch using light transformations (e.g., a small random rotation or flip). Pass this augmented patch through the model to obtain a predicted probability for the "source present" class.

- **Pseudo-Label Generation:** If the model's predicted probability is very high (above 0.95) for a source, we assign a pseudo-label of 1 (source present) to the patch. If the predicted probability is very low (below 0.05), we assign a pseudo-label of 0 (no source). These thresholds (0.95/0.05) reflect a high confidence requirement. Predictions that fall in between (i.e., uncertain cases) are not given any pseudo-label and the patch is skipped for the unsupervised loss.

- **Strong Augmentation & Consistency**: Next, we apply a strong augmentation to the same unlabeled patch (for example, heavy random rotations, crops, or color distortions) and pass this augmented patch through the network, yielding a new prediction p_{strong}. We then enforce consistency between the network's output on this strongly augmented input and the pseudo-label $\hat{y}$ obtained from the weakly augmented input.

- **Consistency Prediction**: Feed the strongly augmented patch into the model to obtain a new prediction (a probability of source present).

- **Unsupervised Consistency Loss**: If a pseudo-label was assigned in step 2, compute a loss that penalizes any discrepancy between the model's prediction from step 4 and the pseudo-label. We use binary cross-entropy for this loss, treating the pseudo-label as the ground truth for the strongly augmented input.

Through this training strategy, the model learns from the ground-truth labels on the bright, clear examples and simultaneously from the vast number of unlabeled examples by enforcing consistency on those. The end result is a model that can detect compact sources with higher sensitivity than would be possible using purely supervised training on the small and incomplete labeled dataset. We next discuss how we evaluate the trained model's performance in a quantitative manner.

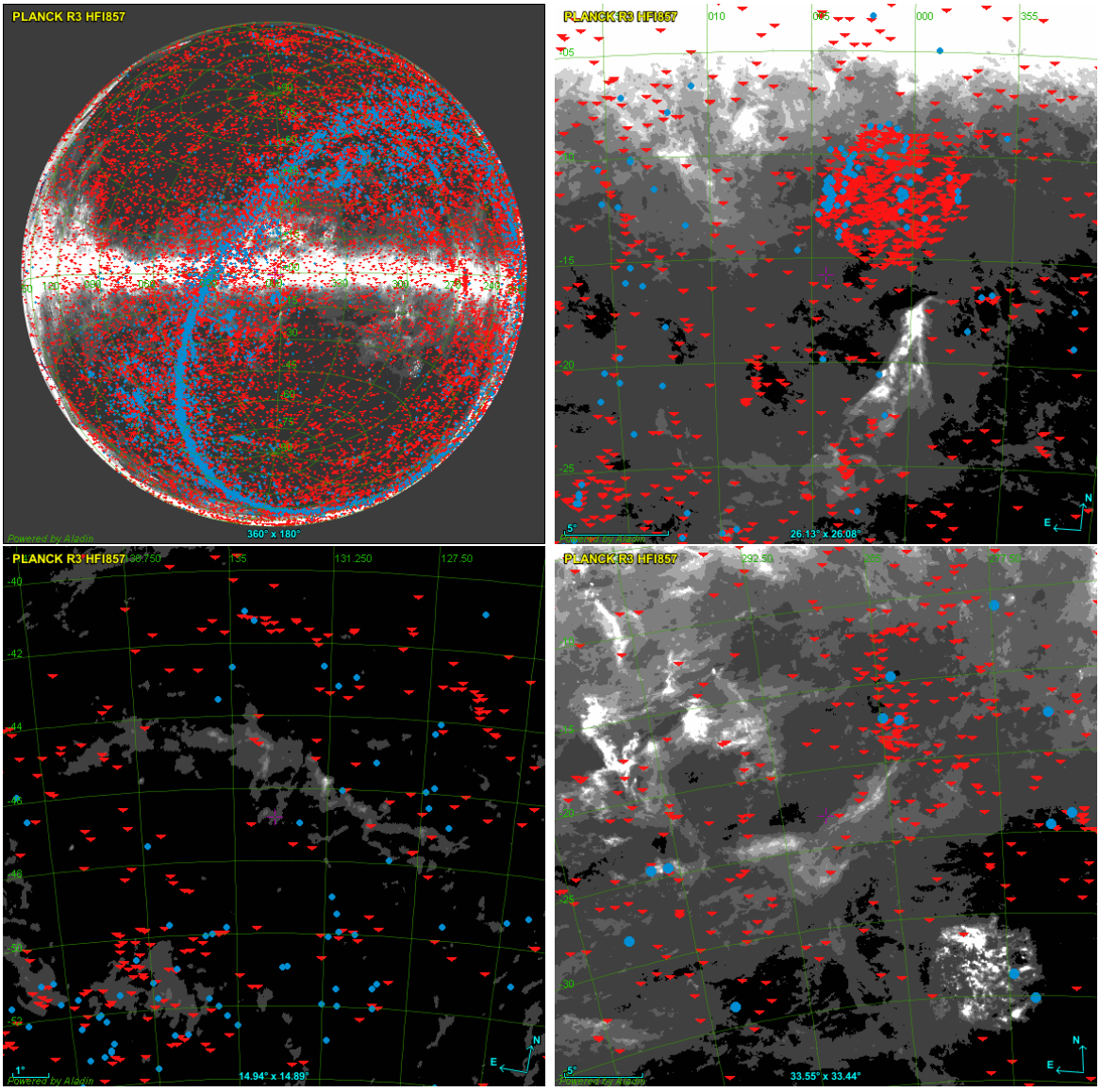


Figure 4.3: Four representative Planck HFI 857 GHz sky views with catalogue overlays. The sphere view represents the full sky with all PCCS2/E catalog sources . Red triangles. mark PCCS2/E sources that our pipeline detects, while blue circles mark sources the pipeline does not detect. Even in the most crowded and bright-foreground regions (upper-right panel), red symbols dominate, showing that the network still recovers the majority of catalogue sources despite severe source confusion and diffuse emission. In the cleaner fields (lower panels), virtually all catalogue sources are detected.

Regional Stratification

One challenge in our project is that the images come from multiple distinct sky regions, each with its a different characteristics (for example, different background noise levels, source densities, or artifact rates). A single model trained on all regions might struggle to accommodate these differences. To address this, we employ a regional stratification strategy: we divide the data into three regional subsets and train a separate model for each region. In practice, the images were sorted according to their region of origin.

We devided the labeled and unlabeled data into into three groups: Region A, Region B, and Region C. This grouping ensured that each subset concontained images with similar astronomical conditions.

We replicated our entire training pipeline for each of these region-specific subsets, training three models (one per region) with identical architecture and hyperparameters. All models were initialized with the same encoder pretrained as a variational autoencoder (VAE) on the combined unlabeled dataset from all regions, providing a common baseline of general features. After this shared initialization, each model was trained independently on its own region's data with no further weight sharing. This approach offered a balance between general feature learning and region-specific specialization: the shared initialization captured broad features, while the independent training allowed each model to focus on the nuances of its region.

As a result, each model specialized in the particular characteristics of its region. For instance, the model trained on Region A learned to recognize Region A's typical noise patterns and source appearances without being confused by the much higher density of sources in Region B. Similarly, the Region B model could focus on its own conditions without interference from Region A's attributes. By isolating the training by region, we hypothesized that these specialist models would outperform a single generalized model, which would have to compromise when dealing with very different data distributions.

Preliminary experiments supported this approach. Each region-specific model achieved a higher F1 score on its own region's validation set than the single model trained on all regions combined achieved on that same region's data. These results validated our hypothesis that stratifying the training by region can capture region-specific details more effectively.

Ensemble Strategy

After training the three regional models, the next step is to combine their outputs to obtain the final detection results for the entire survey. We achieve this using an ensemble strategy that leverages all three models' predictions for each image. In particular, we adopt a strict 'AND' ensemble approach : a patch is classified as a source only if all the corresponding regional models of the regions it belongs to predict it as a source.

For each test image, we run all three regional models and then perform a pixel-wise logical AND on their predicted binary masks. In other words, a pixel is labeled as a source in the final output only if all three models agree that it is a source. This conservative fusion sharply reduces false positives, as spurious detections from individual models are unlikely to align, at least in the overlapped areas. Computationally, the ensemble is efficient: predictions are parallelized, and combining outputs requires only simple logical operations.

4.5 Results and Conclusions

The proposed pipeline achieves excellent detection performance across the sky. By enforcing consensus among specialized models, we obtained a highly reliable catalogue of sources. The best validation F1 scores (0.96, 0.95, and 0.93 for faint, medium, and bright regions, respectively) and the ensemble's test F1 score of 0.96 testify to

Table 4.1: Detection efficiency of the strict–AND ensemble for the two reliability classes and for the full set of PCCS2 & PCCS2E catalogs entries. "Detected'' counts correspond to sources flagged by the pipeline; efficiency is the fraction of detected to total sources in each class.

Reliability class (BEEP)	**Total sources**	**Detected**	**(%)**
High	26 083	24 730	94.8
Low	22 098	14 013	63.4
Total	**48 181**	**38 743**	**80.4**

the effective of identifying most real sources (high recall), while supressing spurious detections to a low level (high precision.). Table 4.1 summarizes the detection efficiency as a function of the Planck reliability class. Quantitatively, the detection rate of bright sources is around 95% in all cases. Spurious detections in each class are controlled at or below the prescribed reliability levels by construction.

Taken together, these results indicate that our semi-supervised, regionally-specialized pipeline is both sensitive and robust. Even in the most challenging parts of the sky (near the Galactic plane), The strict-AND ensemble effectively eliminates unreliable candidates, yielding an output catalog that is self-consistent and stable. Overall, our method recovers nearly all relevant PCCS2/E objects while introducing only a handful of new candidates.

Conclusion

This work introduces an end-to-end deep-learning pipeline for blind detection of compact sources in CMB temperature maps. The sequence begins with an unsupervised VAE that learns generic sky features without labels, moves to FixMatch for leveraging the abundant unlabeled patches, trains separate CNNs tailored to low-, mid-, and high-foreground zones, and finishes with a strict-consensus ensemble that filters out spurious sources. Benchmarking on real Planck PCCS2/PCCS2E data shows that the system retains more than 95% of catalogued sources even under the most stringent reliability cuts (high reliable detected sources), while flagging a small set of additional high-confidence candidates worthy of follow-up (low reliable detected sources). Taken together, the results demonstrate that modern deep-learning techniques provide a robust and scalable solution for extragalactic point-source detection in current and forthcoming cosmological surveys.

Chapter 5

High-Resolution Spectral Analysis of Unevenly Spaced Data Using a Regularization Approach

5.1 Introduction

Frequency-domain analysis is a core method for time series, showing how signal power is distributed over different frequencies. Astronomers routinely use the power spectral density (PDS) of light curves to characterize their variability in light curves (e.g. identifying periodic or quasi-periodic signals superposed on stochastic noise). Classical spectral estimation methods, such as the Fourier transform and periodogram, assume data are evenly sampled in time. However, *unevenly spaced data*—common in astronomy due to interruptions in observations (satellite orbits, Earth occultations, etc.)—violate this assumption and pose significant challenges. Irregular sampling complicates the efficient computation of spectra (the fast Fourier transform cannot be directly applied) and, more importantly, makes the statistical interpretation of periodograms difficult (Vaughan et al., 2011).

In other words, the standard periodogram's distribution under noise is not well-defined for uneven sampling (Gúrpide & Middleton, 2025), which hampers reliable detection of true signals. Astronomical X-ray light curves exemplify these issues: they often consist of gapped or uneven observations and exhibit strong red noise (stochastic variability), so spurious peaks can arise if standard methods are applied incautiously.

To handle irregular timing, the *Lomb-Scargle periodogram* is commonly employed as an extension of the classical periodogram (Lomb, 1976; Scargle, 1982). It provides a way to compute a periodogram without needing uniform time bins and mitigates some statistical issues by normalizing the signal and accounting for observational gaps. While Lomb-Scargle periodograms (LSP) have become a standard tool, they still have limitations. In particular, the LSP's assumptions (a single-frequency sinusoid plus constant offset, with white noise) can be unrealistic for astrophysical time series that contain broadband noise and multiple signals or time-varying amplitudes. As a result, classical periodograms and LSP can miss subtle features or introduce biases, making it hard to distinguish a weak periodic signal from random fluctuations in a noisy, unevenly sampled X-ray light curve.

These limitations motivate the development of more robust, high-resolution spectral analysis techniques that can handle irregular data and complex signal content. One promising solution is to cast spectral estimation as an *inverse problem* and apply regularization. This approach can naturally handle uneven sampling and incorporate

prior information to stabilize the solution, yielding a cleaner estimate of the PDS than classical methods.

5.2 Challenges of Uneven Sampling and Classical Methods

5.2.1 Uneven sampling limitations

Unevenly sampled data violate the assumptions of the standard discrete Fourier transform (DFT), leading to:

- **Loss of FFT efficiency:** Because the Fast Fourier Transform (FFT) assumes uniformly spaced time samples, irregular time steps necessitate the use of direct summation methods rather than an FFT.
- **Spectral leakage and windowing:** Missing data or uneven intervals effectively apply a complex window to the signal, causing power to leak between frequencies, and smearing out sharp features.
- **Undefined noise distribution:** The classical periodogram's known statistical properties for even sampling do not directly carry over to irregular sampling, complicating confidence interval construction and peak significance.

In X-ray astronomy, these issues are pronounced. NuSTAR or other X-ray missions produce data with gaps, leading to strong red-noise phenomena. The periodogram of such a light curve is typically dominated by a power-law continuum with superimposed features like QPOs. Standard approaches (periodogram, LSP) may result in:

- Excess scatter in the estimated PDS, masking or mimicking genuine signals;
- Potential false positives due to random noise fluctuations under uncertain statistics;
- Reduced frequency resolution, unless artificially zero-padded or interpolated (which can introduce artifacts).

5.2.2 Lomb–Scargle Periodogram and Its Assumptions

The Lomb–Scargle method (Lomb, 1976; Scargle, 1982) performs a least-squares fit of a single sinusoid (with an offset) at each trial frequency. Unlike the classical Fourier approach, Lomb–Scargle handles irregularly spaced data directly, without needing to interpolate. Mathematically, for a time series $\{x(t_i)\}_{i=1}^{N}$ observed at irregular times $\{t_i\}$, Lomb–Scargle aims to solve

$$x(t_i) \approx A\cos(\omega t_i) + B\sin(\omega t_i) + C, \tag{5.1}$$

where ω is the angular frequency. The parameters (A, B, C) are estimated by minimizing the sum of squared residuals:

$$\min_{A,B,C} \sum_{i=1}^{N} \Big[x(t_i) - A\cos(\omega t_i) - B\sin(\omega t_i) - C \Big]^2. \tag{5.2}$$

Denoting the data vector by $\mathbf{x} \in \mathbb{R}^N$ and the corresponding regressor matrix by

$$\boldsymbol{\phi}(\omega) \;=\; \begin{bmatrix} \cos(\omega t_1) & \sin(\omega t_1) & 1 \\ \cos(\omega t_2) & \sin(\omega t_2) & 1 \\ \vdots & \vdots & \vdots \\ \cos(\omega t_N) & \sin(\omega t_N) & 1 \end{bmatrix}, \tag{5.3}$$

the ordinary least-squares solution for each fixed ω can be written as

$$\widehat{\boldsymbol{\beta}}(\omega) \;=\; \Big[\boldsymbol{\phi}(\omega)^\top \boldsymbol{\phi}(\omega) \Big]^{-1} \boldsymbol{\phi}(\omega)^\top \mathbf{x}, \tag{5.4}$$

with $\boldsymbol{\beta}(\omega) = [A, B, C]^\top$. A corresponding Lomb–Scargle "power" at each frequency is then computed as

$$P(\omega) \;=\; \tfrac{1}{2}\Big(\widehat{A}^2 + \widehat{B}^2 \Big), \tag{5.5}$$

which quantifies how well a single sinusoid of frequency ω explains the data.

5.2.3 Key Assumptions and Their Consequences

1. **Single-frequency fit:** Each frequency is treated in isolation, so correlations among frequencies are not modeled jointly. If multiple closely spaced signals exist, they can "leak" into each other's frequency bins.

2. **Uncorrelated (white) noise:** The well-defined statistical distribution for Lomb–Scargle peaks relies on the assumption of *i.i.d.* (independent and identically distributed) Gaussian noise. Strong red noise or correlated structure in the residuals can spuriously inflate the power at certain frequencies or broaden signal peaks.

3. **Short or gappy time series:** While Lomb–Scargle accommodates non-uniform sampling, if the data set is too short or irregular, the variance of the estimated parameters can grow large, and the periodogram can become quite noisy.

4. **Single-sinusoid model:** Only one sinusoid plus an offset is fit at each frequency. Complex signals containing multiple harmonics or broad continuum structure may not be well represented by such a restrictive model.

For cases in which the signal comprises multiple subcomponents, intricate broadband structure, or correlated noise, the Lomb–Scargle assumption of a single sinusoid per frequency becomes restrictive. Hence, one naturally seeks global spectral estimation methods—such as those using regularization or Bayesian priors—to simultaneously model structure *across* frequency. By fitting multiple frequencies and exploiting prior

information about the smoothness or sparsity of the spectrum, such methods can stabilize power estimates and surpass the conventional "Rayleigh limit" in scenarios with short or irregular sampling. These topics are discussed in detail in the following sections.

5.3 Regularization Approach to Spectral Estimation

5.3.1 Formulating Spectral Analysis as an Inverse Problem

Let N be the number of data points, $\mathbf{y} \in \mathbb{R}^N$ (or $\mathbb{C}^N$) the observed time-series values at times $t_1, \ldots, t_N$. Suppose we want to estimate the underlying spectrum $\mathbf{x} \in \mathbb{R}^P$ (or $\mathbb{C}^P$) at a chosen set of P frequency bins $\{f_1, \ldots, f_P\}$. We write:

$$\mathbf{y} = \mathbf{A}\,\mathbf{x}, \tag{5.6}$$

where $\mathbf{A}$ is an $N \times P$ matrix encoding the relationship between the time-domain and frequency-domain signals. Typically,

$$A_{n,k} = \exp\bigl(2\pi i\, f_k\, t_n\bigr),$$

for complex exponentials (or sines/cosines for a real-valued model). For evenly spaced data and $P = N$, $\mathbf{A}$ may be invertible. But for *uneven sampling or large* P, $\mathbf{A}$ becomes ill-conditioned or underdetermined: infinitely many solutions $\mathbf{x}$ can fit $\mathbf{y}$ perfectly if we do not constrain the problem further.

5.3.2 Ill-conditioning and the need for regularization

The matrix $\mathbf{A}$ can be highly correlated, especially if $P \gg N$ for super-resolution. Attempting to solve $\mathbf{y} = \mathbf{A}\mathbf{x}$ via ordinary least squares can lead to fitting noise or to wildly oscillatory solutions. Regularization is introduced to impose stability. In a Tikhonov-like framework, we define a penalized cost function:

$$J(\mathbf{x}) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda\, R(\mathbf{x}), \tag{5.7}$$

where $R(\mathbf{x})$ is the *regularization function* (or prior), and $\lambda > 0$ governs the balance between fidelity to data (the $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2$ term) and adherence to prior assumptions (via $R(\mathbf{x})$). A smaller λ attempts to fit the data more closely (risking noise overfitting), while a larger λ enforces smoother or sparser solutions (potentially losing fine detail).

5.3.3 Choice of Regularization Function $R(\mathbf{x})$

The choice of $R(\mathbf{x})$ determines the nature of the prior.

- A simple *ridge* (L2) penalty, $R(\mathbf{x}) = \|\mathbf{x}\|_2^2$, yields smooth solutions but can overly blur sharp spectral lines.
- An *L1* penalty, $R(\mathbf{x}) = \|\mathbf{x}\|_1$, promotes sparsity but is non-differentiable at $x_k = 0$ and can produce spiky solutions.

Ciuciu & Others (2001) introduced a convex function

$$R(x_k) = \sqrt{s^2 + x_k^2},$$

summed over k, which is differentiable everywhere and behaves like $|x_k|$ for large $|x_k|$. This "hyperbolic" or "Cauchy-type" penalty encourages parsimony without harshly driving small coefficients to zero. As shown by Seghouani (2017), it can be effective for short, noisy, and unevenly sampled data.

5.3.4 Solving the Regularized Inverse Problem

Minimizing

$$J(\mathbf{x}) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda \sum_k \sqrt{s^2 + x_k^2},$$

is a convex problem, but there is no closed-form analytic solution. Common numerical methods include:

1. **Gradient-based algorithms**: The gradient of $J(\mathbf{x})$ is well-defined because $\sqrt{s^2 + x_k^2}$ is differentiable. One can use gradient descent or quasi-Newton (e.g., L-BFGS) to iteratively update $\mathbf{x}$.

2. **Iteratively Reweighted Least Squares (IRLS)**: Rewrite $\sqrt{s^2 + x_k^2}$ in a quadratic approximation around the current iterate, solving a weighted least-squares subproblem at each iteration. This approach often converges quickly for these half-quadratic penalties.

3. **Half-Quadratic (HQ) methods**: Introduce auxiliary variables to transform the non-quadratic penalty into a series of simpler subproblems.

Because $J(\mathbf{x})$ is strictly convex, these methods converge to a unique global minimum. In practice, computational cost depends on N and P. For large problems, fast matrix-vector multiplication and iterative solvers (like Conjugate Gradient) may be used.

After solving for $\hat{\mathbf{x}}$, we can form the *power spectral density* as

$$P(f_k) = |\hat{x}_k|^2.$$

This yields a high-resolution spectrum that is less noisy and better at separating closely spaced frequencies or lines than the Lomb-Scargle or raw periodogram, particularly when data are unevenly spaced or short.

5.4 Application to NuSTAR X-ray Data of M82

To test the regularisation scheme under clean, fully controlled conditions, we built a synthetic light curve that combines three sinusoidal components—a lone feature at $\simeq 0.2\,\mathrm{Hz}$ plus a close pair at $\simeq 0.8$ and $0.9\,\mathrm{Hz}$—overlaid with Poisson noise. We then imposed on this mock data set the same visibility gaps that affect the *NuSTAR* observations. As Figure 5.1 shows, the Seghouani-regularised spectrum of the gapped signal

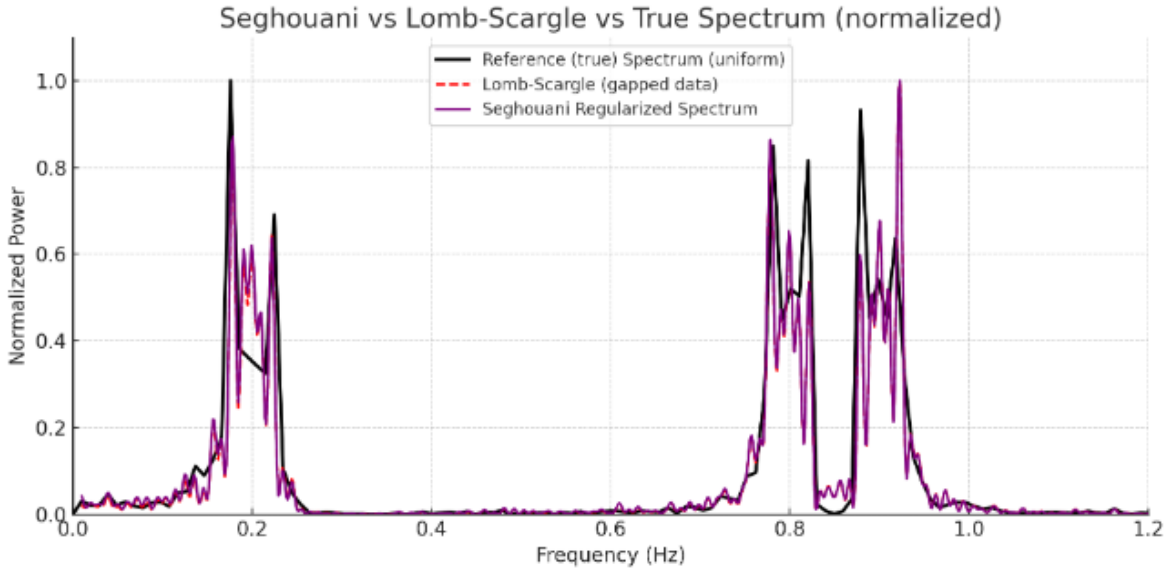


Figure 5.1: Normalised power spectra of the simulated Poisson-noisy light curve containing three sinusoids: one at $\simeq 0.2\,\mathrm{Hz}$ and a closely spaced pair at $\simeq 0.8$–$0.9\,\mathrm{Hz}$. **Black curve**: FFT of the uniformly sampled light curve, marking the true broad peaks. **Red dashed curve**: Lomb–Scargle periodogram of the same signal after two large gaps; irregular sampling fragments each broad feature into a cluster of narrow side-lobes. **Magenta curve**: Seghouani-regularised spectrum of the gapped data, which closely follows the underlying peak envelopes and heights while suppressing most window-induced artefacts.

(magenta) reproduces both the widths and heights of the three true peaks recovered by an FFT of the uniformly sampled series (black); by contrast, the Lomb–Scargle periodogram (red dashed) breaks each broad feature into a forest of narrow side-lobes. This benchmark demonstrates that, in an idealised setting, the regularisation method can suppress window artefacts and preserve multi-component structure.

The outcome is markedly different when the same procedure is applied to the *NuSTAR* light curve of M82. In that real data set, the regularised spectrum shows pronounced distortions: several peaks are broadened or displaced with respect to their Lomb–Scargle counterparts, and weaker harmonics are no longer recovered reliably. These deficiencies point to additional complications present in genuine astrophysical data—overlapping features, a strong red-noise continuum, and non-Gaussian statistics—that interact unfavourably with the current choice of regularisation strength and noise model. Future work will therefore focus on tuning the regularisation parameter, incorporating an explicit red-noise component, and exploring adaptive pre-filtering, with the aim of extending the technique to complex X-ray timing observations.

Appendix A
Complete NuSTAR Data Set and Best-Fit Timing Parameters

Table 1 lists, for every NuSTAR observation analysed in Chapter 3, the best-fit parameters obtained from the broadband power-spectral modelling described in Section 3.4.2. Each row corresponds to one ObsID; asterisks (*) flag observations that are strictly simultaneous with *Chandra* exposures used for source identification. The columns are grouped as follows:

- ***M*** — number of independent 1024-s periodogram segments averaged for that observation;
- $\boldsymbol{rms_{red}}$ — fractional rms (3–79 keV) of the broadband red-noise component;
- **Exponential/Power-law block** — amplitude, break frequency x_0, slope α_0, and high-frequency cut-off adopted for the underlying red-noise continuum when a simple power law was inadequate;
- **Feature block** — parameters of any broad Lorentzian "hump'' required in addition to the continuum (amplitude, centroid x_0, and full width at half-maximum, FWHM);
- **QPO block** — fractional rms of the quasi-periodic oscillation, Lorentzian amplitude and centroid frequency x_1, and FWHM_1. These values underpin the QPO significance and energy-dependence results in Sections 3.4.4–3.5.

All uncertainties are quoted at the 90 % confidence level. Symmetric errors use the "±" notation; asymmetric errors are given as $^{+\mathrm{upper}}_{-\mathrm{lower}}$. A dash (–) indicates that the corresponding parameter was either unconstrained or not required by the preferred model. This comprehensive table is provided to document the full data set and to facilitate reproduction or extension of the analyses presented in Chapter 3.

	M	rms_{red}	Exponential/PowerLaw				Feature			QPO		
			amp	x_0	0	x_{cutoff}	amp	x_0	FWHM	rms_{qpo}	amp_1	x_1
45002*	343	0.114±0.004	–	–	–	–	$1.893^{+0.255}_{-0.260}$	$0.000^{+0.001}_{-0.000}$	$0.020^{+0.015}_{-0.007}$	0.043±0.003	$0.41^{+0.102}_{-0.101}$	$0.111^{+0.0}_{-0.0}$
22002	72	0.037±0.002	–	–	–	–	$0.216^{+0.185}_{-0.105}$	$0.004^{+0.012}_{-0.004}$	$0.024^{+0.031}_{-0.021}$	–	$0.058^{+0.056}_{-0.046}$	$0.146^{+0.0}_{-0.0}$
22004	81	0.054±0.007	–	–	–	–	$1.117^{+0.639}_{-0.450}$	$0.003^{+0.003}_{-0.003}$	$0.002^{+0.002}_{-0.002}$	–	–	–
22008	84	0	–	–	–	–	–	–	–	–	–	–
22010	77	0.117±0.005	–	–	–	–	$1.591^{+0.202}_{-0.229}$	$0.000^{+0.000}_{-0.000}$	$0.014^{+0.008}_{-0.007}$	–	–	–
20002	166	0.010±0.0001	–	–	–	–	–	–	–	0.044±0.002	$0.198^{+0.034}_{-0.033}$	$0.28^{+0.0}_{-0.0}$
20004	170	0.272	$0.002^{+0.002}_{-0.001}$	0.1	$1.735^{+0.182}_{-0.181}$	–	–	–	–	–	–	–
21002*	160	0.047±0.003	–	–	–	–	$0.263^{+0.069}_{-0.069}$	$0.001^{+0.002}_{-0.001}$	$0.102^{+0.054}_{-0.046}$	0.076±0.005	$0.457^{+0.101}_{-0.102}$	$0.123^{+0.0}_{-0.0}$
21004	152	0.046±0.005	$0.164^{+0.065}_{-0.062}$	$0.068^{+0.103}_{-0.068}$	$0.081^{+0.163}_{-0.081}$	$0.259^{+0.190}_{-0.165}$	–	–	–	0.060±0.005	$0.369^{+0.103}_{-0.099}$	$0.178^{+0.0}_{-0.0}$
22002	164	0.177	$0.056^{+0.030}_{-0.022}$	0.1	$0.863^{+0.123}_{-0.112}$	–	–	–	–	0.080±0.005	$1.182^{+0.289}_{-0.285}$	$0.046^{+0.0}_{-0.0}$
22004	180	0.147±0.008	–	–	–	–	$3.064^{+0.595}_{-0.539}$	$0.003^{+0.005}_{-0.003}$	$0.003^{+0.002}_{-0.001}$	0.077±0.004	$0.338^{+0.070}_{-0.067}$	$0.059^{+0.0}_{-0.0}$
19002	56	0.060±0.006	–	–	–	–	$1.088^{+0.448}_{-0.332}$	$0.007^{+0.002}_{-0.001}$	$0.006^{+0.003}_{-0.003}$	0.083±0.001	$0.807^{+0.029}_{-0.028}$	$0.035^{+0.0}_{-0.0}$
19004	305	0.055±0.003	$0.505^{+0.102}_{-0.083}$	$1.161^{+0.196}_{-0.173}$	$0.001^{+0.004}_{-0.001}$	$0.044^{+0.027}_{-0.025}$	–	–	–	0.067±0.003	$0.883^{+0.126}_{-0.170}$	$0.037^{+0.0}_{-0.0}$
92002*	124	0.095±0.005	–	–	–	–	$1.929^{+0.423}_{-0.385}$	$0.004^{+0.001}_{-0.001}$	$0.004^{+0.001}_{-0.001}$	0.080±0.005	$0.903^{+0.201}_{-0.186}$	$0.044^{+0.0}_{-0.0}$
92004*	176	0.075±0.022	–	–	–	–	$0.930^{+0.215}_{-0.181}$	$0.004^{+0.001}_{-0.004}$	$0.012^{+0.020}_{-0.004}$	0.082±0.004	$0.760^{+0.150}_{-0.133}$	$0.051^{+0.0}_{-0.0}$
92006*	583	0.040±0.002	–	–	–	–	$1.142^{+0.217}_{-0.181}$	$0.001^{+0.001}_{-0.001}$	$0.012^{+0.013}_{-0.004}$	0.043±0.003	$0.505^{+0.106}_{-0.100}$	$0.061^{+0.0}_{-0.0}$
92007*	579	0.050±0.003	–	–	–	–	$0.718^{+0.158}_{-0.128}$	$0.000^{+0.002}_{-0.000}$	$0.016^{+0.030}_{-0.004}$	0.070±0.004	$0.402^{+0.088}_{-0.084}$	$0.068^{+0.0}_{-0.0}$
92008*	61	0.062±0.004	$0.569^{+0.143}_{-0.122}$	$0.005^{+0.010}_{-0.005}$	$0.004^{+0.006}_{-0.004}$	$0.029^{+0.018}_{-0.013}$	–	–	–	0.060±0.005	$0.412^{+0.124}_{-0.106}$	$0.108^{+0.0}_{-0.0}$
92009*	341	0.099±0.005	$0.840^{+0.326}_{-0.375}$	$0.004^{+0.006}_{-0.002}$	$0.548^{+0.130}_{-0.166}$	$0.147^{+0.165}_{-0.114}$	–	–	–	0.067±0.004	$0.516^{+0.110}_{-0.104}$	$0.099^{+0.0}_{-0.0}$
92011*	219	0.053±0.010	–	–	–	–	$0.466^{+0.129}_{-0.094}$	$0.000^{+0.022}_{-0.000}$	$0.062^{+0.047}_{-0.029}$	0.069±0.004	$0.960^{+0.182}_{-0.174}$	$0.065^{+0.0}_{-0.0}$
20002	69	0.125	–	–	–	–	–	–	–	–	–	–
20004***	58	0.081±0.001	–	–	–	–	–	–	–	–	–	–
20006	57	0.084±0.006	–	–	–	–	$1.180^{+0.289}_{-0.274}$	$0.001^{+0.003}_{-0.001}$	$0.011^{+0.006}_{-0.005}$	0.055±0.005	$0.087^{+0.029}_{-0.028}$	$0.161^{+0.1}_{-0.1}$
20008	80	0.160±0.036	$0.022^{+0.019}_{-0.014}$	0.1	$0.978^{+0.211}_{-0.186}$	–	–	–	–	0.122±0.011	$0.511^{+0.157}_{-0.139}$	$0.157^{+0.0}_{-0.0}$
05002	73	0.211±0.009	–	–	–	–	$2.293^{+0.318}_{-0.357}$	$0.000^{+0.001}_{-0.000}$	$0.011^{+0.008}_{-0.006}$	–	–	–
37002*	164	0.096±0.004	–	–	–	–	$1.180^{+0.289}_{-0.274}$	$0.001^{+0.003}_{-0.001}$	$0.011^{+0.006}_{-0.005}$	–	$0.087^{+0.029}_{-0.028}$	$0.161^{+0.1}_{-0.1}$
38002	88	0.06	$1.893^{+0.255}_{-0.260}$	0	–	–	–	–	–	–	–	–
38004	81	0.039	–	–	–	–	–	–	–	0.084±0.005	$0.309^{+0.065}_{-0.058}$	$0.221^{+0.0}_{-0.0}$
27002*	173	0.053±0.005	–	–	–	–	$1.206^{+0.481}_{-0.342}$	$0.003^{+0.001}_{-0.000}$	$0.003^{+0.001}_{-0.001}$	0.067±0.004	$0.171^{+0.048}_{-0.028}$	$0.130^{+0.0}_{-0.0}$
27004*	170	0.128±0.018	$0.606^{+0.244}_{-0.322}$	$0.088^{+0.032}_{-0.033}$	$0.094^{+0.140}_{-0.093}$	$0.104^{+0.046}_{-0.043}$	–	–	–	0.036±0.008	$0.404^{+0.344}_{-0.212}$	$0.153^{+0.1}_{-0.0}$
28002	159	0.043±0.003	–	–	–	–	$0.318^{+0.085}_{-0.075}$	$0.004^{+0.013}_{-0.004}$	$0.069^{+0.041}_{-0.036}$	0.079±0.005	$0.413^{+0.093}_{-0.086}$	$0.163^{+0.0}_{-0.0}$
28004	145	–	–	–	–	–	–	–	–	–	–	–
12002*	273	0.141±0.006	–	–	–	–	$2.025^{+0.291}_{-0.349}$	$0.000^{+0.001}_{-0.000}$	$0.014^{+0.012}_{-0.006}$	0.011±0.002	$0.343^{+0.096}_{-0.090}$	$0.140^{+0.0}_{-0.0}$
12004	211	0.078	–	–	–	–	–	–	–	0.094±0.047	$0.267^{+0.043}_{-0.039}$	$0.124^{+0.0}_{-0.0}$
38002	218	0.391±0.097	$0.002^{+0.002}_{-0.001}$	0.1	$1.661^{+0.153}_{-0.136}$	–	–	–	–	–	–	–
32002	117	1.32±0.008	$1.427^{+0.029}_{-0.029}$	0.1	$0.500^{+0.008}_{-0.009}$	–	–	–	–	–	–	–
33002	93	0.141±0.007	–	–	–	–	$2.081^{+0.380}_{-0.359}$	$0.002^{+0.002}_{-0.002}$	$0.010^{+0.004}_{-0.002}$	0.083±0.007	$0.567^{+0.168}_{-0.143}$	$0.129^{+0.0}_{-0.0}$
22003	110	0.076	–	–	–	–	$0.928^{+0.700}_{-0.401}$	$0.004^{+0.014}_{-0.004}$	$0.036^{+0.031}_{-0.031}$	0.036±0.019	$0.179^{+0.114}_{-0.081}$	$0.150^{+0.0}_{-0.0}$
22006	3	0.15	–	–	–	–	–	–	–	–	–	–

Bibliography

Abbott, B. P., et al. 2016, Physical Review Letters, 116, 061102, doi: 10.1103/PhysRevLett.116.061102 16

Ade, P. a. R., Aghanim, N., Argüeso, F., et al. 2014, Astronomy & Astrophysics, 571, A28, doi: 10.1051/0004-6361/201321524 69

—. 2016, Astronomy & Astrophysics, 594, A26, doi: 10.1051/0004-6361/201526914 69, 70

Akaike, H. 1974, IEEE Transactions on Automatic Control, 19, 716, doi: 10.1109/TAC.1974.1100705 57

Akrami, Y., Ashdown, M., Aumont, J., & et al. 2020a, Astronomy & Astrophysics, 644, A99 70

Akrami, Y., Ashdown, M., Aumont, J., et al. 2020b, Astronomy & Astrophysics, 644, A99 70

Alger, M. J., Banfield, J. K., Wong, O. I., et al. 2018, Monthly Notices of the Royal Astronomical Society, 478, 5547, doi: 10.1093/mnras/sty1410 40

Alpar, M. A., & Shaham, J. 1985, Nat., 316, 239 66

Alsing, J., Charnock, T., Feeney, S., & Wandelt, B. 2019, Monthly Notices of the Royal Astronomical Society, 488, 4440 41

Atapin, K., Fabrika, S., & Caballero-García, M. D. 2019, Monthly Notices of the Royal Astronomical Society, 486, 2766, doi: 10.1093/mnras/stz1027 63

Baade, W., & Zwicky, F. 1934, Proceedings of the National Academy of Sciences, 20, 254, doi: 10.1073/pnas.20.5.254 15

Bachetti, M., Harrison, F., Walton, D. J., et al. 2014, Nature, 514, 202 25, 26, 27, 46

Bai, S., Kolter, J. Z., & Koltun, V. 2018, An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. https://arxiv.org/abs/1803.01271 42

Baron, D. 2019, arXiv preprint arXiv:1904.07248 70

Barret, D. 2012, The Astrophysical Journal, 753, 84 36

Barret, D., Kluzniak, W., Olive, J. F., Paltani, S., & Skinner, G. K. 2005, MNRAS, 357, 1288, doi: 10.1111/j.1365-2966.2005.08734.x 60

Barret, D., & Vaughan, S. 2012, The Astrophysical Journal, 746, 131 53, 55

Bartlett, M. S. 1950a, Biometrika, 37, 1 34

—. 1950b, Biometrika, 37, 1 52

Belloni, T. 2010, in Lecture Notes in Physics, ed. T. Belloni, Vol. 794 (Springer), 53–84, doi: 10.1007/978-3-642-02248-7_2 23

Belloni, T., & Hasinger, G. 1990, A&A, 230, 103 55, 63

Belloni, T., Homan, J., Casella, P., et al. 2005, Astronomy and Astrophysics, 440, 207, doi: 10.1051/0004-6361:20042457 48

Belloni, T., Psaltis, D., & van der Klis, M. 2002, The Astrophysical Journal, 572, 392, doi: 10.1086/340290 38, 53, 55

Belloni, T. M., & Stella, L. 2012, Space Science Reviews, 183, 43, doi: 10.1007/s11214-012-9922-0 48

—. 2014, Space Science Reviews, 183, 43 23

Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, The Astrophysical Journal Supplement Series, 208, 20, doi: 10.1088/0067-0049/208/2/20 69

Bhattacharyya, S., & Strohmayer, T. E. 2007, The Astrophysical Journal Letters, 664, L103, doi: 10.1086/520844 21

Bhavanam, S. R., Channappayya, S. S., P. K, S., & Desai, S. 2024, Astrophysics and Space Science, 369, 92 71

Bolton, C. T. 1972, Nature, 235, 271, doi: 10.1038/235271b0 15

Bonavera, L., Gomez, S. S., González-Nuevo, J., et al. 2021, Astronomy & Astrophysics, 648, A50 69

Bonjean, F. 2020, Astronomy & Astrophysics, 635, A10 70

Bouchard, R., Corbel, S., Rodriguez, J., Clavel, M., & Bird, A. J. 2019, The Astrophysical Journal, 878, 90, doi: 10.3847/1538-4357/ab1f6e 42, 43

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. 2015, Time series analysis: Forecasting and control, 5th edn. (Hoboken, New Jersey: John Wiley & Sons) 31, 32

Braun, R., Bourke, T. L., Green, J. A., Keane, E., & Wagg, J. 2015, in Advancing Astrophysics with the Square Kilometre Array, Vol. 215, Sissa Medialab, 174 44

Brightman, M., Walton, D. J., Xu, Y., et al. 2020, The Astrophysical Journal, 889, 71, doi: 10.3847/1538-4357/ab629a 65

Brockwell, P. J., & Davis, R. A. 2016, Introduction to time series and forecasting, 3rd edn. (Cham, Switzerland: Springer), doi: 10.1007/978-3-319-29854-2 31, 32

Brunel, A., Pasquet, J., Pasquet, J., et al. 2019, Monthly Notices of the Royal Astronomical Society, 489, 3794 42

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. 1995, SIAM J. Sci. Comput., 16, 1190, doi: 10.1137/0916069 57

Caballero, I., & Wilms, J. 2012, arXiv preprint arXiv:1206.3124 25

Cackett, E. M., Miller, J. M., Bhattacharyya, S., et al. 2008, The Astrophysical Journal, 674, 415 18

Campana, S., Coti Zelati, F., & D'Avanzo, P. 2013, Monthly Notices of the Royal Astronomical Society, 432, 1695 22

Carvalho, P., Rocha, G., & Hobson, M. 2009, Monthly Notices of the Royal Astronomical Society, 393, 681 70

Casas, J. M., & et al. 2022, Astronomy & Astrophysics, 658, A110, doi: 10.1051/0004-6361/202141874 70, 71

Casella, P., Belloni, T., Homan, J., & Stella, L. 2005, Astronomy and Astrophysics, 429, 403, doi: 10.1051/0004-6361:20041570 23

Cheng, G., Zhou, P., & Han, J. 2016, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2884–2893 43

Chevalier, R. A., & Clegg, A. W. 1985, Nature, 317, 44, doi: 10.1038/317044a0 28

Ciuciu, P., & Others, B. 2001, IEEE Transactions on Signal Processing, 49, 2202, doi: 10.1109/78.950795 85

Colbert, E. J. M., & Mushotzky, R. F. 1999, The Astrophysical Journal, 519, 89, doi: 10.1086/307356 25

Collaboration, P. 2014, Astronomy & Astrophysics, 571, A1 69

Connor, L., & van Leeuwen, J. 2018, The Astrophysical Journal, 865, 15 43

Cooley, J. W., & Tukey, J. W. 1965, Math. Comput., 19, 297, doi: 10.1090/S0025-5718-1965-0178586-1 53

Coppi, P. S. 1999, ASP Conference Series, 161, 375 20

Dauser, T., García, J., Walton, D. J., et al. 2016, Astronomy & Astrophysics, 590, A76, doi: 10.1051/0004-6361/201628135 21

de Grijs, R., O'Connell, R. W., & Gallagher, J. S. 2001, The Astronomical Journal, 121, 768, doi: 10.1086/318769 45

Desai, A., Freeman, C., Wang, Z., & Beaver, I. 2021, arXiv preprint arXiv:2111.08095 43

Dewangan, G. C., Titarchuk, L., & Griffiths, R. E. 2006, The Astrophysical Journal Letters, 637, L21, doi: 10.1086/499235 58

Dieleman, S., Willett, K. W., & Dambre, J. 2015, Monthly Notices of the Royal Astronomical Society, 450, 1441, doi: 10.1093/mnras/stv632 40

Dihingia, I. K., & Fendt, C. 2025, in New Frontiers in GRMHD Simulations (Springer), 327–360 18

Done, C., Gierliński, M., & Kubota, A. 2007, The Astronomy and Astrophysics Review, 15, 1 15, 16, 18, 20

Efron, B. 1987, Journal of the American statistical Association, 82, 171 35

Efron, B., & LePage, R. 1992, Introduction to bootstrap (Wiley & Sons, New York) 35

Einstein, A. 1916, Annalen der Physik, 49, 769, doi: 10.1002/andp.19163540702 15

Fabian, A. C., Rees, M. J., Stella, L., & White, N. E. 1989, Monthly Notices of the Royal Astronomical Society, 238, 729 22

Fabian, A. C., & Ross, R. R. 2010, Space Science Reviews, 157, 167, doi: 10.1007/s11214-010-9699-y 20

Farrell, S. A., Webb, N. A., Barret, D., Godet, O., & Rodrigues, J. M. 2009, Nature, 460, 73, doi: 10.1038/nature08083 26

Feng, H., & Kaaret, P. 2007, The Astrophysical Journal, 668, 941 63

Feng, H., Rao, F., & Kaaret, P. 2010, The Astrophysical Journal Letters, 710, L137 65

Frank, J., King, A., & Raine, D. 2002, Accretion Power in Astrophysics, 3rd edn. (Cambridge University Press) 7, 16, 19

Fürst, F., Walton, D. J., Israel, G. L., et al. 2023, Astronomy and Astrophysics, 672, A140, doi: 10.1051/0004-6361/202245048 66

Gagliano, A., & Villar, V. A. 2023, arXiv preprint arXiv:2312.16687 71

Gal, Y., & Ghahramani, Z. 2016, in international conference on machine learning, PMLR, 1050–1059 44

Galloway, D. K., Chenevez, J., Wörpel, H., et al. 2020, The Astrophysical Journal Supplement Series, 249, 32 16

Gao, Y., Wang, Q. D., Appleton, P. N., & Lucas, R. A. 2003, The Astrophysical Journal, 596, L171, doi: 10.1086/379335 27

García, J., Dauser, T., Reynolds, C. S., et al. 2013, The Astrophysical Journal, 768, 146 20

García, J. A., Steiner, J. F., McClintock, J. E., et al. 2015, The Astrophysical Journal, 813, 84, doi: 10.1088/0004-637X/813/2/84 20

George, D., & Huerta, E. 2018, Physics Letters B, 778, 64 43

Giacconi, R., Gursky, H., Kellogg, E., Schreier, E., & Tananbaum, H. 1971, The Astrophysical Journal, 167, L67, doi: 10.1086/180762 25

Gierliński, M., & Done, C. 2004, Monthly Notices of the Royal Astronomical Society, 347, 885, doi: 10.1111/j.1365-2966.2004.07266.x 18

Gladstone, J. C., Roberts, T. P., & Done, C. 2009, MNRAS, 397, 1836, doi: 10.1111/j.1365-2966.2009.15039.x 25, 26

Godet, O., Lombardi, J., Antonini, F., et al. 2014, The Astrophysical Journal, 793, 105 26

González-Martín, O., & Vaughan, S. 2012, Astron. Astrophys., 544, A80, doi: 10.1051/0004-6361/201219008 39

Griffiths, R. E., Ptak, A. F., Feigelson, E. D., Garmire, G. P., & Townsley, L. K. 2000, Science, 290, 1325, doi: 10.1126/science.290.5491.1325 27

Gúrpide, A., & Middleton, M. 2025, Monthly Notices of the Royal Astronomical Society, 537, 3210, doi: 10.1093/mnras/staf196 81

Hall, P., Horowitz, J. L., & Jing, B.-Y. 1995, Biometrika, 82, 561 35

Hameury, J.-M., Viallet, M., & Lasota, J.-P. 2009, Astronomy & Astrophysics, 496, 413 22

Heil, L., Uttley, P., & Klein-Wolt, M. 2015, Monthly Notices of the Royal Astronomical Society, 448, 3348 23

Hewish, A., Bell, S. J., Pilkington, J. D. H., Scott, P. F., & Collins, R. A. 1968, Nature, 217, 709, doi: 10.1038/217709a0 15

Holmberg, E. 1958, Meddelanden fran Lunds Astronomiska Observatorium Serie II, 136, 1 45

Hon, M., Stello, D., & Yu, J. 2018, Monthly Notices of the Royal Astronomical Society, 476, 3233 42

Huertas-Company, M., Bernardi, M., Decenciére, E., et al. 2018, Astronomy and Computing, 25, 103, doi: 10.1016/j.ascom.2018.08.004 40

Hulse, R. A., & Taylor, J. H. 1975, Astrophysical Journal, 195, L51, doi: 10.1086/181708 16

Huppenkothen, D., Bachetti, M., Stevens, A. L., et al. 2019, ApJ, 881, 39, doi: 10.3847/1538-4357/ab258d 49, 55

Ingram, A., & Done, C. 2011, Monthly Notices of the Royal Astronomical Society, 415, 2323, doi: 10.1111/j.1365-2966.2011.18860.x 65

Ingram, A., Done, C., & Fragile, P. C. 2009, Monthly Notices of the Royal Astronomical Society: Letters, 397, L101, doi: 10.1111/j.1745-3933.2009.00693.x 23, 24, 39, 48

Ingram, A., & Klis, M. v. d. 2013, Monthly Notices of the Royal Astronomical Society, 434, 1476 23

Ingram, A. R., & Motta, S. E. 2019, New Astronomy Reviews, 85, 101524 23, 24, 38, 48, 65

Ioffe, S., & Szegedy, C. 2015, in Proceedings of the 32nd International Conference on Machine Learning (ICML), Vol. 37, 448–456. http://proceedings.mlr.press/v37/ioffe15.html 40, 42

Israel, G. L., Belfiore, A., Stella, L., et al. 2017a, Science, 355, 817, doi: 10.1126/science.aai8635 25, 26

Israel, G. L., Papitto, A., Esposito, P., et al. 2017b, MNRAS, 466, L48, doi: 10.1093/mnrasl/slw216 25, 26

Israel, G. L., Belfiore, A., Stella, L., et al. 2017, Science, 355, 817, doi: 10.1126/science.aai8635 66

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, The Astrophysical Journal, 873, 111 44

Jiang, Y.-F., Stone, J. M., & Davis, S. W. 2014, The Astrophysical Journal, 784, 169 18

Kaaret, P., Feng, H., & Roberts, T. P. 2017a, Annual Review of Astronomy and Astrophysics, 55, 303 16

—. 2017b, Annual Review of Astronomy and Astrophysics, 55, 303 25

Kaaret, P., Prestwich, A. H., Zezas, A., et al. 2001, Monthly Notices of the Royal Astronomical Society, 321, L29, doi: 10.1046/j.1365-8711.2001.04323.x 27

Kaaret, P., Simet, M. G., & Lang, C. C. 2006, ApJ, 646, 174, doi: 10.1086/504830 49

Kawashima, T., Mineshige, S., Ohsuga, K., & Ogawa, T. 2016, Publications of the Astronomical Society of Japan, 68, 83 26

Kay, S. M., & Marple, S. L. 1981, Proc. IEEE, 69, 1380, doi: 10.1109/PROC.1981.12184 34

King, A. R., Davies, M. B., Ward, M. J., Fabbiano, G., & Elvis, M. 2001, Astrophysical Journal Letters, 552, L109, doi: 10.1086/320262 25

Kluźniak, W., & Abramowicz, M. A. 2001, Acta Physica Polonica B, 32, 3605. https://www.actaphys.uj.edu.pl/R/32/11/3605/pdf 24

Kluźniak, W. 2006, Acta Physica Polonica B, 37, 1361 48

Krawczynski, H., Muleri, F., Dovčiak, M., et al. 2022, Science, 378, 650 18

Kretschmar, P., Martínez-Núñez, S., Fürst, F., et al. 2019, arXiv preprint arXiv:1905.08578 16

Krizhevsky, A., Sutskever, I., & Hinton, G. E. 2012, in Advances in Neural Information Processing Systems (NIPS), 1097–1105. https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html 40

Lakshminarayanan, B., Pritzel, A., & Blundell, C. 2017, Advances in neural information processing systems, 30 44

Leahy, D. A., Darbro, W., Elsner, R. F., et al. 1983, The Astrophysical Journal, 266, 160, doi: 10.1086/160766 52

LeCun, Y., Bengio, Y., & Hinton, G. 2015, nature, 521, 436 41

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. 1998, Proceedings of the IEEE, 86, 2278, doi: 10.1109/5.726791 40

Lewin, W. H. G., van Paradijs, J., & Taam, R. E. 1997, Annual Review of Astronomy and Astrophysics, 35, 307, doi: 10.1146/annurev.astro.35.1.307 15

Lin, D., Remillard, R. A., & Homan, J. 2007, The Astrophysical Journal, 667, 1073, doi: 10.1086/521181 18, 20

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. 2017, in Proceedings of the IEEE international conference on computer vision, 2980–2988 44

Linden, T., Kalogera, V., Sepinsky, J. F., Vigna-Gómez, A., & Mushotzky, R. F. 2010, The Astrophysical Journal, 725, 1984, doi: 10.1088/0004-637X/725/2/1984 28

Liska, M., Tchekhovskoy, A., Ingram, A., & van der Klis, M. 2019, Monthly Notices of the Royal Astronomical Society, 487, 550 18

Liu, J.-F., & Bregman, J. N. 2005, ApJ Supp. Ser., 157, 59, doi: 10.1086/427170 9, 64, 66

Lomb, N. R. 1976, Astrophysics and Space Science, 39, 447, doi: 10.1007/BF00648343 67, 81, 82

Magdziarz, P., & Zdziarski, A. A. 1995, Monthly Notices of the Royal Astronomical Society, 273, 837, doi: 10.1093/mnras/273.3.837 20

Makishima, K., Kubota, A., Mizuno, T., et al. 2000, ApJ, 535, 632, doi: 10.1086/308868 25

Manikantan, H., Paul, B., Sharma, R., Pradhan, P., & Rana, V. 2024, Monthly Notices of the Royal Astronomical Society, 531, 530, doi: 10.1093/mnras/stae1170 66

Mapelli, M., Ripamonti, E., Zampieri, L., Colpi, M., & Bressan, A. 2010, Monthly Notices of the Royal Astronomical Society, 408, 234, doi: 10.1111/j.1365-2966.2010.17119.x 27, 28

Marcel, G., Cangemi, F., Rodriguez, J., et al. 2020, Astronomy & Astrophysics, 640, A18 22

Mayya, Y. D., Bressan, A., Carrasco, L., & Hernandez-Martinez, L. 2006, Astrophysical Journal, 649, 172, doi: 10.1086/505169 45

McClintock, J. E., Narayan, R., & Steiner, J. F. 2014, Space Science Reviews, 183, 295, doi: 10.1007/s11214-013-0003-9 18, 21

McHardy, I. M., Koerding, E., Knigge, C., Uttley, P., & Fender, R. P. 2006, Nature, 444, 730, doi: 10.1038/nature05389 39

Mezcua, M. 2017, International Journal of Modern Physics D, 26, 1730021 26

Michelucci, U., & Venturini, F. 2021, Machine Learning and Knowledge Extraction, 3, 357 41

Miller, C. 2002, Nature, 420, 31 22

Miller, J. M. 2007, Annual Review of Astronomy and Astrophysics, 45, 441 20, 21

Miller, J. M., Fabian, A. C., & Miller, M. C. 2004, The Astrophysical Journal, 607, 931, doi: 10.1086/383564 25

Miller, J. M., Homan, J., Steeghs, D., et al. 2006, The Astrophysical Journal, 653, 525, doi: 10.1086/508644 18

Miller, J. M., Reynolds, C. S., Fabian, A. C., Miniutti, G., & Gallo, L. C. 2009, The Astrophysical Journal, 706, L103, doi: 10.1088/0004-637X/706/2/L103 18

Miller, J. M., Parker, M. L., Fuerst, F., et al. 2013, The Astrophysical Journal Letters, 779, L2, doi: 10.1088/2041-8205/779/1/L2 21

Möller, A., & de Boissière, T. 2020, Monthly Notices of the Royal Astronomical Society, 491, 4277, doi: 10.1093/mnras/stz3312 41

Moschou, S., Hicks, E., Parekh, R., et al. 2023a, Machine Learning: Science and Technology, 4, 035032 44

—. 2023b, Machine Learning: Science and Technology, 4, 035032 44

Motta, S., Homan, J., Munoz-Darias, T., et al. 2012, Monthly Notices of the Royal Astronomical Society, 427, 595 24

Motta, S., Rouco-Escorial, A., Kuulkers, E., Muñoz-Darias, T., & Sanna, A. 2017, Monthly Notices of the Royal Astronomical Society, 468, 2311 48

Motta, S., Rodriguez, J., Jourdain, E., et al. 2021, New Astronomy Reviews, 93, 101618 23

Motta, S. E., & Belloni, T. M. 2024, Astronomy and Astrophysics, 684, A209, doi: 10.1051/0004-6361/202347331 24

Motta, S. E., Belloni, T. M., Stella, L., Muñoz-Darias, T., & Fender, R. 2014, Monthly Notices of the Royal Astronomical Society, 437, 2554 24

Motta, S. E., Casella, P., & Henze, T. 2016, Astronomy & Astrophysics Review, 24, 7, doi: 10.1007/s00159-016-0094-2 48, 55, 65

Mucciarelli, P., Casella, P., Belloni, T., Zampieri, L., & Ranalli, P. 2006, Monthly Notices of the Royal Astronomical Society, 365, 1123, doi: 10.1111/j.1365-2966.2005.09754.x 46, 58, 63

Naul, B., Bloom, J. S., Pérez, F., & van der Walt, S. 2018, Nature Astronomy, 2, 151 43

Neal, R. M. 2012, Bayesian learning for neural networks, Vol. 118 (Springer Science & Business Media) 44

Neumann, M., Avakyan, A., Doroshenko, V., & Santangelo, A. 2023, Astronomy & Astrophysics, 677, A134 22

Nowak, M. A. 2000, Mon. Not. R. Astron. Soc., 318, 361, doi: 10.1046/j.1365-8711.2000.03668.x 38

O'Connell, R. W., Gallagher, J. S., Hunter, D. A., & Colley, W. N. 1995, The Astrophysical Journal, 446, 176, doi: 10.1086/175204 27

Padovani, P., Alexander, D. M., Assef, R. J., et al. 2017, Astronomy & Astrophysics Review, 25, 2, doi: 10.1007/s00159-017-0102-9 47

Papamakarios, G., Sterratt, D., & Murray, I. 2019, in Proceedings of Machine Learning Research, Vol. 89, Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics, ed. K. Chaudhuri & M. Sugiyama (PMLR), 837–848. https://proceedings.mlr.press/v89/papamakarios19a.html 41

Pasham, D. R., Cenko, S. B., Zoghbi, A., et al. 2015, The Astrophysical Journal Letters, 811, L11 46

Pasham, D. R., Strohmayer, T. E., & Mushotzky, R. F. 2014a, Nature, 513, 74, doi: 10.1038/nature13707 26, 27, 59, 65, 66

—. 2014b, Nature, 513, 74 26

Patruno, A., & Watts, A. L. 2020, in Timing Neutron Stars: Pulsations, Oscillations and Explosions (Springer), 143–208 25

Penna, R. F., McKinney, J. C., Narayan, R., et al. 2010, Monthly Notices of the Royal Astronomical Society, 408, 752 18

Percival, D. B., & Walden, A. T. 1993, Spectral analysis for physical applications (Cambridge: Cambridge University Press), doi: 10.1017/CBO9780511622762 34

Persic, M., & Rephaeli, Y. 2002, Astronomy & Astrophysics, 382, 843, doi: 10.1051/0004-6361:20011631 27

Pinto, C., & Walton, D. J. 2023, Ultra-Luminous X-ray Sources: Extreme Accretion and Feedback, doi: 10.48550/arXiv.2302.00006 66

Plant, D. S., Fender, R. P., Ponti, G., Muñoz-Darias, T., & Coriat, M. 2015, Astronomy & Astrophysics, 573, A120 18

Priestley, M. B. 1981, Spectral analysis and time series (London: Academic Press) 32, 33

Pruzhinskaya, M. V., Malanchev, K. L., Kornilov, M. V., et al. 2019, Monthly Notices of the Royal Astronomical Society, 489, 3591 43

Psaltis, D., & Norman, C. 2000, On the Origin of Quasi-Periodic Oscillations and Broad-band Noise in Accreting Neutron Stars and Black Holes, arXiv, doi: 10.48550/arXiv.astro-ph/0001391 65

Raissi, M., Perdikaris, P., & Karniadakis, G. E. 2019, Journal of Computational physics, 378, 686 44

Ramos, D., Gonzalez-Rodriguez, J., & Torralba, A. 2018, Entropy, 20, 208, doi: 10.3390/e20030208 75

Ransom, S. M., Eikenberry, S. S., & Middleditch, J. 2002, The Astronomical Journal, 124, 1788Figure, doi: 10.1086/342285 53

Reis, R. C., Fabian, A. C., & Miller, J. M. 2010, Monthly Notices of the Royal Astronomical Society, 402, 836, doi: 10.1111/j.1365-2966.2009.15976.x 18

Remillard, R. A., & McClintock, J. E. 2006, Annual Review of Astronomy and Astrophysics, 44, 49, doi: 10.1146/annurev.astro.44.051905.092532 18, 22, 24, 48

Revnivtsev, M., Churazov, E., Gilfanov, M., & Sunyaev, R. 2001, A&A, 372, 138, doi: 10.1051/0004-6361:20010434 65

Reynolds, C. S. 2014, Space Science Reviews, 183, 277, doi: 10.1007/s11214-013-0006-6 21

Reynolds, C. S., & Nowak, M. A. 2003, Physics Reports, 377, 389, doi: 10.1016/S0370-1573(02)00584-7 21, 22

Rodi, J., Jourdain, E., & Roques, J. 2017, The Astrophysical Journal, 848, 3 22

Ronneberger, O., Fischer, P., & Brox, T. 2015, Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015, 234 75

Sanders, D. B., Mazzarella, J. M., Kim, D.-C., Surace, J. A., & Soifer, B. T. 2003, The Astronomical Journal, 126, 1607, doi: 10.1086/376841 45

Sartorio, N. S., Fialkov, A., Hartwig, T., et al. 2023, Monthly Notices of the Royal Astronomical Society, 521, 4039 22

Scargle, J. D. 1982, ApJ, 263, 835, doi: 10.1086/160554 67, 81, 82

Schwarzschild, K. 1916, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften, 189. https://doi.org/10.1142/9789812796869_0009 15

Seghouani, N. 2017, Monthly Notices of the Royal Astronomical Society, 468, 3312, doi: 10.1093/mnras/stx569 85

Sel, K., Mohammadi, A., Pettigrew, R. I., & Jafari, R. 2023, npj Digital Medicine, 6, 110, doi: 10.1038/s41746-023-00853-4 41

Selvaraju, R. R., Cogswell, M., Das, A., et al. 2017, Proceedings of the IEEE international conference on computer vision, 618 43, 44

Shakura, N. I., & Sunyaev, R. A. 1973, Astronomy and Astrophysics, 24, 337 15, 22

Shallue, C. J., & Vanderburg, A. 2018, The Astronomical Journal, 155, 94, doi: 10.3847/1538-3881/aaaa4a 43

Shumway, R. H., Stoffer, D. S., & Stoffer, D. S. 2000, Time series analysis and its applications, Vol. 3 (Springer) 35

Slijepcevic, I. V., Scaife, A. M., Walmsley, M., et al. 2022, Monthly Notices of the Royal Astronomical Society, 514, 2599 71

Sohn, K., Berthelot, D., Carlini, N., et al. 2020, Advances in Neural Information Processing Systems, 33, 596 71, 77

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. 2014, The journal of machine learning research, 15, 1929 42

Steiner, J. F., McClintock, J. E., Remillard, R. A., et al. 2010, The Astrophysical Journal Letters, 718, L117, doi: 10.1088/2041-8205/718/2/L117 18

Stella, L., & Vietri, M. 1998, ApJL, 492, L59, doi: 10.1086/311075 39

—. 1999, Physical Review Letters, 82, 17, doi: 10.1103/PhysRevLett.82.17 65

Strickland, D. K., & Heckman, T. M. 2009, Astrophysical Journal, 697, 2030, doi: 10.1088/0004-637X/697/2/2030 45

Strohmayer, T. E., & Mushotzky, R. F. 2003, The Astrophysical Journal, 586, L61, doi: 10.1086/374091 26, 27, 46, 58, 63, 65

Sundararajan, M., Taly, A., & Yan, Q. 2017, in Proceedings of Machine Learning Research, Vol. 70, Proceedings of the 34th International Conference on Machine Learning, ed. D. Precup & Y. W. Teh (PMLR), 3319–3328 43, 44

Sunyaev, R. A., & Trümper, J. 1979, Nature, 279, 506, doi: 10.1038/279506a0 22

Sutton, A. D., Roberts, T. P., & Middleton, M. J. 2013, MNRAS, 435, 1758, doi: 10.1093/mnras/stt1412 26

Swartz, D. A., Yukita, M., Wu, K., Freeman, P. E., & Tennant, A. F. 2011, The Astrophysical Journal, 741, 49, doi: 10.1088/0004-637X/741/1/49 27

Tauris, T. M., & van den Heuvel, E. P. J. 2006, Formation and Evolution of Compact Stellar X-ray Sources (Cambridge University Press), 623–655 16

Tetarenko, B., Bahramian, A., Arnason, R., et al. 2016, The Astrophysical Journal, 825, 10 16

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., & Farmer, J. D. 1992, Physica D: Nonlinear Phenomena, 58, 77 35, 36

Timmer, J. 1995, Physical Review E, 52, 4844 36, 37, 38

Titarchuk, L. G. 1994, The Astrophysical Journal, 434, 570 20

Tomsick, J. A., Yamaoka, K., Corbel, S., et al. 2009, The Astrophysical Journal Letters, 707, L87, doi: 10.1088/0004-637X/707/1/L87 18

Tomsick, J. A., Nowak, M. A., Parker, M., et al. 2014, The Astrophysical Journal, 780, 78, doi: 10.1088/0004-637X/780/1/78 22

Tutukov, A. V., & Yungelson, L. R. 1973, Nauchnye Informatsii, 27, 70 16

Uttley, P., Cackett, E., Fabian, A., Kara, E., & Wilkins, D. 2014a, The Astronomy and Astrophysics Review, 22, 1 37

Uttley, P., Cackett, E. M., Fabian, A. C., Kara, E., & Wilkins, D. R. 2014b, The Astronomy and Astrophysics Review, 22, 72 39

Uttley, P., McHardy, I. M., & Papadakis, I. 2002, Monthly Notices of the Royal Astronomical Society, 332, 231 37

van der Klis, M. 1989, International Journal of Modern Physics, 4, 215, doi: 10.1142/S0218271800004010 39

Van der Klis, M. 1989, Fourier techniques in X-ray timing (Springer), 27–69 48, 52

van der Klis, M. 2006, In: Compact stellar X-ray sources. Edited by Walter Lewin & Michiel van der Klis. Cambridge Astrophysics Series, 39 22, 24, 48, 65

van der Klis, M., Jansen, F., van Paradijs, J., et al. 1985, Nature, 316, 225, doi: 10.1038/316225a0 23

VanderPlas, J. T. 2018, The Astrophysical Journal Supplement Series, 236, 16, doi: 10.3847/1538-4365/aab766 67

Vaughan, S. 2005, Astronomy & Astrophysics, 431, 391 37

—. 2013, Philos. Trans. R. Soc. A, 371, 20110549, doi: 10.1098/rsta.2011.0549 36

Vaughan, S., Edelson, R., Warwick, R. S., & Uttley, P. 2003, Mon. Not. R. Astron. Soc., 345, 1271, doi: 10.1046/j.1365-2966.2003.07042.x 39

Vaughan, S., Uttley, P., Markowitz, A. G., et al. 2011, Monthly Notices of the Royal Astronomical Society, 413, 2489, doi: 10.1111/j.1365-2966.2011.18358.x 81

Vavilova, I., Pakuliak, L., Babyk, I., et al. 2020, Surveys, Catalogues, Databases, and Archives of Astronomical Data (Elsevier), 57–102 28

Veilleux, S., Cecil, G., & Bland-Hawthorn, J. 2005, Annual Review of Astronomy and Astrophysics, 43, 769, doi: 10.1146/annurev.astro.43.072103.150610 28

Walton, D. J., Mackenzie, A. D. A., Gully, H., et al. 2022, Monthly Notices of the Royal Astronomical Society, 509, 1587, doi: 10.1093/mnras/stab3001 59

Webster, B. L., & Murdin, P. G. 1972, Nature, 235, 37, doi: 10.1038/235037a0 15

Weedman, D. W. 1983, Astrophysics and Space Science Library, 111, 211, doi: 10.1007/978-94-009-3753-6_28 27

Welch, P. D. 1967, IEEE Transactions on Audio and Electroacoustics, 15, 70, doi: 10.1109/TAU.1967.1161901 34

Whittle, P. 1953, Arkiv för matematik, 2, 423 34

Wijnands, R., & van der Klis, M. 1999, ApJ, 514, 939, doi: 10.1086/306993 65

Xiao, H., Ji, L., Tsygankov, S., et al. 2024, A Systematic Study of Millihertz Quasi-periodic Oscillations in GS 1826-238, doi: 10.48550/arXiv.2410.04963 65

Zdziarski, A. A., & Gierliński, M. 2004, Progress of Theoretical Physics Supplement, 155, 99, doi: 10.1143/PTPS.155.99 22

Zdziarski, A. A., Johnson, W. N., Smith, E. P., Madejski, R. S., & Magdziarz, P. 1996, The Astrophysical Journal, 458, 182 20

Zezas, A., & Fabbiano, G. 2002, The Astrophysical Journal, 577, 726, doi: 10.1086/342356 27, 28

Zhang, C. M., Yin, H. X., Zhao, Y. H., Zhang, F., & Song, L. M. 2006, Monthly Notices of the Royal Astronomical Society, 366, 1373, doi: 10.1111/j.1365-2966.2006.09920.x 48

Zhang, S., Santangelo, A., Feroci, M., et al. 2019, Science China Physics, Mechanics & Astronomy, 62, 29502, doi: 10.1007/s11433-018-9309-2 47

La borsa di dottorato è stata cofinanziata con risorse del
Programma Operativo Nazionale Ricerca e Innovazione 2014-2020, risorse FSE REACT-EU
Azione IV.4 "Dottorati e contratti di ricerca su tematiche dell'innovazione"
e Azione IV.5 "Dottorati su tematiche Green"