# The transient X-ray sky of eROSITA: from prediction, through observation, to interpretation

**Adam Malyali**


München 2021

# The transient X-ray sky of eROSITA: from prediction, through observation, to interpretation

**Adam Malyali**

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

vorgelegt von
Adam Malyali
aus London, England

München, den 13.8.2021

Erstgutachter: Prof. Kirpal Nandra
Zweitgutachter: Prof. Achim Weiss
Tag der mündlichen Prüfung: 29.9.2021

# Contents

# List of Figures

# List of Tables

# Zusammenfassung

Die zeitaufgelöste Röntgenastronomie ermöglicht einen einzigartigen Blickwinkel für die Untersuchung einiger der extremsten, explosivsten und hochenergetischsten Phänomene des Universums. Zwischen 2019 und 2023 wird das Röntgenteleskop *Extended Röntgen Survey with an Imaging Telescope Array* (eROSITA) seine Himmelsdurchmusterung durchführen, bei der es alle sechs Monate den gesamten Röntgenhimmel beobachten wird. Aufgrund dieser regelmäßigen Wiederholung und der im Vergleich zu früheren Röntgenobservatorien verbesserten mittleren Empfindlichkeit des Sichtfeldes, ist eROSITA vielversprechend für die Entdeckung drastischer Veränderungen in der Röntgenemission einer Vielzahl von astrophysikalischen Systemen. Allerdings müssen solche variablen Objekte, auch transiente Ereignisse genannt, zunächst unter den Millionen von Röntgenpunktquellen, die eROSITA voraussichtlich aufspüren wird, identifiziert werden. Da solche transienten Ereignisse relativ kurzlebig sein können, ist es außerdem entscheidend, die interessantesten Ereignisse sofort zu identifizieren. Dadurch können zusätzliche Folgebeobachtungen durchgeführt werden während das Ereignis noch andauert.

Diese Arbeit befasst sich mit der wissenschaftlichen Nutzung der von eROSITA beobachteten variablen Quellen, wobei das Hauptaugenmerk auf der Entdeckung und Analyse von stellaren *Tidal Disruption Event* (TDE) Kandidaten während der Himmelsdurchmusterung liegt.

Kapitel 1 führt in den Stand der modernen zeitaufgelösten Röntgenastronomie ein. Außerdem wird ein Überblick über das eROSITA-Instrument gegeben, gefolgt von einem kleinen Überblick über die Anwendung des maschinellen Lernens in der zeitaufgelösten Astronomie.

Die Kapitel 2, 3 und 4 beziehen sich auf die TDE-Wissenschaft. Genauer gesagt bietet Kapitel 2 einen Überblick über stellare TDEs, beginnend mit einer kurzen theoretischen Einführung in den Zerrrissprozess und die vorhergesagten Beobachtungssignaturen von solchen Ereignissen. Dann wird eine Zusammenfassung der bisherigen Beobachtungen des Feldes gegeben, gefolgt von einer Diskussion darüber, warum diese Ereignisse so wichtig und interessant zu untersuchen sind. Kapitel 3 wurde ursprünglich in Malyali et al. (2019) veröffentlicht und untersucht die Möglichkeit, die transiente Röntgenemission von TDEs, an denen Weiße Zwerge und Schwarze Löcher mittlerer Masse beteiligt sind, mit eROSITA während der Himmelsdurchmusterung nachzuweisen. Kapitel 4 wurde in Malyali et al. (2021) veröffentlicht und präsentiert die Analyse einer Reihe von Multi-Wellenlängen-Beobachtungen des außergewöhnlichen nuklearen Transienten AT 2019avd, der während der ersten eROSITA Himmelsdurchmusterung entdeckt wurde. AT 2019avd zeigt eine Reihe von Merkmalen, die noch nie zuvor in einem anderen nuklearen Transient gesehen wurden. Wir argumentieren, dass der physikalische Mechanismus, der diese exotische Variabilität antreibt, wahrscheinlich ein Anderer ist als der in allen

anderen bisher bekannten nuklearen Transienten. Außerdem demonstrieren wir, wie außerordentlich schwierig es sein kann, zwischen TDE- und nicht-TDE-induzierten Transienten zu unterscheiden.

Kapitel 5 präsentiert detailliert die Arbeit an einer Pipeline, die für die automatische Identifizierung und Klassifizierung der transienten und variablen Quellen von eROSITA entwickelt wurde. Der Quellenklassifikator wird auf den Multi-Wellenlängen-Merkmalen (abgeleitet aus eROSITA-, Gaia- und CatWISE-Quellenkatalogen) eines Satzes bekannter Quellen unter Verwendung eines Zufallsforsts trainiert und anschließend verwendet, um wahrscheinlichkeitstheoretische Klassifikationen von Quellen vorzunehmen, die während des dritten eROSITA All-Sky Survey entdeckt wurden. Die Leistung der Quellenklassifizierung wird erheblich verbessert, wenn Multi-Wellenlängen-Merkmale verwendet werden, im Gegensatz zu Merkmalen, die nur auf den eROSITA-Quellenkatalogen basieren. Die gesamte Pipeline ist äußerst nützlich, um extreme Flussänderungsereignisse sofort zu identifizieren und dem Benutzer eine schnelle Klassifizierung des Ereignisses zu ermöglichen.

Kapitel 6 schließt mit einer Zusammenfassung dieser Arbeit und einem Ausblick auf die Transientenforschung mit eROSITA in der nahen Zukunft.

# Abstract

Time domain X-ray astronomy offers a unique vantage point for studying some of the most extreme, explosive, high-energy phenomena in the Universe. Between 2019 and 2023, the extended Röntgen Survey with an Imaging Telescope Array (eROSITA) will be performing its All-Sky Survey, where it will repeatedly observe the entire X-ray sky every 6 months. By virtue of its cadence and improved field-of-view averaged sensitivity relative to previous X-ray observatories, eROSITA holds vast promise for detecting drastic, time-variable changes in the X-ray emission from a diverse population of astrophysical systems. However, such objects must first be identified amongst the millions of X-ray point sources that eROSITA is expected to detect. Furthermore, since transient events do not last forever, it is critical to promptly identify the most interesting events, so that additional follow-up observations may be obtained whilst the transient is still ongoing.

This thesis is concerned with the scientific exploitation of the transient sky of eROSITA, with a dominant interest in detecting and analysing stellar tidal disruption event (TDE) candidates during the All-Sky Survey.

Chapter 1 introduces the state of modern time domain X-ray astronomy. In addition, an overview of the eROSITA instrument is presented, followed by a mini-review of the application of machine-learning within time-domain astronomy.

Chapters 2, 3, and 4 relate to TDE science. More specifically, Chapter 2 provides an overview of stellar TDEs, beginning with a short theoretical introduction to the disruption process and the predicted observational signatures from such events. Then, a summary of the observational status of the field is given, followed by a discussion of why these events are so important and interesting to study. Chapter 3 was initially published in Malyali et al. (2019), and explores the feasibility of detecting the transient X-ray emission from TDEs involving white dwarfs and intermediate mass black holes with eROSITA during its All-Sky Survey. Chapter 4 was published in Malyali et al. (2021), and presents an analysis of a set of multi-wavelength observations of the exceptional nuclear transient, AT 2019avd, which was discovered during the first All-Sky Survey. AT 2019avd shows a set of features that have never been seen before in the same nuclear transient, and we argue that the physical mechanism driving such exotic variability is likely different to that in all other known nuclear transients to date. Furthermore, we highlight how it can be exceptionally difficult to distinguish between TDE and non-TDE induced nuclear transients.

Chapter 5 presents a detailed report of a pipeline developed for the automatic identification and classification of eROSITA's transient and variable sources. The source classifier is trained on the multi-wavelength features (derived from eROSITA, Gaia and the CatWISE source cat-

alogues) of a set of known sources using a random forest, and is subsequently used to make probabilistic classifications of sources discovered during the third eROSITA All-Sky Survey. The source classification performance is vastly improved when multi-wavelength features are used, as opposed to only features based on the eROSITA source catalogues. The overall pipeline is extremely useful for promptly identifying extreme flux change events and providing the user with a quick classification of the event.

Chapter 6 concludes with a summary of this thesis and an outlook for transient science with eROSITA in the near future.

# Chapter 1

# Introduction

> " *Gone is the classical conception of the Universe as a serene and majestic ensemble...*"
>
> Giacconi (2003)

During its four year all sky survey, the extended Röntgen Survey with an Imaging Telescope Array (eROSITA; Merloni et al. 2012; Predehl et al. 2021) on board the Russian-German *Spectrum-Röntgen-Gamma* satellite (*SRG*; Sunyaev et al. 2021) will offer an unprecedented insight into the high-energy universe. This thesis focuses on exploiting the transient[1] sky seen by eROSITA. Reflecting the multi-disciplinary nature of modern time domain astronomy (TDA), the work herein features an interplay of simulation of observations, applied machine learning, analysis of astronomical observations, and astrophysical interpretation.

In this introduction, I will discuss:

- A brief history of X-ray astronomy and the discovery of the dynamic X-ray sky, followed by an overview of the rich diversity of X-ray transient phenomena known today (section 1.1).
- An overview of the eROSITA instrument. This will begin with the science drivers for building such an instrument, and how these shaped its final design. Then, its operations and survey strategy will be discussed, followed by the expected insight that eROSITA will provide into the dynamic X-ray sky (section 1.2).
- The launch of the eROSITA mission in the context of the recent transformation of TDA over the last 30 years (section 1.3).

Finally, I will conclude with an outline of this thesis in section 1.4.

[1] For the remainder of this thesis, I will broadly consider a transient to be either a one-off destructive event (e.g. the tidal destruction of a star as it torn apart in the vicinity of a super massive black hole, or supernova), or to be an eruptive variable (e.g. a large amplitude outburst from an X-ray binary).

## 1.1 The dynamic X-ray sky

### 1.1.1 Early insights into the high-energy universe

The serendipitous discovery of the brightest persistent X-ray source in the sky, Sco X-1 (Giacconi et al., 1962), in 1962, during a mission that was searching for X-rays from the surface of the Moon, marked the first detection of X-rays from a non-solar origin, and heralded in an era of major discoveries that transformed our view of high-energy astrophysical phenomena. Whilst there were several candidates for extraterrestial X-ray sources prior to this (e.g. flaring stars, novae and supernova remnants), such X-ray sources had been expected to be at much fainter X-ray fluxes than Sco X-1, if assuming Galactic-like distances, and therefore too faint to be detected with existing instrument sensitivities at the time. After tighter constraints were later obtained on the X-ray position of Sco X-1 (Gursky et al., 1966), which subsequently aided identifying the optical counterpart as a $13^{\text{th}}$ magnitude star (Sandage et al., 1966), it was established that Sco X-1 exhibited an extremely large X-ray-to-optical luminosity ratio, relative to other known stellar objects. The source thus radiated most of its energy through high energy photons, and it was clear that a novel (previously unanticipated) physical mechanism must be producing such radiation. Only one year after the optical counterpart identification, Shklovsky (1967) proposed that the X-rays in this system were being produced via a neutron star accreting matter from its companion star in a close binary system.

Additional sounding rocket flights and balloon-based missions in the years after the discovery of Sco X-1 began to further unveil the diverse nature of the X-ray sky. Not only were new compact binaries discovered within the Milky Way (e.g. Cyg X-1, Bowyer et al. 1965), but the existence of extra-galactic X-ray sources was also established through the detections of X-ray emission from the galaxy M87 and the quasar 3C 273 (Friedman & Byram, 1967), and the discovery of extended X-ray emission from the extremely hot ($\sim 10^8$ K) gas in the inter-galactic medium of the Coma galaxy cluster (Boldt et al., 1966). Through these repeated observations of the small set of known sources at the time, it was also found that a subset of these X-ray point sources showed significant variations in their X-ray fluxes between visits (e.g. Cyg X-1, Byram et al. 1966). In 1967, a new X-ray source, Cen X-2, was detected that appeared roughly as bright as Sco X-1, but had not been detected in observations of the same region only four months previously (Harries et al., 1967). Follow-up X-ray observations found Cen X-2 to decay several orders of magnitude over the following months (Chodil et al., 1968). Described initially as an 'X-ray nova', Cen X-2 was the first X-ray transient ever discovered.

Within under a decade, X-ray observations had completely transformed our perceptions of what constitutes the high energy Universe. Whilst many additional important early results were obtained with X-ray detectors flying aboard rockets and balloons, X-ray astronomy only grew into a format more similar to today after the launch of space based satellites, starting with Uhuru in 1970 (Giacconi et al., 1971; Jagoda et al., 1972). Whilst such space based missions were necessary to avoid the absorption of X-rays by the Earth's atmosphere, they also enabled observations to be performed with longer exposure times, and offered improved pointing stability relative to rockets and balloons. Crucially for TDA, this allowed for easier repeated observations of the same source, and for more robust detections and characterisation of X-ray variabil-

ity (e.g. Tananbaum et al. 1972a,b).

### 1.1.2 Today's view of the X-ray transient and variable sky

Since Uhuru, numerous X-ray missions have contributed towards our current understanding of the X-ray dynamic sky (e.g. *Ariel 5*, Smith & Courtier 1976; *The European X-ray Observatory Satellite, EXOSAT*, Taylor et al. 1981; *Ginga*, Tsunemi et al. 1989; *ROSAT*, Trümper 1982; *ASCA*, Tanaka & Inoue 1994; the *Rossi X-ray Timing Explorer*, *RXTE*, Gruber et al. 1996; *BeppoSAX*, Boella et al. 1997; *Chandra*, Weisskopf et al. 2000; *XMM-Newton*, Jansen et al. 2001; *INTEGRAL*, Winkler et al. 2003; *Swift*, Gehrels et al. 2004; *HETE-2*, Ricker et al. 2003; *MAXI*, Matsuoka et al. 2009; *NuSTAR*, Harrison et al. 2013; *NICER*, Gendreau et al. 2016). We now know that a rich variety of transient classes exist, which collectively span ∼20 orders of magnitude in luminosity, and vary on timescales from milliseconds to years (Fig. 1.1).

Stepping through the X-ray luminosity-timescale phase space plot in Fig. 1.1 in order of increasing X-ray luminosity, the following section briefly discusses the physical and observational characteristics of the main X-ray transient and variable sources currently known. We omit discussion here of some of the more exotic, less well-studied X-ray transients, such as those produced via supernovae shock breakout (e.g. Soderberg et al. 2006), accretion induced white-dwarf collapse events (Metzger et al., 2008; Yu et al., 2019), or soft-gamma repeaters (Kouveliotou et al., 1998).

#### Stellar flares

Magnetic reconnection events in the corona of cooler stars (spectral types F to M) produce X-ray outbursts (Fig. 1.2), originating from the transient heating and subsequent cooling of coronal plasma during such events (Güdel, 2004). Based on the light curves of variable F to M stars seen with *XMM*, Pye et al. (2015) characterise the variability from these systems as either: i) stochastic (right panel of Fig. 1.2), or ii) flaring (left panel of Fig. 1.2), where the X-ray flux rises rapidly to peak luminosities of $L_{\rm X} \sim 10^{29} - 10^{32}$ erg s$^{-1}$ and decays back to a quiescent level with $L_{\rm X} \sim 10^{27} - 10^{31}$ erg s$^{-1}$, over timescales between $\sim 10^2$ s and a few $10^4$ s. For flaring systems, the flare rise timescale is proportional to the flare decay timescale, and the peak luminosity is proportional to their quiescent luminosity (Pye et al., 2015).

*Ultra-cool dwarfs* (UCDs), typically of spectral type later than M7, also show outbursts similar to coronal flares, although these have generally been observed at lower peak luminosities ($L_{\rm X} \lesssim 10^{30}$ erg s$^{-1}$, Gupta et al. 2011; De Luca et al. 2020). Furthermore, the origin of the X-ray flaring in UCDs is currently unclear, since these systems are not expected to have hot coronas that produce flaring (due to their lower temperatures producing neutral atmospheres with negligible magnetic activity).

In addition, *young stellar objects* (YSOs), with peak $L_{\rm X}$ between $10^{28}$ and $10^{33}$ erg s$^{-1}$ in the 0.5-8 keV band (e.g. Feigelson et al. 2002; Ozawa et al. 2005), are also known to be X-ray variables. However, they show a more diverse range of X-ray variability behaviours relative to stellar coronal flares, with this added complexity potentially originating from the interactions of the magnetic fields of the YSO and its circumstellar disc (Montmerle et al., 2000).

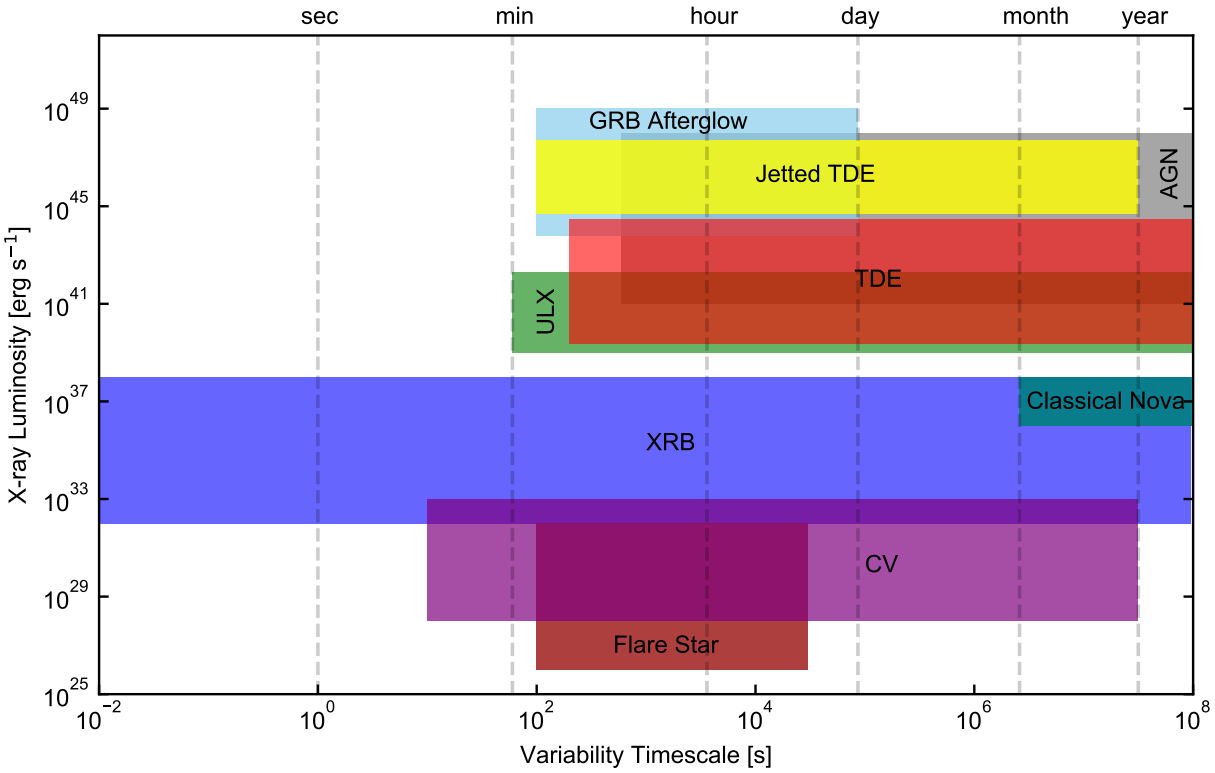


Figure 1.1: An overview of the X-ray transient and variable phase space, adapted from Soderberg (2010) and Merloni et al. (2012). The variability timescale provides a rough estimate of the outburst duration for a given transient class, with some source classes showing an extremely broad range of variability timescales. The X-ray luminosities have been obtained from as close to the 0.2-10 keV band as possible, depending on the quoted luminosities (and timescales) in the following literature: gamma-ray burst (GRB) afterglows (e.g. D'Avanzo et al. 2012), active galactic nuclei (AGN; e.g. Boller et al. 2021), jetted tidal disruption events (jetted TDE; Burrows et al. 2011; Levan et al. 2011; Cenko et al. 2012; Pasham et al. 2015), tidal disruption events (TDEs; e.g. Saxton et al. 2020) ultra-luminous X-ray sources (ULX; Feng & Kaaret 2007; Irwin et al. 2016; Israel et al. 2017), classical nova (e.g. Henze et al. 2009), X-ray binaries (XRB; Done et al. 2007; Kotze & Charles 2012; Romano et al. 2015; Sidoli et al. 2019), cataclysmic variables (CV; e.g. Reis et al. 2013; Schwope et al. 2014) and flare stars (Pye et al., 2015).

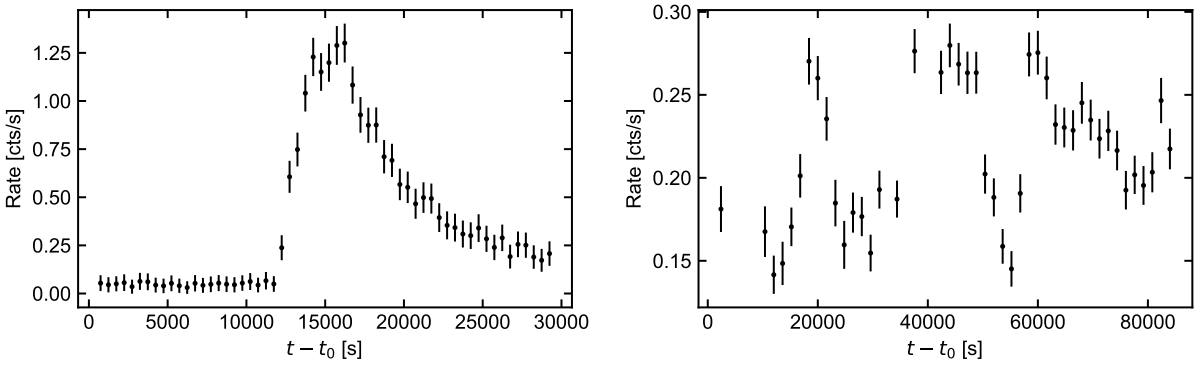


Figure 1.2: *XMM* light curves portraying stellar X-ray variability in the 0.2-12 keV band, initially reported in Pye et al. (2015). *Left*: EPIC-MOS1 light curve of a coronal flare from 2MASS J23163068+7905362 (500s binning). *Right*: stellar stochastic variability seen in the EPIC-PN light curve of 2XMM J040721.7-121003 (1600s binning). $t_0$ is set to the start time of each *XMM* observation.

### Cataclysmic Variables

*Cataclysmic variables* (CVs) are semi-detached binary systems consisting of a white-dwarf (WD) accreting matter from a secondary star, with the nature of the accretion process strongly dependent on the WD's magnetic field strength. For non-magnetic systems (with magnetic field strength $B \lesssim 0.1$G), material is accreted onto the WD via an accretion disc, whereas for WDs with $B \gtrsim 10^4$G (*polars*), accretion proceeds via ionised material from the donor star being channeled along the magnetic field lines of the WD and onto its magnetic poles, with the strong magnetic fields preventing formation of an accretion disc around the WD. There also exists *intermediate polars* (IPs, with $0.1\,\mathrm{G} \lesssim B \lesssim 10^4\,\mathrm{G}$; e.g. Ramsay et al. 2008), whereby an accretion disc does form around the WD, but its inner radii are truncated and material is accreted onto the magnetic poles (similarly to polars). The X-ray variability of CVs is diverse and is highly dependent on the magnetic field strength of the WD. For the non-magnetic CVs, persistent X-ray emission with $L_\mathrm{X} \sim 10^{29} - 10^{32}$ erg s$^{-1}$ is produced from the boundary layer between the WD's surface and the inner accretion disc, with the disc emission predominantly in the optical and UV bands.

On the basis of their optical light curve behaviour, non-magnetic CVs can be subdivided into being *nova*-like and *dwarf-nova*-like systems. The former undergo long periods where the optical flux is roughly constant, and then undergo large amplitude outbursts (6-16 mag increase) over month long timescales. *Classical novae* are nova-like systems where such a major outburst has only been observed to once, whereas *recurrent novae* show classical nova-like outbursts more than once. These major optical outbursts are triggered by the thermonuclear burning of hydrogen-rich material that has accumulated onto the surface of the WD. This subsequently heats the surface of the WD, with the system brightening in the X-rays and rising to peak luminosities of $L_\mathrm{X} \sim 10^{36} - 10^{38}$ erg s$^{-1}$, at least a month after the optical outburst and lasting over month-

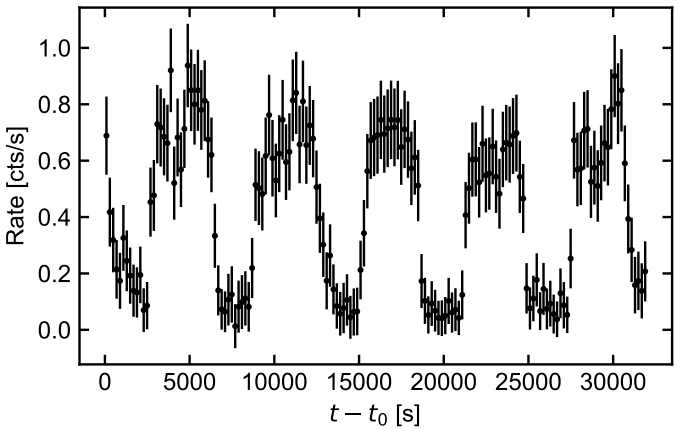


Figure 1.3: *XMM* EPIC-PN light curve in the 0.2-12 keV band of the magnetic CV Swift J0706.8+0325 with 200s binning. The X-ray emission from this system is modulated with a period of ∼1.7 hours (Bernardini et al., 2017), with the spin period of the WD equal to the orbital period of the binary. $t_0$ is set to the start time of the *XMM* observation.

to-year long timescales[2] (Starrfield et al. 1990). After brightening in the X-rays, the CV enters a super-soft X-ray state (SSS), where the X-ray spectrum can be well modelled by a blackbody with temperature $15 \lesssim kT \lesssim 80$ eV (Kahabka & van den Heuvel, 1997).

Dwarf nova systems show recurrent ∼2-6 mag outbursts in their optical light curves, with this variability thought to be caused by the accretion disc instability (Meyer & Meyer-Hofmeister, 1981). The time between successive outbursts is observed to be proportional to the energy of the earlier outburst, and the X-ray variability of dwarf nova systems over the course of these cycles is complex and currently not well understood (e.g. Wheatley et al. 2003; Zhang et al. 2017).

Most polars and IPs are hard X-ray emitters (Mukai, 2017), with the bulk of this being produced from Brehmsstrahlung radiation released in strong shocks at the base of the accretion column onto the WD's poles ($L_X \lesssim 10^{33}$erg s$^{-1}$; Aizu 1973). A subset of these systems also show a soft X-ray excess thought to originate from a hot spot near the base of the accretion column with $kT \sim 10 - 100$ eV (e.g. Traulsen et al. 2010), which has been heated through absorption of the Brehmsstrahlung emission (Cropper, 1990). Both the rotation of the WD and the orbital motion of the binary can produce occultations of the X-ray bright hot spot, thus producing periodic modulations (Fig. 1.3) of the X-ray light curves of magnetic CVs (which also typically show stochastic flickering in the X-rays due the erratic accretion onto the WD; e.g. Anzolin et al. 2010).

### X-ray binaries

*X-ray binaries* (XRBs) consist of a compact object (neutron star, NS; or black hole, BH) accreting matter from a donor star. Systems where the donor mass is $\lesssim 1 M_\odot$ are labelled *low-mass X-ray*

[2]However, this phase can turn-off within ∼100days and the lifetime of the SSS phase is quite variable.

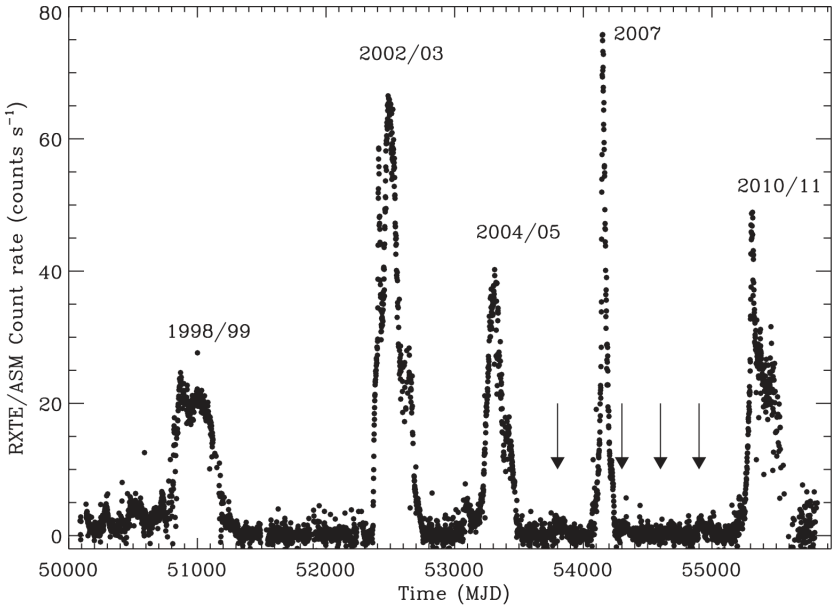


Figure 1.4: *RXTE* All Sky Monitor light curve of the erratic, irregular outbursts from the low mass black hole XRB GX 339-4 in the 1.5-12 keV energy range. The black arrows mark on epochs of smaller amplitude outbursts of the system which were below the detection sensitivity of the All Sky Monitor. Figure from Corbel et al. (2013).

*binaries* (LMXBs; e.g. Liu et al. 2001), whereas *high-mass X-ray binaries* (HMXBs) are systems where the donor mass is $\gtrsim 10M_{\odot}$. The X-ray variability of the binary can vary significantly depending on the properties of the compact object and the donor star.

The vast majority of known *black hole X-ray binaries* (BHBs) are LMXBs, with the donor star feeding the black hole's accretion disc via Roche lobe overflow. These systems typically spend long periods in quiescence with $L_{\rm X} < 10^{33}$ erg s$^{-1}$, followed by days to months-long outbursts that reach peak X-ray luminosities up to $10^{36} - 10^{38}$ erg s$^{-1}$ (Fig. 1.4). These outbursts recur irregularly over month-to-year long timescales, are thought to be triggered by accretion disc instabilities (e.g. Lasota 2001), and the general evolution of each outburst in hardness-luminosity space is the same across all low mass BHBs (e.g. Fender et al. 2004).

NS LMXBs are predominantly persistent X-ray emitters with $L_{\rm X} > 10^{36}$ erg s$^{-1}$, although a small number of systems do show outburst behaviour similar to BH LMXBs (Done et al., 2007). A key observational signature for the presence of a NS in an XRB are Type I X-ray bursts. These are characterised by fast, order of magnitude rises (∼second-long timescales) in the X-ray luminosity of the system, followed by a slowed decay over ∼ 10s to several minutes-long timescales. These bursts stem from the unstable thermonuclear burning of material accreted onto the surface of the NS (such burning is not possible for BHs since they have no hard surface to accumulate material), and show a range of recurrence timescales (from ∼ 10s to several hours).

### Ultra-luminous X-ray sources

*Ultra-luminous X-ray sources* (ULXs) are point-like, off-nuclear sources in their host galaxies, with luminosities exceeding $10^{39}\,\mathrm{erg\,s^{-1}}$, the Eddington luminosity of a $10M_{\odot}$ BH. Such systems are currently thought to be either intermediate mass BH (IMBH) candidates accreting at sub-Eddington rates (e.g. Pasham et al. 2014), or stellar mass ($\lesssim 20M_{\odot}$) compact objects with relativistically beamed emission (King et al., 2001) or undergoing super-Eddington accretion (Begelman, 2002), such as the NS in the ULX of NGC 5907, which is estimated to be accreting at $\sim$ 1000 times Eddington with $L_X$ ~a few $10^{41}\,\mathrm{erg\,s^{-1}}$ (Israel et al., 2017).

It is likely that the current population of known ULXs is actually a heterogeneous mix of super-Eddington NSs, sub-Eddington IMBHs, and background AGN that (falsely) appear to be associated with an extended foreground galaxy, thus there is yet a well defined variability characteristic of ULXs. Nonetheless, several systems show pulsations in their X-ray light curves, signposting the presence of a rotating NS (e.g. Israel et al. 2017; Carpano et al. 2018), whereas others show irregular, stochastic variability (Earnshaw et al., 2016), or even fast outbursts that rise ~2 orders of magnitude over timescales of a minute, and decay over the following hour post-peak (Irwin et al., 2016).

### Active galactic nuclei

*Active galactic nuclei* (AGN) are accreting supermassive black holes (SMBHs) of masses $10^6$ – $10^9 M_{\odot}$ in the centres of massive galaxies (Lynden-Bell, 1969), emitting with bolometric luminosities $10^{42}$ – $10^{48}\,\mathrm{erg\,s^{-1}}$. The extremely luminous output from these systems is produced by an optically thick, geometrically thin, radiatively efficient accretion disc surrounding the SMBH (Shakura & Sunyaev, 1973). Whilst the emission from this disc peaks in the rest-frame UV (e.g. Shields 1978), non-jetted AGN are typically X-ray bright due to the inverse Compton scattering of seed photons from the accretion disc by the 'corona' – a hot ($\sim 10^9$K), compact, optically thin plasma (e.g. Haardt & Maraschi 1991, 1993), lying a few gravitational radii above the disc (e.g. Reis et al. 2013).

Whilst AGN exhibit variability across all wavelengths, the most rapid variations are seen in the X-ray light curves, which display variability from timescales of minutes to years. X-ray variability is often studied and characterised based on the light curve's power spectral density (PSD). The X-ray PSD can be modelled by a broken power law, with slope $\sim -1$ at low frequencies, and $\lesssim -2$ for frequencies higher than the break frequency (e.g. Edelson & Nandra 1999). The characteristic timescale for the AGN's variability is then obtained from the inverse of the break frequency, and has been found to be inversely proportional to the black hole mass (McHardy et al., 2006). The X-ray variability information can also be used to constrain the physical conditions within the vicinity of the accreting BH (e.g. via X-ray reverberation mapping, see review by Uttley et al. 2014).

There is also an increasing number of AGN being identified which show significant changes to their observed characteristics over time. This generally constitutes large amplitude flux variations in excess of the stochastic variance, or major X-ray and/ or optical spectral changes (Fig. 1.5). These may be induced by intrinsic or extrinsic factors to the accreting system. The

former are largely thought to be produced by extreme changes to the accretion rate, potentially through state transitions in the inner disc (Noda & Done, 2018; Ross et al., 2018), radiation pressure instabilities in the disc (Śniegowska & Czerny, 2019) or tidal disruption events (further details presented in Chapter 2 of this thesis). Extrinsic factors that may significantly alter the appearance of an AGN are obscuration events (e.g. Matt et al. 2003), nuclear[3] supernovae (e.g. Graham et al. 2017), or microlensing events of background AGN by stars in a foreground galaxy (Meusinger et al., 2010; Lawrence et al., 2012; Graham et al., 2017).

#### Gamma-ray Burst Afterglows

*Gamma-ray bursts* (GRBs) are the most luminous explosions in the Universe (peak isotropic luminosities $\sim 10^{51} - 10^{52}\,\mathrm{erg\,s^{-1}}$). Based on the duration of their outburst, GRBs are divided into short ($\lesssim 2$ s) and long ($\gtrsim 2$ s) variants (Kouveliotou et al., 1993), with the former being produced by the merging of a NS-NS binary, and the latter from the collapse of a massive star into a NS or BH (Woosley, 1993). The 'prompt' emission in a GRB is thought to be produced through internal dissipation within a relativistic jet launched by the stellar collapse or merger (and can also produce a short-lived X-ray flash in addition to the gamma-ray emission e.g. Heise et al. 2001), whilst the afterglow originates from shocks produced as the ejecta collide with the circum-burst medium. The X-ray afterglows of GRBs (Fig. 1.6) are X-ray transients that show a $t^{-1.2}$ power-law decline (Nysewander et al., 2009) in their X-ray emission from peak luminosities of $\sim 10^{47} - 10^{49}\mathrm{erg\,s^{-1}}$ (D'Avanzo et al., 2012), and can decay by factors of 1000-10000 over timescales of ∼10 days (e.g. Gendre et al. 2008).

## 1.2 An overview of eROSITA

eROSITA is the primary instrument on board *SRG*, and was developed under the leadership of the Max Planck Institute for Extraterrestial Physics in Garching, Germany. In the following section, the science drivers for eROSITA are discussed, and a brief overview of its main features is given. Then, the operation of the telescope during its all sky survey phase will be explained, followed by the expected impact that eROSITA will have on X-ray TDA.

### 1.2.1 Main scientific motivations

The primary science driver for eROSITA is to detect a large ($\sim 10^5$) sample of galaxy clusters (Merloni et al., 2012). Beginning as overdensities in the primordial Universe (e.g. Bardeen et al. 1986), these systems have hierarchically grown over time to become the most massive, gravitationally-bound collapsed structures in the Universe. Such a large sample of galaxy clusters will firstly enable the large scale structure (LSS) of the Universe to be mapped out, since these systems are expected to form at the nodes of the cosmic web (e.g. Springel et al. 2005). Secondly, through the usage of various scaling relations, the observed X-ray luminosity of these systems can be used as a proxy for the cluster mass (e.g. Reiprich & Bohringer 2002), and the cluster

[3]Nuclear in this case refers to supernovae that are astrometrically consistent with the nucleus of their host galaxy.

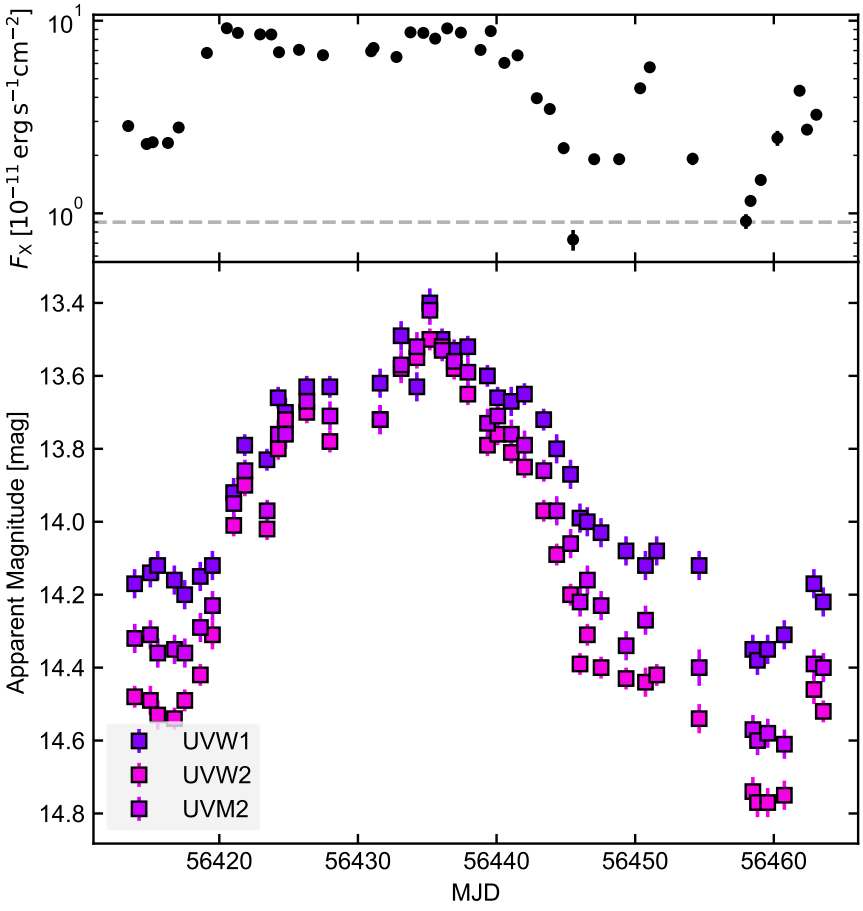


Figure 1.5: An example changing-look event seen in NGC 2617 (Shappee et al., 2014). After detection of an optical outburst from the system, a subsequent monitoring campaigned observed an increase in the X-ray flux by an order of magnitude. The black markers are inferred 0.3-10 keV fluxes, whilst the grey dashed line denotes the 0.2-12 keV flux inferred from an archival *XMM* observation six years before the optical outburst. This X-ray increase is followed, with a time lag of $\sim 2-3$ days, by a large increase in the UV flux. The system also underwent an extreme optical spectral change relative to archival observations, where broad Balmer emission lines appeared in spectra taken after the initial optical outburst. The 'changing-look' classification thus stems from the system transitioning from a Type 1.8 (no broad Balmer emission lines) to Type 1 AGN (with broad Balmer emission lines). Figure adapted from data presented in Shappee et al. (2014).

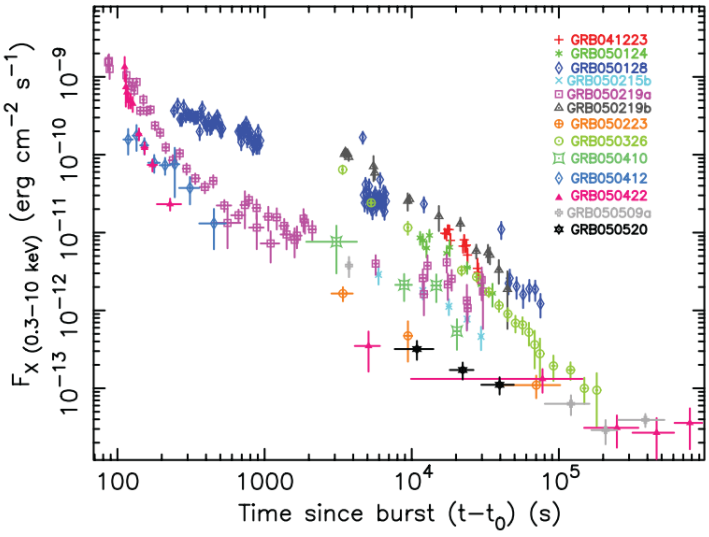


Figure 1.6: Observed frame 0.3-10 keV X-ray light curves of a population of GRB afterglows followed up using the *Swift* observatory. Plot from Nousek et al. (2006).

mass function may be inferred (i.e. the number of clusters of a given mass at a given redshift; e.g. Bahcall & Cen 1993). Since this function is highly sensitive to the expansion history of the Universe, and thus its underlying cosmology, then X-ray surveys of galaxy clusters may be used to provide constraints on cosmological parameters. For such a survey to be competitive with other cosmological probes (e.g. Ia supernovae, Schmidt et al. 1998; Perlmutter et al. 1999; the cosmic microwave background, e.g. Planck Collaboration XIII 2016; gravitational lensing, Heymans et al. 2013), Haiman et al. (2005) estimated that $\sim 10^5$ clusters would need to be detected. This result thus led to the requirement that eROSITA must have high sensitivity towards detecting soft X-rays, so that it could efficiently detect X-ray emission from the hot intra-cluster medium. Furthermore, eROSITA needed to be an imaging instrument, and have a large field-of-view (FoV) with well-behaved vignetting and point spread functions (PSF) against off-axis angle (to attain higher angular resolution[4] and a large FoV-averaged effective area), such that it would be able to rapidly survey large sky areas that were necessary for detecting the large cluster sample.

A secondary motivation was to produce the first all sky imaging survey in the hard X-rays, in order to obtain a census of the AGN population that includes the most obscured sources. Whilst *ROSAT* was excellent for discovering a huge population of AGN during its all sky survey in the 1990s, it was only sensitive in the 0.1-2.4 keV energy range, and so its AGN population was generally biased towards systems with low obscuration; eROSITA being sensitive to the detection of harder X-rays would thus partially alleviate such biases. eROSITA is expected to detect approximately three million X-ray selected AGN during its first four years of operation, expanding upon the population known pre-eROSITA by roughly an order of magnitude. Combined with extensive follow-up campaigns (particularly optical spectroscopic), eROSITA will enable the construction

[4]This is needed for distinguishing point sources from extended sources.

of rich multi-wavelength datasets that will be invaluable for studying accreting SMBHs across the mass scale and over different phases of their evolution.

### 1.2.2 Instrumentation

eROSITA consists of 7 very similar telescope modules (TM), with each module comprising a mirror assembly (MA, containing an X-ray baffle, a mirror module and electron deflectors), and a camera assembly (CA, containing a filter wheel, a camera and the camera electronics). The mirror module in each MA is made from a set of 54 co-axially nested mirrors with a gold-coated reflecting surface. Each mirror is based on a Wolter-I design (Wolter, 1952) where it uses grazing incidence reflection to focus X-rays onto a focal plane 1.6m from the principle plane. An X-ray baffle, located at the front of each MA, reduces the amount of stray light[5] passing through the MA by ~92% (Friedrich et al., 2014; Predehl et al., 2021), whilst the magnetic electron deflectors at the rear of the MA greatly reduce the number of low energy electrons striking the CA. Each TM has the same field of view (~1 degree diameter), such that there is effectively a 7-fold redundancy across the 7 TMs. The PSF of eROSITA has a half-energy width (HEW, the angular diameter within which 50% of a point source's photons are detected) of ~ 16" on-axis at 1.5 keV (averaged over the 7 TMs), with this increasing for higher off-axis angles; the HEW averaged over the FoV is ~ 26" (Predehl et al., 2021).

A pn-CCD lies in the focal plane of each TM. Each CCD is composed of 384x384 pixels, with each pixel capturing a 9.6x9.6 arcsec sky area. eROSITA's CCDs offer an improved energy resolution (FWHM of ~80 eV at 1.5 keV; Meidinger et al. 2020) over those aboard *Chandra* (~95 eV at 1.5 keV[6]) and *XMM* (~110 eV at 1.5 keV; Strüder et al. 2001)[7]. Unlike several other X-ray telescopes, there are no chip gaps on the CCDs, leading to an uninterrupted imaging across the entire FoV. Whilst the design of the pn-CCDs is very similar to those in the EPIC-PN camera of *XMM*, each CCD in eROSITA additionally features a shielded framestore area separate from the CCD. After each (nominal) 50ms integration, the charge in the image area of the CCD is then shifted to the framestore area within ~ 0.12 ms. Readout is then performed from this framestore area within ~ 9.2 ms, with this system greatly reducing the number of photons from striking the CCDs during readout and thus the number of out-of-time events.

The filter wheel, located in front of the camera in each CA, can be rotated to be in either OPEN, CALIBRATION, CLOSED or FILTER positions. The OPEN position was used primarily at the start of the mission for outgassing of the instrument, where this process effectively tries to remove any contaminant gasses present within the telescope that might be detrimental to telescope operation (such as if they were to condense on the CCDs). The CALIBRATION position contains a source of radioactive $^{55}$Fe, which irradiates a target plate and produces a set of fluorescence lines. The measured CCD response to this emission is then used for further instrumental calibration post-launch (further details presented in Dennerl et al. 2020). The CLOSED position

[5]The stray light dominantly originates from photons outside the FoV that would reach the CCDs via a single reflection.

[6]`https://cxc.harvard.edu/cal/Acis/`.

[7]Although higher spectral resolutions may be obtained through usage of the grating spectroscopy instruments onboard *Chandra* and *XMM*.

shields each camera from non-instrumental X-rays, and is thus used primarily for instrument protection and for when measuring the instrumental background. The FILTER position is the default mode for eROSITA observations. Five out of the seven CCDs have a 200nm Al filter mounted on-chip, with a 200nm polyimide foil in their FILTER mode of their filter wheel, whilst two CCDs (TM5 and TM7) have no on-chip filter, but a 100nm Al filter on a 200nm polyimide foil on their filter wheel's filter instead. Whilst these filters reduce eROSITA's sensitivity to detecting X-rays below 1 keV, they are present primarily to block out the incidence of optical-UV light on the CCDs, and thus reduce the number of spurious events caused by optical loading.

The on-axis effective area curve[8] of eROSITA is compared with other missions in the left panel of Fig. 1.7. For on-axis observations, eROSITA is roughly as sensitive as *XMM* (using all three of its telescope modules, i.e. PN and 2 MOS) for detecting photons in the 0.5-2 keV energy band, whereas *XMM* is more sensitive outside of this range. Over the 0.2-10 keV range, eROSITA also has a larger effective area than both *Chandra* ACIS-I and *Chandra* HRC-I.

The most significant enhancement of eROSITA relative to previous X-ray missions is clearest when comparing its *grasp* (the product of the size of its FoV with the FoV-averaged effective area, as a function of energy) with its predecessors, where the the grasp of an instrument quantifies how fast an instrument can survey the sky. In the right-hand panel of Fig.1.7, the grasp is compared with *Chandra*, *XMM* and *Chandra*; between ∼0.3keV and 2.3keV, eROSITA has the highest grasp of all major imaging X-ray missions (eROSITA's grasp is ∼5x greater than *XMM*'s at 1 keV), resulting in it being able to image larger sky regions to a given depth relative to its predecessors, within a given period of time.

### 1.2.3 Operations

On 13th July 2019, *SRG* was launched successfully from the Baikonur Cosmodrome in Kazakhstan, and began a ∼100 day journey towards the second Lagrangian (L2) point of the Earth-Sun system, ∼1.5 million km away from Earth. During this journey, the eROSITA instrument initially underwent a commissioning phase for each of its 7 cameras, followed by a series of performance verification and calibration studies beginning from the 18th October 2019 (Dennerl et al., 2020; Predehl et al., 2021). After a number of trajectory corrections, *SRG* entered a halo orbit about L2 on the 21st October 2019, where it will complete one full revolution every 180 days. eROSITA, and the Mikhail Pavlinski ART-XC telescope (herein PART-XC) alongside it on-board *SRG*, are the first X-ray telescopes to operate at L2, with previous X-ray satellite missions being based at much lower altitudes in either circular (e.g. *Chandra*) or elliptical orbits (e.g. *XMM* and *Chandra*) about Earth. The main benefits of an L2 orbit for *SRG* are that near-continuous observations are possible, and that it was expected to be (and is) less prone to soft proton flaring events that affect X-ray telescopes on trajectories that pass through Earth's magnetosphere (such as *XMM* and *Chandra*).

eROSITA's first four years of operation are being spent in its all sky survey phase (eRASS)[9], where it will scan the entire sky once every 6 months, and repeat this eight times (eRASS1-8).

[8] Effectively how sensitive eROSITA is to the detection of photons as a function of their energy.

[9] The all sky survey will be followed by a three year pointed observation phase.

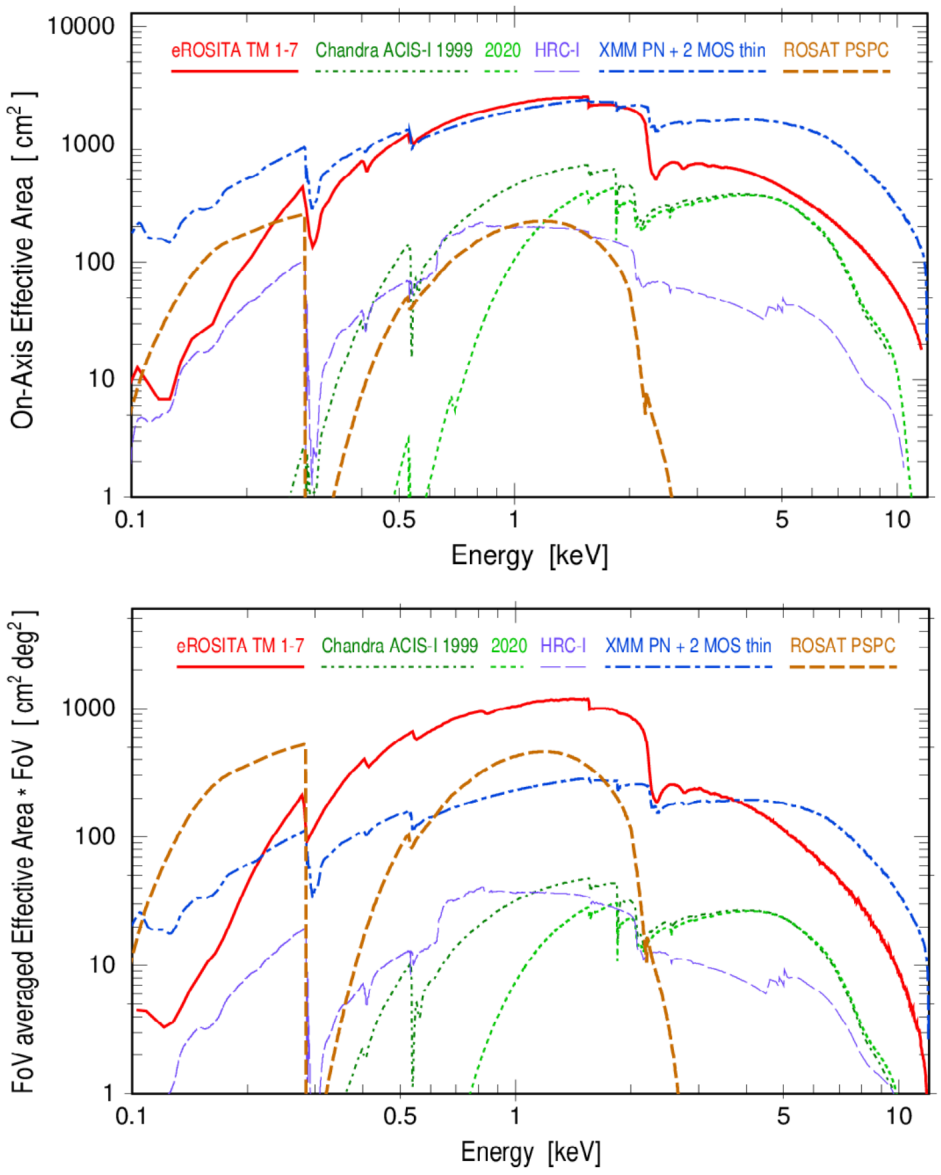


Figure 1.7: *Top panel*: A comparison of the summed on-axis effective area against energy of eROSITA's 7 TMs with other major X-ray telescopes; on-axis pointed observations with eROSITA are competitive with *XMM*'s PN and 2 MOS thin cameras in the 0.3 keV to 2.3 keV range. *Bottom panel*: A comparison of the *grasp* (FoV multiplied by the FoV averaged effective area) versus energy of eROSITA with other major X-ray instruments; eROSITA offers the fastest survey speeds of all current X-ray telescopes in the 0.3 keV to 2.3 keV range. Figures from Predehl et al. (2021).

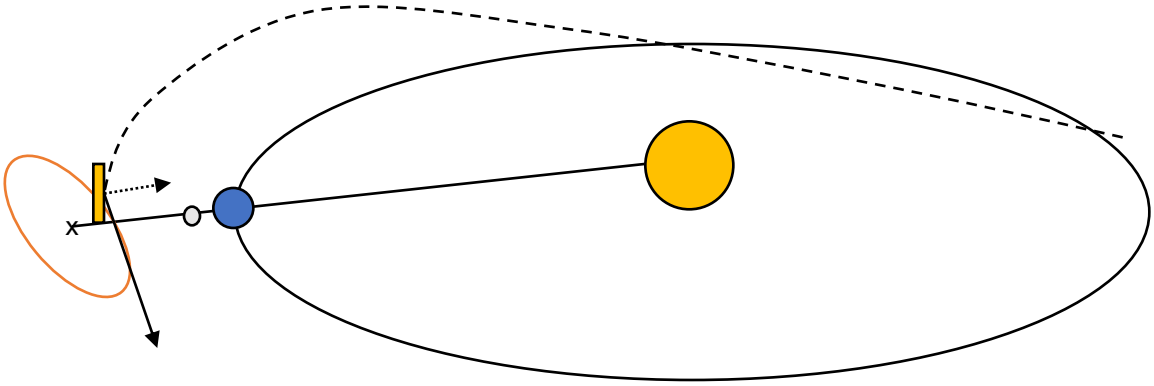

Figure 1.8: eROSITA's orbit during its all sky survey. eROSITA (yellow rectangle) completes a halo orbit (orange ellipse) about L2 (black cross) of the Earth (blue circle)-Sun (yellow circle) system, once every 180 days. eROSITA rotates every 4 hours about an axis pointed in the direction of the Sun (black dotted arrow pointing towards Sun), with eROSITA's pointing direction perpendicular to this rotation axis (solid black arrow). The black dashed line traces out eROSITA's past trajectory during the survey. Diagram constructed from the following DLR video of eROSITA's operations (`https://vimeo.com/342935738`).

Ground contact is made with eROSITA once-per-day for the purposes of transmitting its telemetry for further processing on Earth, and also for eROSITA/ spacecraft commands to be sent, such as those needed for occasional orbit correction manoeuvres. In its eRASS phase, eROSITA rotates every 4 hours about an axis pointing a few degrees away from the Earth-Sun axis. Since eROSITA corotates with the Earth around the Sun, the plane of the sky that eROSITA scans rotates ∼1 degree each day, thus enabling an all sky scan every 6 months. eROSITA's unique cadence results in a smooth but non-uniform exposure map of the whole sky being produced (Fig. 1.9), with the lowest exposures being in the ecliptic plane, and the highest at the ecliptic poles. For the lowest exposure regions, each source will be observed roughly ∼6 times in one day[10], with each visit roughly 40s long spaced 4 hours apart, and this is repeated every half a year, for four years. The number of visits that a source receives, and the number of consecutive days that the source is observed for in a given eRASS, increases towards the ecliptic poles (Fig. 1.2.3).

### 1.2.4 eROSITA's impact on time-domain astrophysics

eROSITA will detect several million X-ray point sources during its four year all sky survey (Merloni et al., 2012). Whilst the vast majority of these will effectively be observed at a constant flux during this time, a subset of these will be variable and/ or transient in nature. The scanning pattern of eROSITA enables a probe of variability across a broad range of timescales- from within

[10]This assumes a constant scanning rate during the All-Sky Survey. During operations, the actual scanning rate is slightly non-uniform, meaning that in some sky regions the scanning rate is faster than ∼1 degree each day and some sources in equatorial fields may have fewer than 6 visits in a day, as shown in Fig. 1.2.3 (J. Robrade, priv. comm.)

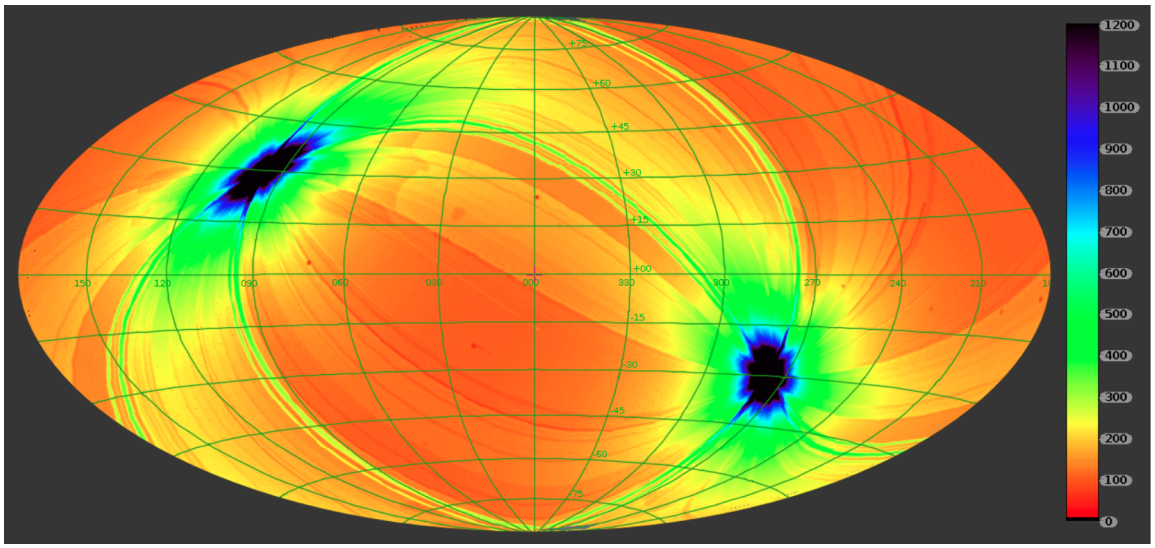


Figure 1.9: Effective exposure map in the 0.6 – 2.3 keV band (corrected for vignetting) from eRASS1 observations. The effective exposure is computed through multiplying the exposure time for a sky position with the ratio of the eRASS1-averaged effective area at that sky position, to the on-axis effective area. The values next to the colour bar correspond to the effective exposure times. eROSITA's cadence results in the ecliptic poles being visited more frequently than regions closer to the ecliptic equator, such that a highly non-uniform exposure map is created over the whole sky. Within $\sim$ 5 degrees of the two ecliptic poles, there is a very steep gradient in the exposure times, with this peaking at about $\sim$10ks. The map is plotted in galactic coordinates, using an Aitoff projection. Figure from Predehl et al. (2021).

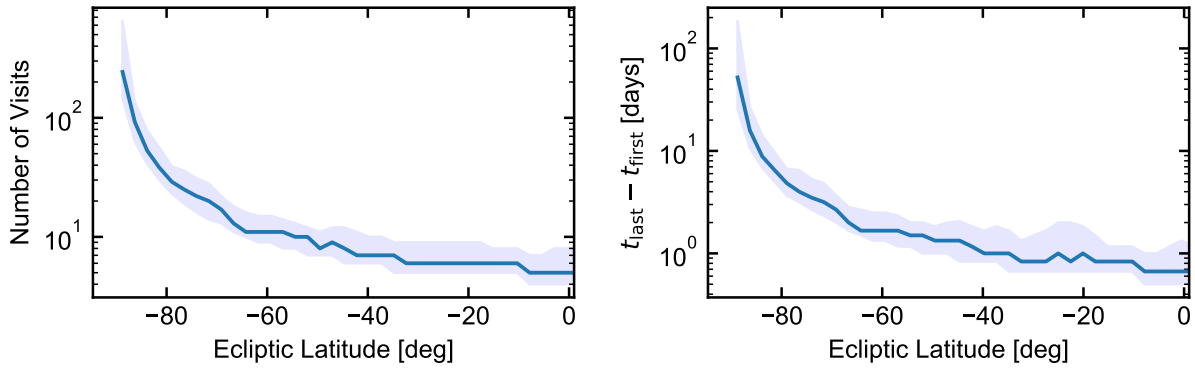


Figure 1.10: Left: The number of visits that a point source receives per eRASS (where a visit is defined as a single passage of a source through eROSITA's FoV) against the ecliptic latitude. Right: the time difference between the first and last visit of each source in eRASS1, against the ecliptic latitude. The data in each figure has been extracted from a preliminary set of eRASS1 light curves generated for nearly every point source detected in eRASS1. In each figure, the solid blue line represents the median value for the given ecliptic latitude bin, whilst the lighter blue region encloses 68% of the observed values.

the ∼40s it takes for a source to pass through the centre of eROSITA's FoV, to the variability between visits of the source every four hours, and to the variability between each eRASS. This, combined with its large grasp in the soft X-rays, results in eROSITA offering an unprecedented insight into the dynamic X-ray sky.

However, eROSITA is predominantly not a transient telescope, and has no dedicated instrument for following-up bright flaring events it discovers during its all sky survey. It also has a predefined survey strategy, such that it is not able to rapidly change its pointing direction in response to interesting new transients. X-ray observatories with extremely wide-area telescopes/ monitors, such as the Burst Alert Telescope on *Swift*, or the Gas Slit Camera on *MAXI*, may still be more suitable than eROSITA for the detection, and/or prompt follow-up, of the brightest transients. Another small disadvantage of eROSITA's cadence during the eRASS is that it will not be able to continuously monitor a source for longer than 40s, which may be a limitation for studying the variability of certain classes of X-ray transient (e.g. inferring recurrence timescales for outbursts, or a periodicity in the light curve of an X-ray binary). Relative to *RXTE* and *MAXI*, eROSITA also offers lower cadence monitoring of sources, excluding eROSITA's coverage of the ecliptic poles, so it may be less useful for studying bright X-ray variables such as the Galactic XRB population, and may provide weaker constraints on when a transient ignited relative to these missions.

It is important to stress, however, that the examples of eROSITA's weaker performance mentioned above predominantly apply to only the brightest transients (and involve comparing eROSITA with X-ray telescopes tailor made for promptly detecting and following-up such events). Thanks to its ability to detect transients down to fainter fluxes, eROSITA is likely to be a transient discovery engine. For the majority of the classes of X-ray transient, eROSITA is expected to significantly expand the size of the known population, and will also provide X-ray light curves that can be used for studies of variability using increased ensemble sizes. Furthermore, it is expected to be extremely useful for the study of certain transient classes. For example, Khabibullin et al. (2014) predict that it will detect 100s-1000s of TDEs per year, which would represent an order of magnitude increase over the population of TDEs known pre-eROSITA. Through its high survey speed capabilities relative to previous X-ray missions (and with serendipity), it is also highly likely to discover examples of X-ray variability phenomena that have yet to be observed.

## 1.3 The launch of the *SRG*/ eROSITA mission in the context of the recent transformation of time domain astronomy

Over the 30 years preceding eROSITA's launch, TDA has undergone a significant transformation. Boosted largely by the adoption of CCDs and the co-improvement of computing power during this time, the search for transients has grown increasingly systematic, with modern synoptic[11] surveys yielding an ever-increasing number of followed-up events. The following section looks

[11] A synoptic sky survey in the literature generally refers to a survey that covers a wide sky area repeatedly over multiple epochs.

at this transformation, particularly focusing on the increasing prevalence of machine-learning within TDA to help handle the large datasets produced from these new surveys. The initial focus in the following discussion is on the optical domain, since the largest change to TDA methodologies has occurred at these wavelengths. The relevance and applicability of these advances to transient science with eROSITA is later highlighted in section 1.3.2.

A major breakthrough in modern TDA came from the supernova search programs during the 1990s. Benefiting from the switch from photographic plates to CCDs, telescopes were able to produce digital images of large sky areas. By repeatedly observing previously imaged sky regions, and combined with improved image subtraction algorithms[12], it suddenly became far easier for astronomers to identify new transients in the most recent images. These advances led to shorter timescales between observation and identification of transient candidates, enabling follow-up observations (spectroscopy and multi-epoch photometry) to be obtained of the most interesting events whilst they were still in outburst. The most significant outcome of these developments was the increased sample sizes of Ia supernovae built up, which when coupled with the discovery that Ia could be used as standard candles for measuring cosmic distances, led to constraints on cosmological parameters that suggested the accelerating expansion of the Universe (Schmidt et al., 1998; Perlmutter et al., 1999).

A number of additional time-domain surveys have been performed since then, and whilst wide-field surveys have been undertaken in other wavebands, such as the *Galaxy Evolution Explorer* in the UV (*GALEX*; Martin et al. 2005), the *Two Micron All-Sky Survey* (*2MASS*; Skrutskie et al. 2006) in the NIR and the *Wide-Field Infrared Survey Explorer* mission (*WISE*; Wright et al. 2010) in the MIR, nearly all of these TDA surveys were in the optical domain (e.g. ASAS-3, Pojmański 2001; ROTSE-III, Akerlof et al. 2003; Supernovae Legacy Survey, Astier et al. 2006; SkyMapper, Keller et al. 2007; Palomar Quest, Djorgovski et al. 2008; Catalina Real-Time Transient Survey, Drake et al. 2009; Palomar Transient Factory, Rau et al. 2009; ASAS-SN, Shappee et al. 2014; *Transiting Exoplanet Survey Satellite*, *TESS*, Ricker et al. 2014; Pan-STARRS, Chambers et al. 2016; ATLAS, Tonry et al. 2018; Zwicky Transient Facility, Bellm et al. 2019). Across each successive survey generation, camera technology has generally improved (e.g. via higher quantum efficiency in the CCDs, which allow a deeper limiting magnitude to be obtained for a given exposure), and (some) telescopes have increased their collecting areas; Bellm et al. (2019) report that the (state-of-the-art) Zwicky Transient Facility offers a $\sim$3 order of magnitude improvement in survey speed over traditional photographic surveys. In addition, improvements in the computing and software aspects enabled prompt exploitation of the observations (further details are presented in the following section). Each survey has different sky coverage, cadences, photometric filters, cameras, fields of view, depths, and thus collectively, probe a variety of time domain phenomena.

Time-critical information about transients discovered through these surveys, that may be of interest to the wider-astronomy community, can now be rapidly disseminated through Virtual Observatory Events (VOEvents[13]), astronomer's telegrams (ATels; Rutledge 1998), the Transient

[12] These subtract a reference image of a given sky region from the latest image of this region. Transients 'pop' up as the residuals of this subtraction process.

[13] Although this format is slightly less common now.

Name Server[14], or through GRB Coordinate Networks (GCNs). A subset of past (e.g. CRTS, Palomar-Quest) and current (e.g. ATLAS, ZTF) transient surveys also release *all* alerts during public survey observations to the public nearly immediately after image processing and transient identification.

Modern TDA thus now consists of a set of synoptic sky surveys that act as the workhorses for transient discovery, and are connected to a large, global network of telescopes capable of promptly following up the most interesting events. An excellent example of how far the field has transformed over the past thirty years is epitomised by the recent identification of a kilonova as the electromagnetic counterpart to a NS-NS merger event (Smartt et al., 2017). On August 17, 2017 (UT), the Advanced LIGO (Aasi et al., 2015) and Advanced Virgo (Acernese et al., 2015) gravitational wave detectors observed the gravitational wave emission from the inspiral and merging of a NS-NS binary. ∼27 minutes after this detection, the source and its localisation sky map was shared with the collaboration, and only ∼11 hours after the GW event (Coulter et al., 2017), its associated optical transient was identified. Only ∼34 hours after the reporting of the GW event, the first optical spectrum of this transient was obtained. This discovery brought astronomy into the multi-messenger era.

### 1.3.1 The rise of the machines

As alluded to in the previous section, there are two main approaches for exploiting time domain datasets in astronomy. The first is a retrospective approach, whereby one generally looks back through the time-varying datasets produced by a survey, and uses these to try and answer a question of interest. For example, one may study the distribution of flare amplitudes in the light curves of an AGN population, or produce a systematic study of the PSD of AGN light curves and explore how these PSDs depend on black hole mass (e.g. Simm et al. 2016). The key drawback of such an approach is that transients of interest may only be identified at late-times after their 'active' phase, and it is not possible to obtain additional follow-up observations during the outburst. For certain classes of transient, this can be detrimental to understanding of the physics of the underlying system, and also for obtaining a robust classification of the event (such as for the kilonova discovery in Smartt et al. 2017, whereby additional photometric and spectroscopic observations were needed to assist identification). Thus the alternative approach is to adopt a pro-active search for new transients during an ongoing survey.

Prior to the latest generation of wide-field, high cadence surveys, the rate of transient detection per night was typically low enough in order to allow for human inspection of each transient candidate. For each source, an expert astronomer would first check whether the transient is real or bogus, and if real, what type of transient it likely is. However, the deluge of data flooding in from surveys over the last 10 years has meant that the rate of increase of newly detected transients dwarfed the rate of transient vetting by humans. So although these surveys offered the potential for detecting and following-up vast numbers of new transients, they also amplified the two following problems:

1. Is the transient real or bogus (e.g. caused by imperfect image subtraction in an optical

[14] `https://www.wis-tns.org/`

survey, or a spurious X-ray source triggered by optical loading from an optically bright star), and is it possible to identify the bogus cases, thus improving the purity of transients detected by a survey?

2. What type of transient is this event?

Machine-learning (ML) based solutions to these problems have typically been most effective over the last decade[15]. One of the main reasons for this efficacy is that ML algorithms can be trained to recognise patterns and correlations in high-dimensional datasets that may be imperceptible to humans, or would be exceptionally difficult to describe using multi-dimensional thresholding. ML approaches can also be extremely fast, which is well suited to the needs of the TDA community as it begins to explore transients of ever shorter timescales, where response times between transient discovery and follow-up should be minimised. Finally, ML algorithms, when being applied, do not suffer from the subjective nature of a human's judgement[16] and offer the chance for reproducible results (i.e. if a human had to classify a set of 1000 objects twice, the results from the two runs may differ, whereas a ML-trained classifier will offer a consistent performance across each run). In the following, a brief overview of the main applications of ML within TDA is presented.

The process of selecting transients rarely produces a pure set of objects which have truly increased in flux between visits of a given sky region. In optical surveys, various artefacts may be introduced to residual images produced from the reference image subtraction process (e.g. imperfect bright star masking, imperfect PSF matching, readout streaks), which may mimic transients. Similarly, artefacts may be introduced into radio imaging data from the imperfect imaging process (Frail et al., 2012). In the X-ray domain, many factors, such as spacecraft pointing instabilities, soft proton flares or optical loading, may produce spurious flux increases of a source. Such impurities contaminate transients, and if left unaccounted for, can waste follow-up resources and the time of any astronomer inspecting the datasets. A number of papers have used ML for classifying a given transient candidate as real/ spurious based upon the output from image subtraction (see Duev et al. 2019 and references therein), whilst Farrell et al. (2015) use a random forest classifier (Breiman, 2001) to identify spurious sources in the 3XMM catalogues with 95% accuracy. Such works demonstrate the strong capabilities of ML for vastly cutting down the number of spurious transients identified.

A second key application of ML within TDA is for the automatic classification of each observed transient in a survey. Driven by the large numbers of transients detected by optical synoptic surveys, and in particular in preparation for the start of operations of LSST, the bulk of the literature on this has focused on classifying optical transients based on their observed multi-filter photometric light curves. Earlier work generally considered a retrospective classification of each light curve (see mini-overview in Lochner et al. 2016), whilst more recently, Muthukrishna et al. (2019) and Möller & de Boissière (2020) have developed classifiers (based on deep recurrent neural networks) to produce probabilistic classifications of continuously evolving light curves

[15]Fortunately, the development and improved accessibility of machine learning algorithms to the wider scientific community coincided with the onset of these new surveys.

[16]However, they can of course be affected by human judgement through the training dataset.

during an ongoing survey. Some of the main principles of these photometric light curve classification efforts are also transferable to other wavebands; for example, Sooknunan et al. (2021) adapt the approaches to optical transient classification in Lochner et al. (2016) and Revsbech et al. (2018) for the classification of radio transients. In the high-energy domain, Lo et al. (2014) and Farrell et al. (2015) used a random forest classifier to automatically classify the variable and transient sources within the 3XMM source catalogue. However, there have been no attempts to date to run an automated classification of transients found live during an ongoing X-ray survey.

A looming issue in transient astronomy is that the number of alerts (potentially interesting objects) detected by wide field surveys is increasing much faster than the rate at which objects can be followed up, so it's also becoming increasingly important to prioritise which objects warrant additional 'expensive' follow-up. The current bottleneck in spectroscopic follow-up (Kulkarni 2020 estimate only ∼10% of transients reported to the Transient Name Server in 2019 were spectroscopically classified), will only be amplified by the next generation of time-domain surveys, such as LSST. To circumvent this issue, various new ML based anomaly detection techniques are currently being explored to aid discovering the most extreme, rare transients (Ishida et al., 2019; Villar et al., 2020). Such approaches may yield good targets for triggering dedicated follow-up campaigns to study new physics, and for optimising the scientific yield of next-generation surveys.

### 1.3.2 Benefits of the transformation of time domain astronomy for eROSITA

The advances in TDA over the last 30 years will provide extensive support for exploiting the transient sky of eROSITA. Firstly, for a large fraction of newly detected transients, it will now be possible to look at the variability in the years preceding the eROSITA detection, as observed by the multiple synoptic sky surveys that have operated during this period, with this variability information likely to be extremely valuable for source classification and characterisation. Secondly, the growing number of applications of ML in TDA has demonstrated the efficacy of such approaches for dealing with large numbers of transients uncovered by a given survey, and how these may be suitable for implementing in pipelines analysing eROSITA's transients. Finally, the generally improved infrastructure and connectivity of modern transient astronomy means that it will now be much easier to quickly trigger and coordinate follow-up campaigns of transients of interest.

## 1.4 Outline of this thesis

This thesis is structured as follows:

As the main focus of this thesis is the search for TDEs in galaxy nuclei, Chapter 2 presents a general overview of TDEs that covers the basic theoretical predictions of these events, the observational status of the field, and the motivations for studying TDEs. This section thus provides the necessary background for Chapters 3 and 4.

In Chapter 3, we estimate the rate of eROSITA detecting the luminous X-ray transients produced in the aftermath of the tidal disruption of white dwarfs by intermediate mass black holes

(WDTDEs). Relative to TDEs involving main-sequence stars and supermassive black holes, WDTDEs are more explosive phenomena and evolve over much shorter timescales. This work was performed in advance of the launch of *SRG*, and with the initial aim of understanding (and preparing for) whether we might be able to use such events to probe the elusive population of intermediate mass black holes.

Chapter 4 contains an analysis and interpretation of the spectacular nuclear transient AT 2019avd. This source was detected during eRASS1, where it was initially designated as a strong TDE candidate on the basis of its ultra-soft, large amplitude X-ray flaring from the nucleus of a previously inactive galaxy. However, AT 2019avd also showed a double-peaked optical light curve (in the year preceding the eRASS1 detection) that can not be easily explained in the canonical TDE scenario, and showed such an extreme set of multi-wavelength properties that no current models for nuclear variability can explain well what triggered such an outburst.

Chapter 5 details a pipeline developed for the identification of transients during the eROSITA All-Sky Survey, and the application of machine learning for an automated classification of transients detected by eROSITA. Operating within a person-power limited consortium, there is not enough time to manually vet and classify each of the transients that are identified during the eROSITA All-Sky Survey. The described pipeline in this section is therefore used to assist astronomers with promptly identifying the more interesting transients uncovered by eROSITA.

Finally, in Chapter 6, I present an outlook for eROSITA transient studies, and summarise the main conclusions of this thesis.

# Chapter 2

# Stellar Tidal Disruption Events

This chapter presents an overview of the physics behind stellar tidal disruption events, followed by a short review of the observational state of the field.

## 2.1 Theoretical overview

A star will be tidally disrupted if it passes too close to a black hole (BH), where strong tidal forces in the vicinity of the BH exceed the self-gravity of the star. Disruption takes place if the pericentre radius, $r_{\rm p}$, of the star's orbit, is less than the tidal radius (Hills, 1975)[1]:

$$\begin{aligned} r_{\rm t} &= \left(\frac{M_{\rm bh}}{M_\star}\right)^{1/3} R_\star \\ &\approx 7\times10^{12}\left(\frac{M_{\rm bh}}{10^6 M_\odot}\right)^{1/3}\left(\frac{M_\star}{M_\odot}\right)^{-1/3}\frac{R_\star}{R_\odot}\,{\rm cm} \end{aligned} \tag{2.1}$$

where $M_{\rm bh}$ is the BH's mass, $M_\star$ and $R_\star$ are the star's mass and radius respectively.

However, if $r_{\rm t}$ lies inside the event horizon (also known as Schwarzschild radius for a non-spinning BH) of the BH ($r_{\rm s} = 2GM_{\rm bh}/c^2$ for a non-spinning BH), then no tidal disruption occurs and the star is swallowed whole. Since $r_{\rm t} \propto M_{\rm bh}^{1/3}$ and $r_{\rm s} \propto M_{\rm bh}$, there exists an upper limit on $M_{\rm bh}$ that can tidally disrupt an object of a given mass (the Hills mass; Hills 1975):

$$M_{\rm H} = M_\star\left(\frac{c^2 R_\star}{2G}\right)^{3/2} \approx 1\times10^8\left(\frac{R_\star}{R_\odot}\right)^{3/2}\left(\frac{M_\star}{M_\odot}\right)^{-1/2} M_\odot \tag{2.2}$$

thus a solar-like object can only be tidally disrupted by non-spinning BHs with masses $\lesssim 10^8 M_\odot$. The scaling of $M_{\rm H} \propto \rho_\star^{-1/2}$ also leads to a wide range of Hills masses for different types of stellar objects. For example, $M_{\rm H} \sim 3\times10^9 M_\odot$ for a red giant (e.g. $M_\star = 1M_\odot$, $R_\star \sim 10R_\odot$), whereas $M_{\rm H} \sim 1\times10^5 M_\odot$ for a white dwarf (e.g. $M_\star = 0.6M_\odot$, $R_\star \sim 0.01R_\odot$).

[1] $r_{\rm t}$ is computed by finding the radius, $r$, where the tidal force around the BH, $F_{\rm t} = GM_{\rm bh}M_\star R_\star/r^3$, is equal to the self-gravity of the star, $F_{\rm self} = GM_\star^2/R_\star^2$.

The broad picture of a stellar tidal disruption in Newtonian physics is well captured by the 'impulse' approximation model. Initially, the star is in hydrostatic equilibrium, and is scattered, via two-body gravitational relaxation (e.g. Magorrian & Tremaine 1999), onto effectively a parabolic orbit[2] about the BH that sees its pericentre radius passing within the tidal radius.

Since the tidal force about the BH scales $\propto r^{-3}$, the star effectively remains unperturbed up until $r_{\rm t}$. As it passes within $r_{\rm t}$, the star is instantaneously disrupted and torn apart. Each fluid element of the star keeps its velocity prior to disruption, but a spread of specific energies in the debris, $\Delta\epsilon$, is induced as a result of different parts of the debris lying at different positions in the BH's gravitational field. $\Delta\epsilon$ can be computed from the difference of the gravitational potential, $\phi(r)$, across the star's radius, using a first-order Taylor expansion (Lacy et al., 1982):

$$\begin{aligned}\Delta\epsilon &= \phi(r+R_\star) - \phi(r) \\ &= \frac{-GM_{\rm bh}}{r+R_\star} - \frac{-GM_{\rm bh}}{r} \\ &\approx \frac{GM_{\rm bh}R_\star}{r^2},\end{aligned} \tag{2.3}$$

which at the tidal radius is equal to:

$$\begin{aligned}\Delta\epsilon &= \frac{GM_{\rm bh}R_\star}{R_\star^2}\left(\frac{M_\star}{M_{\rm bh}}\right)^{2/3} \\ &= \frac{GM_\star}{R_\star}\frac{M_{\rm bh}}{M_\star}\left(\frac{M_{\rm bh}}{M_\star}\right)^{2/3} \\ &= \epsilon_\star\left(\frac{M_{\rm bh}}{M_\star}\right)^{1/3}\end{aligned} \tag{2.4}$$

where $\epsilon_\star = GM_\star/R_\star$ is the specific binding energy of the star, and $|\,\Delta\epsilon\,| \gg \epsilon_\star$ for tidal disruptions involving BHs with $M_{\rm bh} > 10^6 M_\odot$. The induced spread in $\epsilon$, which is effectively frozen-in at the tidal radius (Stone et al., 2013), subsequently governs the evolution of the debris. Assuming a symmetric spread in $\epsilon$ about 0, approximately half of this has $-\Delta\epsilon < \epsilon < 0$ and remains bound to the BH, whilst the other half has $0 < \epsilon < \Delta\epsilon$ and is ejected from the system on hyperbolic orbits.

The most bound debris, with specific energy $\sim -\Delta\epsilon$, is on highly eccentric orbits with a semi-major axis[3]:

$$\begin{aligned}a_{\rm min} &= \frac{R_\star}{2}\left(\frac{M_{\rm bh}}{M_\star}\right)^{2/3} \\ &\approx 4\times10^{14}\left(\frac{R_\star}{R_\odot}\right)\left(\frac{M_{\rm bh}}{10^6 M_\odot}\right)^{2/3}\left(\frac{M_\star}{M_\odot}\right)^{-2/3}\ {\rm cm}\end{aligned} \tag{2.5}$$

[2]The majority of stars are scattered into the loss cone for tidal disruption from orbits with apocentre radii much larger than the tidal radius, thus the orbit is highly eccentric and is approximated to be parabolic.

[3]Using the relation between specific energy, $\epsilon$, and semi-major axis, $a$, for Keplerian orbits $\epsilon = -GM_{\rm bh}/2a$.

and will fall back to pericentre on a timescale (using Kepler's third law):

$$\begin{aligned} t_{\rm min} &= \frac{\pi}{\sqrt{2G}}\left(\frac{R_\star^3}{M_\star}\right)^{1/2}\left(\frac{M_{\rm bh}}{M_\star}\right)^{1/2} \\ &= 41\left(\frac{R_\star}{R_\odot}\right)^{3/2}\left(\frac{M_\star}{M_\odot}\right)^{-1}\left(\frac{M_{\rm bh}}{10^6 M_\odot}\right)^{1/2}\ \text{days}. \end{aligned} \tag{2.6}$$

The rate of debris fallback to pericentre can then be computed using the chain rule (Rees, 1988; Phinney, 1989):

$$\dot{M}_{\rm fb} = \frac{{\rm d}M}{{\rm d}t} = \frac{{\rm d}M}{{\rm d}\epsilon}\frac{{\rm d}\epsilon}{{\rm d}t}, \tag{2.7}$$

where ${\rm d}M/{\rm d}\epsilon$ is the distribution of the specific energies in the bound debris, which was initially assumed to be uniform in Rees (1988) and equal to $M_\star/2\Delta\epsilon$. Using again the Keplerian relation $\epsilon = -GM_{\rm bh}/2a$ and Kepler's third law, the specific energy scales $\propto t^{-2/3}$, such that $\dot{M}_{\rm fb} \propto {\rm d}\epsilon/{\rm d}t \propto t^{-5/3}$. Through requiring that $M_\star/2$ must eventually fall back onto pericentre after $t_{\rm min}$:

$$M_\star/2 = \int_{t_{\rm min}}^{\infty} A t^{-5/3}{\rm d}t, \tag{2.8}$$

and solving for the constant of proportionality, $A$, then the evolution of the mass fallback rate is:

$$\dot{M}_{\rm fb} = \dot{M}_{\rm peak}\left(\frac{t}{t_{\rm min}}\right)^{-5/3} \tag{2.9}$$

where the peak mass fallback rate, $\dot{M}_{\rm peak}$ is:

$$\begin{aligned} \dot{M}_{\rm peak} &= \frac{M_\star}{3t_{\rm min}} \\ &\approx 3\left(\frac{M_{\rm bh}}{10^6 M_\odot}\right)^{-1/2}\left(\frac{R_\star}{R_\odot}\right)^{-3/2}\left(\frac{M_\star}{M_\odot}\right)^{2} M_\odot\,{\rm yr}^{-1}. \end{aligned} \tag{2.10}$$

Through comparing this with the Eddington accretion rate onto the BH:

$$\dot{M}_{\rm Edd} = \frac{4\pi G M_{\rm bh} m_{\rm p}}{\sigma_{\rm T}\eta c}, \tag{2.11}$$

where $m_{\rm p}$ is the proton mass, $\sigma_{\rm T}$ is the Thomson cross-section, and $\eta$ is the radiative efficiency of accretion, then the ratio:

$$\begin{aligned} \frac{\dot{M}_{\rm peak}}{\dot{M}_{\rm Edd}} &= \frac{\sqrt{2}\eta c\sigma_{\rm T}}{12\pi^2 m_{\rm p}\sqrt{G}} \\ &\approx 130\left(\frac{\eta}{0.1}\right)\left(\frac{M_{\rm bh}}{10^6 M_\odot}\right)^{-3/2}\left(\frac{R_\star}{R_\odot}\right)^{-3/2}\left(\frac{M_\star}{M_\odot}\right)^{2} \end{aligned} \tag{2.12}$$

implies that super-Eddington mass fallback rates are expected in the early stages of TDEs (although disruptions of Sun-like stars about BHs with $M_{\rm bh} \gtrsim 3 \times 10^7 M_\odot$ will have sub-Eddington peak fallback rates).

After disruption, the BH's tidal forces shape the stellar debris into elongated, thin debris streams on highly eccentric orbits about the BH. In the canonical TDE scenario, it is assumed that this debris is then able to promptly circularise around the BH to form a quasi-circular accretion disc with outer radius at $\sim 2r_{\rm t}$. The dominant energy dissipation mechanism here is currently thought to be due to shocks produced from the intersection of the debris streams (Guillochon & Ramirez-Ruiz, 2015). On the additional assumption that the viscous timescale in this nascent disc is much less than the fallback timescale of the most bound debris, then the accretion rate traces the mass fallback rate. The bolometric luminosity of the tidal disruption flare is then:

$$\begin{aligned} L(t) &= \eta \dot{M}_{\rm acc} c^2 \approx \eta \dot{M}_{\rm fb} c^2 \\ &\approx \eta \dot{M}_{\rm peak} c^2 \left(\frac{t}{t_{\rm min}}\right)^{-5/3}, \end{aligned} \tag{2.13}$$

and approximating the emission as blackbody with characteristic radius set to $r_{\rm t}$ (Ulmer, 1999), then the effective temperature at peak is (assuming Eddington-limited):

$$\begin{aligned} T_{\rm eff} &\approx \left(\frac{L_{\rm Edd}}{4\pi r_{\rm t}^2 \sigma_{\rm SB}}\right)^{1/4} \\ &\approx 3 \times 10^5 \left(\frac{M_{\rm bh}}{10^6 M_\odot}\right)^{1/12} \left(\frac{R_\star}{R_\odot}\right)^{-1/2} \left(\frac{M_\star}{M_\odot}\right)^{-1/6} \text{K}. \end{aligned} \tag{2.14}$$

For the disruption of a solar-like star with $M_{\rm bh} = 10^6 M_\odot$, this peaks in the far-UV/ soft X-ray range ($kT \sim 30$ eV).

## 2.2 Observations

### 2.2.1 X-ray selected TDEs

The first strong TDE candidates were discovered using the *ROSAT* observatory (Trümper, 1982) in the 1990s. Broadly consistent with the observational signatures predicted in Rees (1988), these events (NGC 5905, Bade et al. 1996; Komossa & Bade 1999; RX J1242.6–1119, Komossa & Greiner 1999; RX J1624+7554, Grupe & Leighly 1999; RX J1420+5334, Greiner et al. 2000) were ultra-soft X-ray sources (blackbody temperatures, $30 \lesssim kT \lesssim 100$ eV; example X-ray spectrum shown in Fig. 2.1) associated with galaxies that showed no evidence for pre-existing AGN activity in follow-up spectroscopic observations. These candidates showed large amplitude X-ray flux declines (factors up to several thousand) over the course of follow-up observations in the years following their initial *ROSAT* detection, and combined with 0.1-2.4 keV peak luminosities between $\sim 7 \times 10^{42}$ erg s$^{-1}$ and $4 \times 10^{44}$ erg s$^{-1}$, were extremely likely to be linked with accretion onto supermassive black holes. Since the X-ray light curves were sparsely followed-up, it was

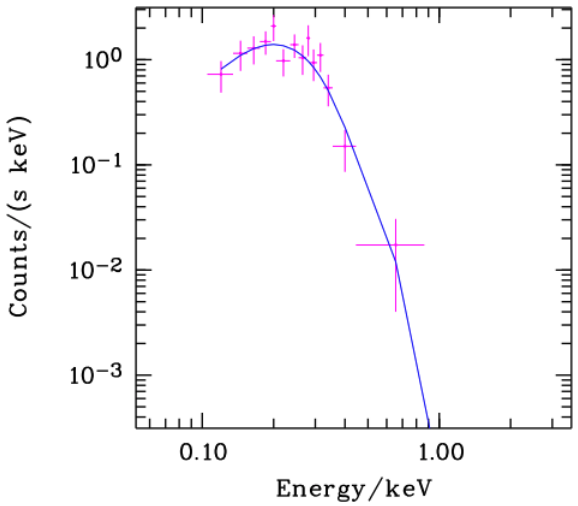


Figure 2.1: Ultra-soft *ROSAT* spectrum of the TDE candidate RX J1420+5334 (Greiner et al., 2000). The pink markers show the binned observed spectrum, whilst the blue line is the best fitting blackbody model with $kT = 38 \pm 10$ eV. Figure from Greiner et al. (2000).

not possible to place tight constraints on the decay slope of the events and compare with the predicted $L \propto t^{-5/3}$ rate.

The launch of *Chandra* (Weisskopf et al., 2000), *XMM-Newton* (Jansen et al., 2001), and the *Swift* observatory (Gehrels et al., 2004) in the early 2000s enabled the identification of an additional set of X-ray selected tidal disruption flares. These TDE candidates all originated from previously quiescent galaxies out to a redshift of $z = 0.2$, with some also located in dwarf galaxies (Maksym et al., 2013; Lin et al., 2017a), highlighting how TDEs may prove valuable for uncovering otherwise quiescent intermediate mass black holes.

These additional TDE candidates were identified through: i) cross-matching the *ROSAT* source catalogues with the *XMM* and *Chandra* serendipitous source catalogues and searching for objects associated with galaxies that showed large amplitude flux declines (Cappelluti et al., 2009; Khabibullin & Sazonov, 2014), ii) searching for large amplitude flux declines in cross-matched *XMM* and *Chandra* source catalogues (e.g. Lin et al. 2017b), or iii) searching for bright new X-ray transients discovered by the *XMM*-Slew survey and promptly following these up (Esquej et al., 2007; Saxton et al., 2012; Mainetti et al., 2016; Saxton et al., 2017, 2019). The majority of these candidates again showed ultra-soft X-ray spectra, typically dominated by a blackbody component with $40 \lesssim kT \lesssim 280$ eV, and soft-band observed peak luminosities ranging from $\sim 5{\times}10^{41}$ erg s$^{-1}$ to $\sim 3{\times}10^{44}$ erg s$^{-1}$. More recently however, two TDE candidates detected in the *XMM*-Slew survey also showed strong X-ray emission above 2 keV, with the X-ray spectrum of XMMSL2 J144605.0+685735 (Saxton et al., 2019) and XMMSL1 J074008.2-853927 (Saxton et al., 2017) being best modelled by a power-law of slope $\sim 2.6$ and $\sim 2$ respectively (XMMSL1 J074... also has a weak blackbody-like soft excess with $kT \sim 86$ eV). Whilst this suggests that the X-ray emission from TDEs is not limited to being ultra-soft and that a fraction of TDEs may also be capable of forming harder X-ray spectra, the driving mechanisms for

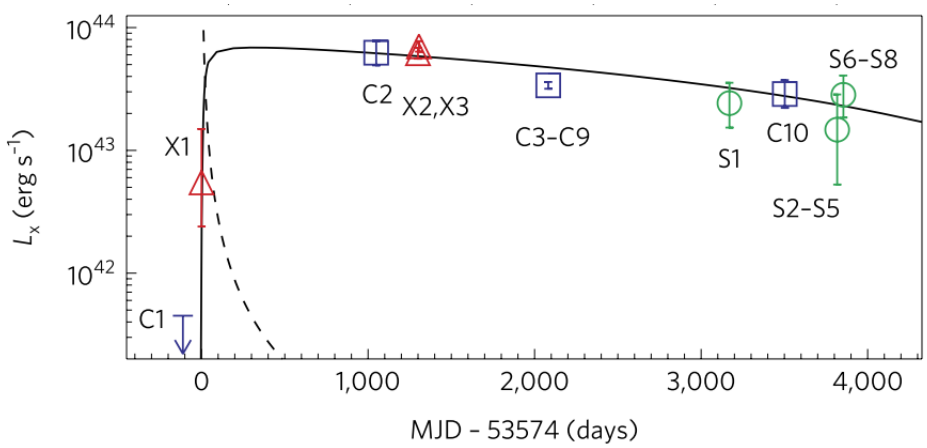


Figure 2.2: X-ray light curve of a slowly evolving TDE candidate, 3XMM J150052.0+015452 Lin et al. (2017a). Blue, red and green markers represent observations obtained with *Chandra*, *XMM* and *Swift* XRT respectively. This candidate was found based on cross-matching the *XMM* and *Chandra* serendipitous source catalogues, which led to a very sparsely sampled light curve. Figure from Lin et al. (2017a).

this spectral diversity are currently unclear (see discussion in Saxton et al. 2017). The X-ray light curves of these candidates, while sparsely and irregularly sampled (e.g. only one datapoint obtained every few years for Lin et al. 2017b; see also Fig. 2.2), typically show large amplitude flux increases relative to previous flux upper detection limits, and/ or major flux declines over timescales of months to years (a review of these events is presented in Saxton et al. 2020). Prior to the launch of eROSITA, no X-ray selected, non-jetted TDEs have yet shown simultaneous transient optical emission (detectable above their host galaxy), although several have been observed to show a declining UV flux (Fig. 2.3)

There also exists a population of relativistic tidal disruption flares (Swift J164449.31573451, Bloom et al. 2011; Burrows et al. 2011; Zauderer et al. 2011; Levan et al. 2011; Swift J2058.4+0516, Cenko et al. 2012; Swift J1112.2-8238, Brown et al. 2015), which again are associated with previously quiescent galaxies, but show starkly different X-ray properties to the X-ray selected TDEs described above. All of these relativistic events initially emitted a several day long gamma-ray burst that triggered the *Swift* Burst Alert Telescope (BAT). Prompt follow-up observations revealed peak observed isotropic X-ray luminosities between $\sim 6 \times 10^{46}$ erg s$^{-1}$ and $10^{48}$ erg s$^{-1}$, and displayed hard X-ray spectra with power-law slopes between $\sim 1.3$ and $\sim 2$. The X-ray light curves were extremely variable in the early observations (showing dips and a potential quasi-periodicity, Reis et al. 2012), and each of these events launched a relativistic jet, producing a luminous radio transient (Bloom et al., 2011; Zauderer et al., 2011; Cenko et al., 2012; Brown et al., 2017). On the basis of this rapid X-ray variability seen in these systems, the X-ray emission is thought to originate at the base of this jet, as opposed to shock fronts formed from the jet expanding in to the external ambient medium (e.g. Zauderer et al. 2013). In addition, X-ray monitoring campaigns showed that the 0.2-10 keV flux declined for nearly a year after the initial gamma-ray trigger, before showing a sharp drop-off at late times (factors of $\sim 170$ $\sim 160$ and in

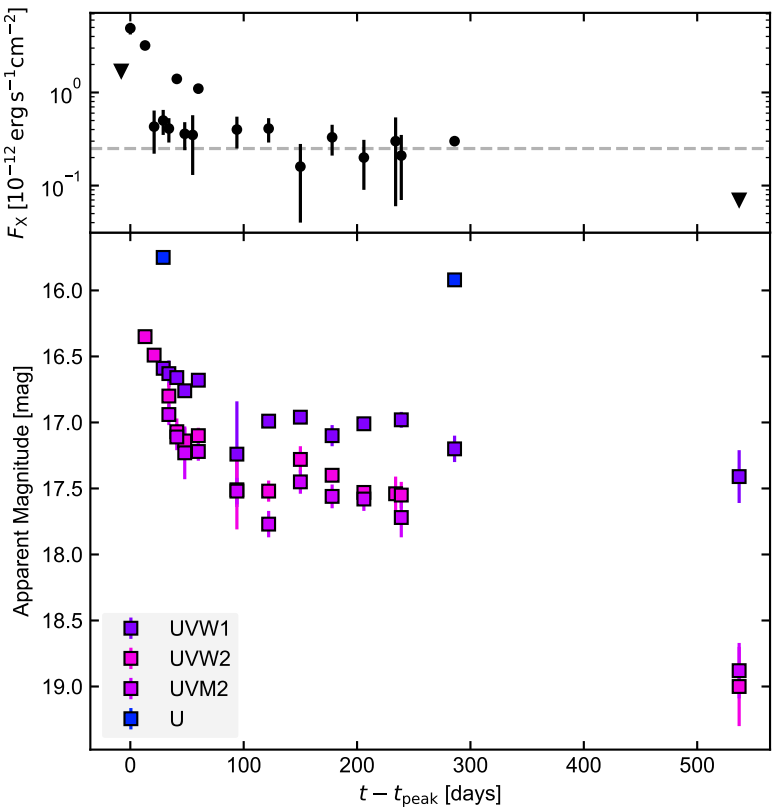


Figure 2.3: X-ray (0.2-2 keV flux) and UV light curve of the TDE candidate XMMSL1 J074008.2-853927 (Saxton et al., 2017). Black markers denote X-ray observations where the source was detected, whilst the black triangles represent $2\sigma$ flux upper limits. The grey dashed line denotes the $2\sigma$ upper flux detection limit in the 0.2-2 keV band from *ROSAT* observations in 1990 (when the source was not previously detected). The peak observed 0.2-2 keV flux is ~20 times brighter than this. Figure plotted using data from Saxton et al. (2017).

J164449.31573451 and J2058.4+0516 respectively; Zauderer et al. 2013; Pasham et al. 2015), with this potentially associated with the accretion rate transitioning from super to sub-Eddington and the shutdown of the jet (e.g. Zauderer et al. 2013).

In total, approximately 20 X-ray selected TDE candidates have been discovered since the launch of *ROSAT* in the 1990s (Saxton et al., 2020). This sample is expected to grow significantly during the four year eROSITA All-Sky Survey, where (pre-launch) estimates predicted the discovery of 100s-1000s of new TDEs every six months (Khabibullin et al., 2014; Jonker et al., 2020).

### 2.2.2 Optically selected TDEs

The vast majority of recently discovered TDE candidates have been through wide-field, high cadence optical surveys, with $\sim$ 30 discovered over the last decade. Major contributions to this discovery rate have been provided by the All Sky Automated Survey for Super Novae (ASAS-SN; Shappee et al. 2014; Kochanek et al. 2017), and the Zwicky Transient Facility (ZTF; e.g. Bellm et al. 2019; Graham et al. 2019).

The key observational characteristics of these events (see van Velzen et al. 2020 for a detailed review) are blue (mean $g - r \lesssim 0$ mag), nuclear transients (positions astrometrically consistent with their host galaxy centres), which rise to peak luminosities $\sim 10^{43}$ to $\sim 10^{44}$ erg s$^{-1}$ over timescales of weeks, and decay over month to year long timescales (Fig. 2.4); this translates into longer/ shorter rise and decay on timescales compared with AGN/ supernovae flares (Fig. 2.5). The optical spectra of these systems in the first $\sim$100 days post-peak typically show dominant blue continua, with various broad Balmer, He II and/ or Bowen (Bowen, 1928) emission lines (with full width at half maximum $\lesssim 10^4$ km s$^{-1}$).

One of the main puzzles of TDE science introduced by the discovery of optically-selected TDEs, is that these systems rarely show transient X-ray emission. For example, 25% of optically selected TDEs in van Velzen et al. 2021 were X-ray bright. This was not a predicted observational signature in the canonical TDE scenario described initially in Rees (1988). In cases where the optically-selected TDE is X-ray bright, then the characteristic blackbody temperatures and radii of the optical emission are a few $\times 10^4$ K ($\sim$1 eV) and $\sim 10^{14} - 10^{15}$ cm, but $\sim 10^6$ K ($\sim$100 eV) and $\sim 10^{11}$ cm for the X-ray emission. Whilst the soft X-ray emission in TDEs is generally consistent with being produced from the innermost regions of the nascent accretion disc, the mechanism responsible for the optical emission is currently not clear, since the inferred radii (that the optical radiation is emitted) are $\sim$2-3 orders of magnitude too large to be associated with emission at the outer radius of a circularised accretion disc ($\sim 2r_t$; equation 2.1).

It has been suggested that the optical emission may instead originate from the reprocessing of the X-ray emission from an accretion disc, by obscuring gas that lies along the line-of-sight to the disc (e.g. Loeb & Ulmer 1997; Ulmer et al. 1998; Roth et al. 2016; Roth & Kasen 2018). Strong evidence for this is based on the detection of Bowen emission lines in optically-bright, X-ray faint TDEs (e.g. Leloudas et al. 2019). To produce such lines requires a high flux of photons with energies above the ionisation potential of He II (54.4 eV), thus TDEs showing Bowen lines, but no associated transient X-ray emission, should host accreting SMBHs which are X-ray faint due to the strong reprocessing of their X-rays into optical, possibly by a much larger debris envelope

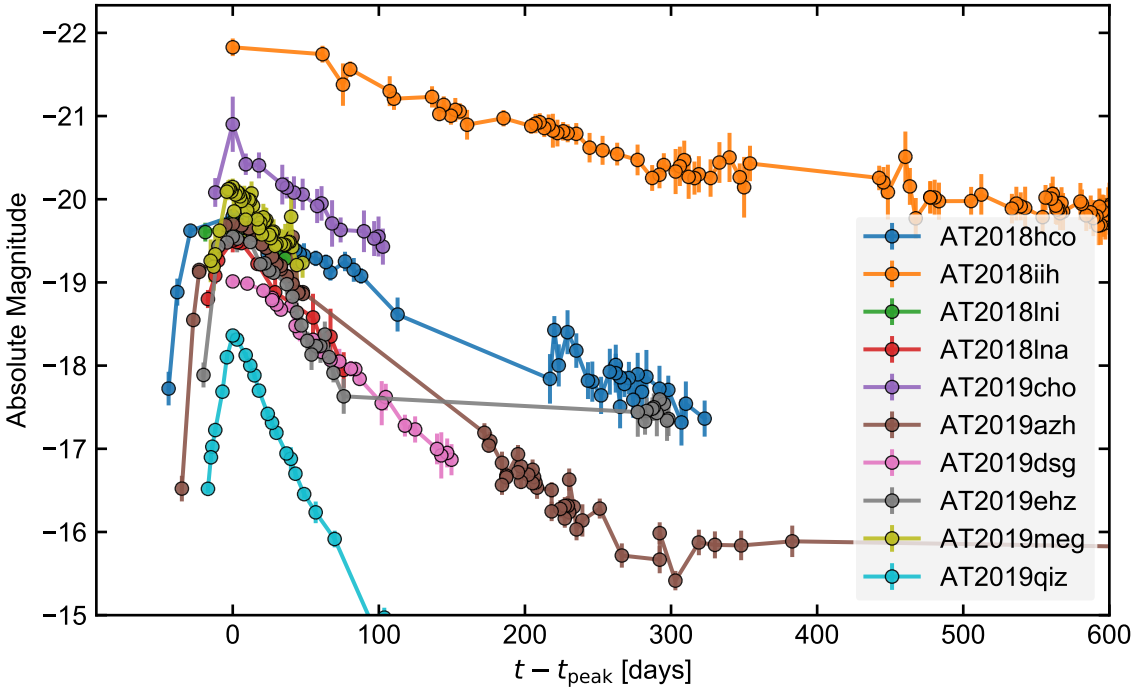


Figure 2.4: $g$-band light curves of a subset of the population of TDEs identified by ZTF in van Velzen et al. (2021). Contrasting X-ray selected TDEs to-date (such as in Fig. 2.2), wide-field and high cadence optical surveys offer light curves with much higher frequency time-sampling.

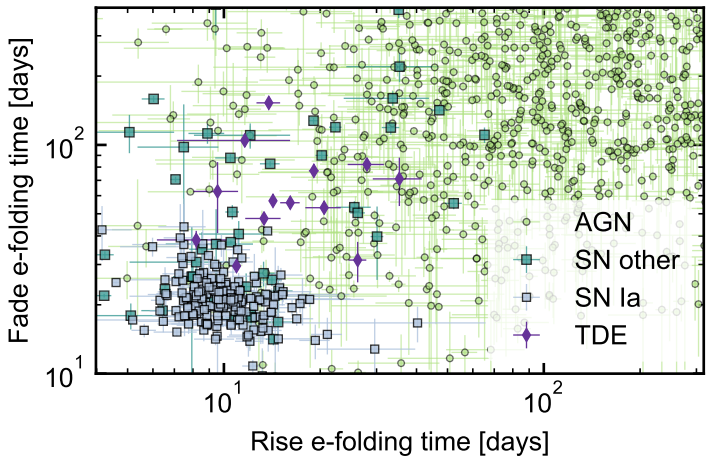


Figure 2.5: Rise and decay timescales for the set of spectroscopically classified nuclear transients identified by ZTF, adapted from data initially presented in van Velzen et al. (2021).

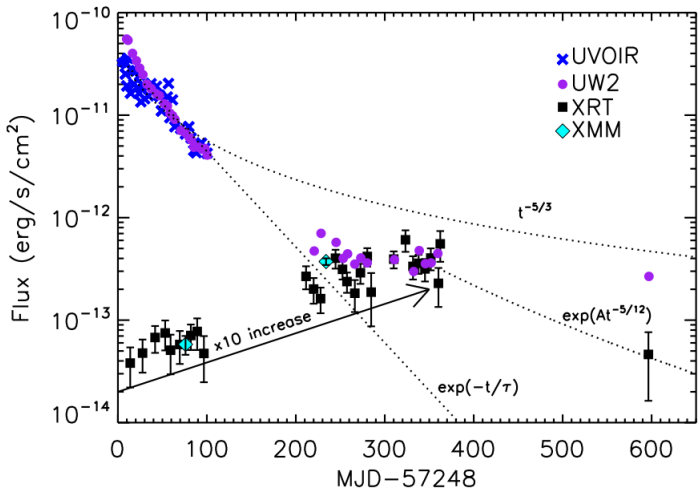


Figure 2.6: Late-time X-ray brightening of the optically-selected TDE candidate ASAS-SN 15oi Gezari et al. (2017). Figure from Gezari et al. (2017).

that forms in the aftermath of the initial disruption. Under such a reprocessing scenario, the apparent dichotomy between X-ray bright and faint systems may be explained through a viewing angle dependence (Dai et al., 2018), if the geometry of the reprocessor is linked with the disc orientation.

Alternatively, the optical emission may be produced by shocks formed at the stellar debris stream self-intersections (Shiokawa et al., 2015b; Piran et al., 2015). The optical emission would then appear to originate from a photosphere of radius $\sim a_{\rm min}$ (equation 2.5), which is roughly consistent with the inferred blackbody radii of the optical emission. Under this model, there can be inefficient circularisation of the debris post-disruption, delayed formation of an accretion disc, and subsequently weak X-ray emission in these systems when followed-up at early times. This scenario is supported by the discovery of several optically-selected TDEs (Gezari et al., 2017; Hinkle et al., 2020; Kajava et al., 2020) that have shown a late-time X-ray brightening relative to their optical emission (e.g. Fig. 2.6).

## 2.3 Motivations for studying TDEs

Aside from constituting spectacular, cataclysmic astrophysical events that are interesting to study in themselves, TDEs hold promise for offering additional insights into a range of phenomena.

A robust detection of a TDE reveals the presence of a (potentially) previously unknown BH candidate. While it is widely accepted that SMBHs reside in the centres of galaxies, the BH occupation fraction in galaxies grows increasingly uncertain towards the low mass end of the galaxy mass function (see Chapter 3 of this thesis for further discussion). Since the rate of TDEs in a galaxy is expected to increase as $M_{\rm bh}$ decreases (Stone & Metzger 2016 report an empirical rate scaling proportional to $M_{\rm bh}^{-0.4}$), then a large fraction of observed TDEs may originate from BHs with $M_{\rm bh} \lesssim 10^6 M_\odot$, and the observed TDE population may be used to place constraints on

the BH occupation fraction (e.g. van Velzen 2018). TDEs have already been used to identify IMBH candidates that had previously shown no strong evidence for being there, such as in globular clusters (Maksym et al., 2013) or dwarf galaxies (Lin et al., 2017a). Additionally, the vast release of energy from the flare may allow us to probe the pristine, sub-parsec circumnuclear environment of the disrupting BH, such as through radio monitoring of TDE-launched outflows (Alexander et al., 2016), or mid-infrared dust echoes (van Velzen et al., 2016).

Since the timescales that TDEs evolve over are linked to the BH mass (equation 2.6), the light curve is expected to encode information about the properties of the disrupting BH, and it is hoped that TDEs may eventually be able to 'weigh' black holes. If this is achieved, then TDEs would therefore offer a unique mass probe into quiescent BHs that are too distant for their sphere of influence to be spatially resolved (and therefore have a kinematic mass estimate available; e.g. Kormendy & Ho 2013). However, TDE science is still a long way from robustly understanding how the BH mass, stellar properties and stellar orbit affect the observed TDE light curve (see discussion in Mockler et al. 2019).

TDEs also provide the unique opportunity to study accretion physics in SMBHs over accretion rates that vary from being quiescent-like, to super-Eddington and then to sub-Eddington (and effectively quiescence) over realistically observable timescales. The super-Eddington accretion rates predicted in TDEs near their peak fallback rates may also provide promising laboratories for studying the launching of super-Eddington winds (Kara et al., 2018), or for probing the conditions needed for jet formation (Krolik & Piran, 2012).

Lastly, a more general motivation is that we simply do not fully understand the physics of TDEs, and further study of these events might enable us to explain why their observed properties deviate from the canonical TDE scenario (e.g. Rees 1988). It is also still not clear how to distinguish between TDE and non-TDE induced flares in the centres of galaxies, especially in cases where the flare has occurred in a system where the SMBH was previously active (e.g. Trakhtenbrot et al. 2019b). With a rapidly increasing number of nuclear transients being discovered by wide-field, high cadence surveys, learning how to differentiate between TDEs and non-TDEs is therefore growing increasingly important to solve and warrants further exploration.

# Chapter 3

# eROSITA Detection Rates for Tidal Disruptions of White Dwarfs by Intermediate Mass Black Holes

White dwarf-black hole tidal disruption events (herein WTDEs) present an opportunity to probe the quiescent intermediate mass black hole population in the universe. We run an extensive set of Monte-Carlo based simulations to explore *SRG*/eROSITA's detection sensitivity to WTDEs as a function of black hole mass, redshift and time offset between event flaring and it first being observed. A novel estimate of WTDE rate densities from globular clusters and dwarf galaxies is also presented. We combine this with estimated detection sensitivities to infer the rate of eROSITA detecting these events. Depending on the estimate of the intrinsic rate of WTDEs, we anticipate that eROSITA may detect $\sim 3$ events over its 4 year all-sky survey. eROSITA will be most sensitive to systems with black hole masses above $10^4 M_\odot$, and is most likely to catch these within 5 days of flaring.

This work was performed in advance of the launch of eROSITA, and the contents of this chapter were originally published in Malyali et al. (2019).

## 3.1 Introduction

Intermediate-mass black holes (IMBHs), with masses in the range $10^3 - 10^5 M_\odot$, are predicted to reside at the centres of globular clusters Colbert & Mushotzky (1999); Fabbiano et al. (2001); Gultekin et al. (2004) and dwarf galaxies (Dong et al., 2007b; Greene, 2012), yet current observational evidence for the existence of IMBHs is uncertain (see Mezcua 2017 for a review).

One potential avenue for probing the quiescent IMBH population is through Tidal Disruption Events (TDEs), whereby perturbations to a star's orbit lead to it passing too close to the BH and being ripped apart by tidal forces (Hills, 1975; Lacy et al., 1982; Rees, 1988). Stellar populations surrounding IMBHs are currently weakly constrained observationally, although it is expected that the rate of tidal disruption is higher for main sequence (MS) than white dwarf (WD) stars (Ramirez-Ruiz & Rosswog, 2009).

WDs can only be disrupted by black holes with mass $\lesssim 10^5\ M_\odot$ and are swallowed whole by black holes more massive than this, with gravitational radiation likely being the only observational signature in such cases (Luminet & Pichon, 1989b; East, 2014). After disruption, accretion of the bound stellar debris onto the central BH leads to a soft X-ray flare of thermal radiation – for a review of TDE observations see Komossa (2015) and references therein. Theoretical and computational work has predicted a wealth of additional signatures such as Type Ia supernovae-like optical transients due to thermonuclear burning triggered by extreme tidal compression of the white dwarf perpendicular to the orbital plane (Luminet & Pichon, 1989a; Rosswog et al., 2008, 2009; Haas et al., 2012; Kawana et al., 2018), modulation of the luminosity output due to changes in accretion rate produced by elliptical orbit behaviour of the WD (Zalamea et al., 2010), and the launching of relativistic jets (De Colle et al., 2012; Krolik & Piran, 2012; Shcherbakov et al., 2013). A number of candidates for WTDEs have already been reported; some with X-ray and radio jet signatures (Krolik & Piran, 2011; Shcherbakov et al., 2013) and several with non-jetted X-ray signatures (Jonker et al., 2013; Glennie et al., 2015). The rapid decay rates involved (the whole flaring episode occurs over hours to weeks depending on WTDE configuration, as opposed to months to years for MS TDEs – see Fig. 11 of Law-Smith et al. 2017a for comparison of different TDE timescales) makes confident classification of WTDEs difficult as there is usually insufficient multi-wavelength evidence to choose one model over competing models, such as flaring stars or off-axis GRBs.

eROSITA (extended Roentgen Survey with an Imaging Telescope Array) (Merloni et al., 2012; Predehl, 2017), which is on board the Russian-German *Spektrum-Roentgen-Gamma* (*SRG*) mission, was launched in July 2019. Its first four years of operation will be dedicated to an X-ray all-sky survey where it will be ∼20 times more sensitive than its predecessor, ROSAT, in the soft X-ray band (0.5 − 2 keV) and will be the first imaging survey to cover the whole hard X-ray (2-10keV) sky. Khabibullin et al. (2014) explored the rate of eROSITA detecting tidal disruption flares of main sequence (MS) stars by supermassive black holes and estimated $\sim 10^3$ new TDE candidates per all sky scan (ie. every six months) – see also Thorp et al. (2019). However, the more explosive, shorter timescales of WTDEs has yet to be explored and given the extent of theoretical work put into predicting the observational signatures of these transients, it is useful to understand eROSITA's detection sensitivity to these events.

In this work, we explore eROSITA's detection sensitivity to WTDEs through an extensive set of Monte-Carlo based simulations. Based on previous theoretical and computational work we explore the observational signatures of these events, focusing on non-jetted signatures since the underlying jet formation mechanisms are not well understood. We introduce the simulation framework and underlying theory that we use to model WTDEs and associated X-ray observational properties in Section 3.2. eROSITA's detection sensitivity to these events for different WD-BH configurations is explored via simulations in Section 3.3. This is then combined with the intrinsic rate of WTDEs in the local universe to estimate eROSITA's detection rate in Section 3.4. We discuss caveats of our modeling and draw up an approach for multiwavelength follow-up in Section 5.5, before presenting conclusions in Section 4.6.

## 3.2 Simulating eROSITA observations

Following launch, the *SRG* satellite will enter an orbit around the Second Lagrangian (L2) point of the Earth-Sun system, where eROSITA will perform eight All-Sky Surveys (eRASS 1-8) during its first four years. The satellite will rotate once every four hours around an axis pointed a few degrees away from the Sun, with observations moving by 1 degree per day (at the equator) perpendicular to this axis. Point sources will pass through eROSITA's 1° diameter FOV six times per day. With each passage lasting ∼40s, this will lead to ∼240s exposure for each source per eRASS. We note that this is a minimum exposure for a sky region. Due to the non-uniformity of exposure across the sky, the Ecliptic poles will be visited more frequently and sources within these regions will have longer exposures (Merloni et al., 2012). This scanning strategy allows eROSITA to probe variability on timescales from ∼ 40s, to days and up to years, albeit with an inhomogeneous and often sparse light curve coverage.

We used the Monte Carlo based code SIXTE[1] (Schmid, 2012) to simulate eROSITA observations of patches of the sky during the first all-sky survey, eRASS1, where we utilise a spacecraft attitude file to model the time-dependence of eROSITA's pointing direction[2]. SIXTE requires specification of an instrument description file, and an instrument-independent sky model contained in a `simput` file[3], with details of each source's sky position, flux in a reference energy band, X-ray spectrum and light curve (LC). A photon population is then simulated given this sky model and propagated through the instrument model to produce a set of simulated event files. Whilst computationally demanding, this approach has the advantage that it realistically models instrument specific effects on photon propagation and detection which cannot be modelled analytically.

The sky models used in this work consist of a population of WTDEs, an AGN population, and a soft X-ray background (SXB) originating from collisionally-ionized diffuse gas within the Solar system and Galaxy. Details of each of these components are provided in Sections 3.2.1, 3.2.2 and 3.2.3 respectively. We also add a particle background component that is uniform over the detectors (not passing through eROSITA's mirror system) and implemented within the SIXTE simulator based on Tenzer et al. (2010). For simplicity, we do not include galaxy cluster and stellar populations, but discuss the effects of this omission in Section 5.5.

### 3.2.1 WTDEs

#### Temporal properties

A star will be tidally disrupted if it passes too close to a BH. For a system containing a BH of mass $M_{\rm bh}$ and a star with mass and radius of $M_*$ and $R_*$, respectively, this occurs for stellar orbits where the pericenter radius, $r_{\rm p}$, is less than the tidal radius (Hills, 1975), approximately defined

[1] `http://www.sternwarte.uni-erlangen.de/research/sixte/`

[2] Whilst the final strategy could change for operational reasons, it is expected to be very similar.

[3] `https://www.sternwarte.uni-erlangen.de/research/sixte/simput.php`

as:

$$r_{\rm t} \simeq \left(\frac{M_{\rm bh}}{M_*}\right)^{1/3} R_* \tag{3.1}$$

which is the distance from the BH whereby the star's self-gravitational forces are overcome by the extreme tidal forces acting on it.

Stars are swallowed by a non-spinning BH without tidal disruption if $r_{\rm t} < r_{\rm g} = 2GM_{\rm bh}/c^2$, where $r_{\rm g}$ is the BH's gravitational radius. This places an upper constraint on black hole mass for tearing a star apart (eg. Hills 1975; Rosswog et al. 2009):

$$M_{\rm bh} < 2.5 \times 10^5 \left(\frac{R_*}{10^9{\rm cm}}\right)^{3/2} \left(\frac{M_*}{0.6M_\odot}\right)^{-1/2} M_\odot \tag{3.2}$$

on condition of ignoring strong relativistic effects.

Following pericenter passage for single, highly disrupted encounters, approximately half of the material of the star is ejected from the system, whilst the other half remains bound to the BH in Keplerian orbits, subsequently falling back to the pericenter with fallback timescale (Evans & Kochanek, 1989):

$$t_{\rm fb} \simeq \frac{r_{\rm p}^3}{\sqrt{GM_{\rm bh}}R_*^{3/2}} \tag{3.3}$$

where $t_{\rm fb}$ is dependent upon the energy distribution of the bound debris (Rees, 1988; Lodato et al., 2009). The fall back rate onto the black hole is:

$$\dot{M}_{\rm fb} = \frac{1}{3}\frac{M_*}{t_{\rm fb}}\left(\frac{t_{\rm fb}}{t}\right)^{5/3} \tag{3.4}$$

and peaks at:

$$\dot{M}_{\rm peak} \approx \frac{1}{3\delta^{5/3}}\frac{\sqrt{GM_{\rm bh}}R_*^{3/2}}{r_{\rm p}^3}M_*, \tag{3.5}$$

where the value of $\delta$ is dependent upon the type of star. For WDs, $\delta \approx 3.33$ for $M_{\rm wd} > 0.5M_\odot$ (Evans & Kochanek, 1989), and $\delta \approx 5.5$ for $M_{\rm wd} \leq 0.5M_\odot$ (Lodato et al., 2009). $\dot{M}_{\rm peak}$ occurs at a time:

$$t_{\rm peak} = \delta\frac{r_{\rm p}^3}{\sqrt{GM_{\rm bh}}R_*^{3/2}} \tag{3.6}$$

after the disruption (Shcherbakov et al., 2013). The strength of the encounter is typically characterised by the parameter $\beta = r_{\rm t}/r_{\rm p}$.

The bound stellar debris circularizes around the black hole to form an accretion disc post-disruption. Early studies assumed that the accretion rate onto the BH, $\dot{M}_{\rm acc}$, would trace $\dot{M}_{\rm fb}$ if the viscous timescale, $t_{\rm visc}$, is less than the fallback timescale, yet the physical justification for this link is becoming more uncertain (Dai et al., 2013; Hayasaki et al., 2013, 2016; Piran et al., 2015; Shiokawa et al., 2015a; Bonnerot et al., 2016). Significant deviation between $\dot{M}_{\rm fb}$ and $\dot{M}_{\rm acc}$ can occur if there is little energy dissipation per orbit (low viscosity) and material is slow to circularise around the BH (Ramirez-Ruiz & Rosswog, 2009). This is more likely to occur in

WTDEs whereby small self-intersection angles between debris streams (due to smaller general relativistic orbit precession effects), combined with the intersection point being approximately at the apogee, initially leads to low levels of energy dissipation per orbit and inefficient circularisation of debris around the BH (Guillochon & Ramirez-Ruiz, 2015). MacLeod et al. (2016) estimate that this leads to a lowering of $\dot{M}_{\rm acc}$ by a factor of $\sim$10 relative to $\dot{M}_{\rm fb}$ (0.1 term in equation (3.7)).

Initial mass fallback rates are likely to be highly super Eddington (eg. Loeb & Ulmer 1997; MacLeod et al. 2016; Law-Smith et al. 2017a), such that the event is radiatively inefficient and its luminosity output Eddington-limited (Haas et al., 2012; Khabibullin et al., 2014), remaining approximately constant and only starting to decay once accretion rates become sub-Eddington. We note that the process of how falling-back debris is accreted onto the central object is still not well understood, especially for super-Eddington fallback rates. Furthermore, whilst super-Eddington accretion rates persist, the environment may favour jet production. If a jet is launched, it is likely to dominate the power output of the event if viewed on-axis, where the radiation will likely be non-thermal, whereas off-axis emission is likely to be thermal. There is also the possibility of radiatively-driven winds originating from the accretion disc, but we do not consider these here.

Following MacLeod et al. (2014), if $t_{\rm visc} > t_{\rm fb}$, $\dot{M}_{\rm acc}$ no longer traces $\dot{M}_{\rm fb}$, with viscous expansion of the disc modifying accretion timescales according to (Cannizzo & Gehrels, 2009; Cannizzo et al., 2011):

$$\dot{M}_{\rm acc}(t) = 0.1\dot{M}_{\rm peak}\left(\frac{t}{t_0}\right)^{-4/3}, \tag{3.7}$$

where $t_0 = 4/9\alpha^{-1}r_{\rm p}^{3/2}(GM_{\rm bh})^{-1/2}$ and $\alpha = 0.1$ for thick discs. The form of $\dot{M}_{\rm acc}$ during its rise is currently not well constrained; we model it crudely here using a half-Gaussian centred on $t_{\rm peak}$ with $\sigma = t_{\rm peak}/4$, and normalise its peak to $L_{\rm Edd}$. In this work, we assume that the observable X-ray light curve will trace $\dot{M}_{\rm acc}$ and plot examples of our simulated light curves following the above description in Fig. 3.1.

The white dwarf experiences extreme compression perpendicular to the orbital plane, with this being maximised as it crosses pericenter. This is expected to trigger thermonuclear burning if its timescale is much shorter than the dynamical timescale (see eg. Luminet & Pichon 1989a; Brassart & Luminet 2008; Rosswog et al. 2009; Haas et al. 2012). This burning originates in the tidal debris, with emission likely resembling Type Ia SNe-like light curves in the optical band (Rosswog et al., 2008, 2009; MacLeod et al., 2016; Kawana et al., 2018). Whilst thermonuclear burning may affect the amount of bound debris and subsequent accretion rate, no robust models of this currently exist and we neglect its impact in this work.

### Spectral properties

Following the modelling of emission from thick accretion discs (see Balbus & Hawley 1998 for a review), the innermost sections of the disc, $R < 5R_S$ (as in Ulmer 1999 and Khabibullin et al.

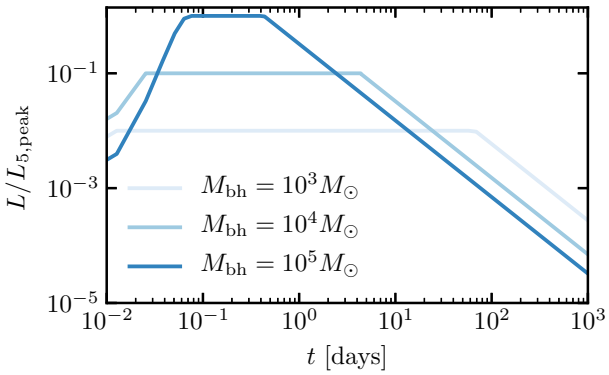


Figure 3.1: Simulated light curves for WTDEs based on the theory described in Section 3.2.1. Higher $M_{\mathrm{bh}}$ have faster rises to peak $L$, with fastest decay rates in light curve for $10^5 M_\odot$ BHs. For comparison of peak luminosities of these events, we normalise each light curve by $L_{5,\mathrm{peak}}$, the Eddington luminosity of a $10^5 M_\odot$ BH.

2014), radiate as a blackbody at temperature:

$$T_{\mathrm{bb}} \simeq \left( \frac{L_{\mathrm{Edd}}}{4\pi\sigma(5R_S)^2} \right)^{1/4} \mathrm{K} \tag{3.8}$$

where $L_{\mathrm{Edd}}$ is the Eddington luminosity of the BH. For IMBHs with masses of $10^3 - 10^5 M_\odot$, this results in X-ray spectra with temperatures $\sim 0.2$keV. The soft emission from these events can be significantly affected by the interstellar medium's (ISM's) absorption along the line of sight. For each TDE, we set the Galactic neutral hydrogen column density, $N_{\mathrm{H}}$, to $5 \times 10^{20}\mathrm{cm}^{-2}$, close to the estimated extra-galactic median of $N_{\mathrm{H}}$ (Esquej et al., 2008a). We model the WTDE's spectrum using the XSPEC model `tbabs(zbbody)`, but neglect modelling any spectral evolution.

An optically-thick extended envelope may form from stellar debris around the black hole, reprocessing a fraction of the X-ray accretion flare into the UV and optical bands (eg. Loeb & Ulmer 1997; Ulmer et al. 1998), with the extent of reprocessing potentially viewing angle dependent (Roth et al., 2016; Dai et al., 2018). As we are interested in exploring eROSITA's sensitivity to the X-ray signatures of WTDEs, we keep to this simplified model of emission from the event and ignore reprocessing in this work (further discussion in section 3.5.2).

### 3.2.2 AGN population

We construct a realistic AGN population based on Section 2 of Clerc et al. (2018). AGN are generated down to several orders of magnitude below the eROSITA flux detection limit for point sources at the end of the four year all-sky survey. To lower SIXTE's computational expense, we then split the generated AGN `simput` file into two parts by selecting those with flux above and below $f_{\mathrm{split}} = 10^{-16}\mathrm{ergs}^{-1}\mathrm{cm}^{-2}$ in the 0.5-2 keV band respectively. AGN with flux above $f_{\mathrm{split}}$ are stored in their own `simput` file. For AGN with flux below $f_{\mathrm{split}}$, we stack the spectra of all sources below this limit and assign this spectrum to a new, artificial extended source of same dimensions

as the original AGN population for the sky patch. This extended source is then placed into a separate `simput` file and models the cosmic X-ray background (CXB) for simulations. Whilst flaring AGNs may represent a large fraction of detected transient sources in eRASS, we do not model AGN time variability in this work and discuss this further in Section 3.5.2.

### 3.2.3 Soft X-ray background

For each sky patch simulated, we determine the diffuse X-ray background due to the SXB and CXB at that position using the *HEASARC Soft X-Ray Background Tool*[4] to obtain the integrated flux and spectrum for the associated sky position. Using XSPEC, we fix $N_{\rm H}$ and fit the spectrum with the model: `apec+wabs(apec+powerlaw)`, obtaining a set of best fitting parameters for each component (`apec` is a model of the X-ray emission of collisionally-ionized diffuse gas, whilst `powerlaw` is an approximate model of the CXB at the given sky position). We then use the best fitting parameter values to construct a `simput` file for the SXB using `apec+wabs(apec)`. Information about the *powerlaw* parameters is discarded as we model the CXB using the stacked AGN below the flux detection limit as described in Section 3.2.2.

## 3.3 eROSITA detection sensitivity

The standard processing of the eROSITA all-sky survey data will use a tiling system that divides the sky into 3.6deg ×3.6deg sky fields. The exposure of each tile over an eRASS is dependent upon its sky position; those near the ecliptic equator have short exposures and are visited only over a single Earth day, whereas those at the poles are visited over consecutive Earth days and have longer exposures. Polar fields therefore present the opportunity to discover lower flux transients, but cover a smaller fraction of the sky.

Given the observational signatures of WTDEs described in Section 3.2.1 and for different sky positions, we are interested in exploring the detection sensitivity of eROSITA to different WTDE parameter configurations. To do this for every eROSITA skyfield would be too computationally expensive; we instead look at this across two different skyfields labelled *2090* (equatorial) and *110135* (intermediate), with exposure times of ∼0.2ks and ∼0.5ks respectively in eRASS1[5].

### 3.3.1 Grid simulations

Each WTDE can be parametrised in terms of $(M_{\rm bh}, M_{\rm wd}, \beta, z, \mathcal{T})$, where $\mathcal{T} = t_{\rm first} - t_{\rm flare}$, with $t_{\rm first}$ being the time when the sky position of that event first enters eROSITA's FOV during eRASS1 and $t_{\rm flare}$ is the time of the disruption. However, the probability of a given event being detected is most dependent on the subset $(M_{\rm bh}, z, \mathcal{T})$. This is due to higher $M_{\rm bh}$ values leading to faster decay of $\dot{M}_{\rm acc}$ and greater peak luminosities since $L_{\rm Edd} \propto M_{\rm bh}$; whilst as $z$ increases, the measured flux in the soft X-ray regime drops due to increasing distance to the event and redshifting of the intrinsically soft spectrum. Positive $\mathcal{T}$ values have the TDE flare occurring before they first enter

[4] `https://heasarc.gsfc.nasa.gov/cgi-bin/Tools/xraybg/xraybg.pl`

[5] Successive eRASS scans will also have similar exposures, but we only consider eRASS1 in this work.

the FOV (and vice versa) – so as $\mathcal{T}$ increases, the flux of the source when first observed decreases. We note that as some skyfields are visited over multiple days during an eRASS, a negative $\mathcal{T}$ does not prevent detection. $M_{\rm wd}$ only ranges between 0.2-1.4$M_\odot$ and modifies $\dot{M}_{\rm peak} \propto M_{\rm wd}$, thus has a smaller impact on the detection efficiency, $\mathcal{D}$, in comparison.

We construct a grid of WTDEs in each skyfield, with spacings between events of 0.8° in RA and Dec. For each skyfield, we consider five different $M_{\rm bh}$ values: $10^3$, $10^{3.5}$, $10^4$, $10^{4.5}$ and $10^5 M_\odot$. $M_{\rm bh}$ is fixed within each grid, and we consider a low and high $z$ set of WTDEs in each skyfield where $z$ increases uniformly between 0.01-0.16 and 0.17-0.32 respectively. For each $M_{\rm bh}$ grid, we vary $\mathcal{T}$ of all WTDEs according to: [-0.75, -0.25, 0, 0.25, 1, 5, 10, 20, 45, 90]d and [-5, -2.5, -1, 0, 1, 5, 10, 20, 45, 90]d for the 2090 and 110135 skyfields respectively. For all simulations, we set $M_{\rm wd} = 0.5 M_\odot$, close to the measured mean mass from SDSS observations (Kepler et al., 2007), but also sufficiently small so that WDs are disrupted instead of being swallowed whole for all simulated $M_{\rm bh}$ values (ie. the condition in equation 3.2 remains valid). We further simplify modelling via fixing $\beta = 1$ for all disruptions, since encounters where $r_{\rm p} \approx r_{\rm t}$ are anticipated to be most common (eg. Rees 1988; Guillochon & Ramirez-Ruiz 2015; Stone & Metzger 2016).

This leads to 200 WTDE grids in total[6], and we then run 10 differently seeded SIXTE simulations of each one over eRASS1. The resulting event files are merged with the simulated event files of the AGN and SXB over eRASS1 (sources in the AGN `simput` are shuffled in position for each seeded run and then simulated). Subsequent source detection and characterisation is performed on the merged event files using the most recent release of the eROSITA analysis software, *eSASS*. Using the software *stilts*[7], we then match the input population of TDEs to the list of eSASS detected sources based on the position estimates. To reduce the number of false matches between input population and detected AGN, we also impose the constraint that eSASS estimated the detected, matched source to have $< 30\%$ and $< 5\%$ of their photon counts between 2-5keV and above 5keV respectively.

### 3.3.2 Detection efficiency

$\mathcal{D}(M_{\rm bh}, \mathcal{T}, z)$ for each configuration of parameters is estimated by computing the fraction of times each WTDE was detected by eSASS across the 10 seeded runs (Fig. 3.3.2). To reduce Poisson noise in estimated $\mathcal{D}(M_{\rm bh}, \mathcal{T}, z)$ due to running only 10 seeded simulations, we smooth $\mathcal{D}$ by applying a Savitzky-Golay filter[8] (Savitzky & Golay, 1964) to each curve, with window length 3 and degree of polynomial of 1.

From Fig. 3.3.2, we see that for both skyfields, events involving higher black hole masses will be detectable out to a larger maximum redshift, $z_{\rm max}$, relative to lower $M_{\rm bh}$. Under these conditions, eROSITA probes each event when $L \approx L_{\rm Edd} \propto M_{\rm bh}$; either due to $\dot{M}_{\rm acc}$ being capped at $\dot{M}_{\rm Edd}$, or $L$ having decayed insignificantly since flaring. This is in contrast to the detectability of TDEs from MS stars with eROSITA, where Khabibullin et al. (2014) estimate $z_{\rm max}$ to be

[6] i.e. (5 different $M_{\rm bh}$) × (2 $z$ ranges) × (10 $\mathcal{T}$ values) for each skyfield

[7] `http://www.star.bris.ac.uk/~mbt/stilts/`

[8] using the `scipy.signal.savgol_filter` implementation.

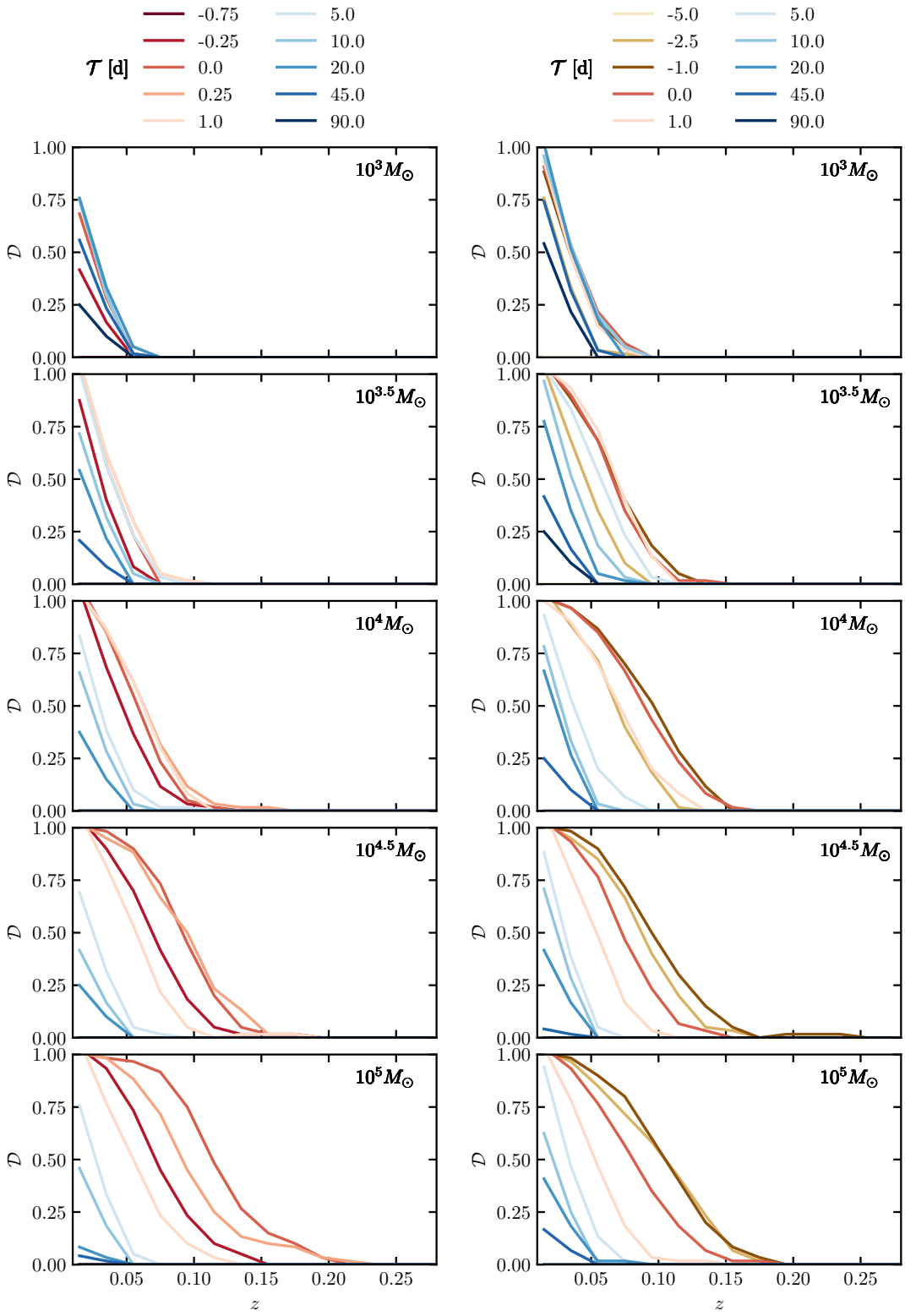


Figure 3.2: Estimated detection efficiency, $\mathcal{D}$, of WTDEs as function of $z$ for different IMBH masses (top-to-bottom) and different $\mathcal{T}$ (note the different color coding). Left and right columns show the equatorial and intermediate skyfields, respectively.

independent of $M_{\rm bh}$. This likely stems from the larger decay timescales of MS TDEs (months to years) causing smaller drops in luminosity as a function of $\mathcal{T}$, such that on average they should be observed with eROSITA much closer to their $L_{\rm Edd}$. The blackbody temperature $T_{\rm bb} \propto M_{\rm bh}^{-1/4}$ (section 3.2.1) for WTDEs has relatively little impact on $\mathcal{D}$, in comparison to $\mathcal{T}$ or $z$.

The longer exposure times of intermediate[9] relative to equatorial skyfields generally boost $\mathcal{D}$ for most $M_{\rm bh}$, though the magnitude of this boost is dependent on $\mathcal{T}$ and $M_{\rm bh}$ (Fig. 3.3.2). For the general case of a constant luminosity source, $\mathcal{D}$ should always be greater for an intermediate compared with equatorial sky field if running detection over the stacked events of an eRASS. However for faster transient sources, Fig. 3.3.2 highlights a more subtle feature of eROSITA's transient detection ability – $\mathcal{D}$ is determined by a trade-off between exposure, average effective area over this exposure, and the luminosity and decay rate of the transient. Across all skyfields, those with longer exposures see a decrease in the rate at which a source located in one moves through eROSITA's FOV. For example, an equatorial field has 6 consecutive visits over an Earth day every 4 hours, whereas an intermediate field has the same frequency of visits but over a $\sim$ 5 day period. The effective area the source is observed with peaks during the middle of this period, when the source passes through the centre of eROSITA's FOV, and decreases either side of this. This means that on the first eROSITA day a source is observed, the average effective area would be greater for the equatorial skyfield relative to the intermediate one, so a source of a given luminosity would have a greater chance of being detected in the equatorial field in this case. If one considers fast decaying transients (such as for the $M_{\rm bh} = 10^5 M_\odot$ WTDEs), the same effect will be present. However in this case, the brightest stage of a WTDE could be observed on average with a lower effective area for the intermediate field compared to an equatorial field. With the source flux decaying so rapidly for these events, by the time it later gets observed in an intermediate field at a higher effective area, its flux has decreased so greatly that the added exposure time has little impact on increasing $\mathcal{D}$. This effect contributes to the observed behaviour of $\mathcal{D}$ curves in Fig. 3.3.2, and will be more noticeable the faster the transient decays (ie. for the more massive IMBHs) and around $\mathcal{T} \sim$ 0d.

Regardless of skyfield (and as expected), our best chance for detection of a WTDE is catching the event close to its flaring, with detection sensitivity rapidly dropping for higher $M_{\rm bh}$ in the following days. For instance, between $\mathcal{T} =$ 1d and $\mathcal{T} =$ 5d for the equatorial skyfield, the redshift at which we detect 50% of events drops from 0.07 to 0.04 for $M_{\rm bh} = 10^5 M_\odot$, but is roughly constant during this time interval at $z \approx$0.04 for $M_{\rm bh} = 10^3 M_\odot$. For $\mathcal{T} >$ 5d, we are only able to detect events up to $z \sim$ 0.05 for all IMBHs. Negative $\mathcal{T}$ values see the WTDE flaring after eROSITA first visits it during an eRASS, with intermediate-like fields presenting the best opportunity to detect negative $\mathcal{T}$ cases. For these, eROSITA will catch the event flaring in the middle of a multi-day visit and thus offers the potential to sample the rise, peak, and decay of its light curve.

[9] These fields will have deeper exposure times as a result of multiple visits of the skyfield over consecutive days, as opposed to only a single day.

Table 3.1: Estimated rates of WTDEs due to elliptical and spiral galaxies computed from the mean of estimated values up to $z = 0.24$ from Fig. 9 of Fragione et al. (2018).

| $M_*$ ($M_\odot$) | $\mathcal{R}_e(M_*)$ (Gpc$^{-3}$yr$^{-1}$) | $\mathcal{R}_s(M_*)$ (Gpc$^{-3}$yr$^{-1}$) |
|---|---|---|
| $10^{10}$ | 18 | 14 |
| $5 \times 10^{10}$ | 6 | 4 |
| $10^{11}$ | 7 | 3 |

## 3.4 Estimate of rate

In this section, we will first derive an updated estimate on the rate density of WTDEs in the local universe, before using this along with the inferred detection sensitivities to estimate eROSITA's detection rate of these events.

### 3.4.1 Intrinsic rate of WTDEs

Following Shcherbakov et al. (2013), we assume that the IMBH population in the local universe mainly resides in globular clusters (GCs) and dwarf galaxies (DGs), with these two being the main contributors to the WTDE rate density[10], $\mathcal{R}$. Furthermore, we assume that GCs host BHs with $10^3 M_\odot < M_{bh} < 10^4 M_\odot$, and DGs those with $M_{bh} > 10^4 M_\odot$.

The contribution to $\mathcal{R}$ from GCs is estimated based on Fragione et al. (2018), which uses a semi-analytic model to evolve a primordial population of globular clusters within massive host galaxies over cosmic time. The evolution takes into account mass loss via stellar winds, the loss of stars through evaporation, two body interactions and tidal stripping by the host galaxy. Furthermore, they also include interactions between the IMBH and stellar mass black holes that can eject the IMBH from its GC due to recoil from asymmetric gravitational wave emission. Combined with estimates of the local number density of globular clusters in Rodriguez et al. (2015), they present estimates of the WTDE rate density in the local universe from elliptical and spiral galaxies for three different $M_*$.

To estimate the total rate density from the estimates for specific host galaxy masses in table 3.1, we sum the weighted mean contribution from elliptical and spiral hosts:

$$\mathcal{R}_{GC} = \frac{\sum_i \Phi(M_{*,i})\mathcal{R}_e(M_{*,i})}{\sum_i \Phi(M_{*,i})} + \frac{\sum_i \Phi(M_{*,i})\mathcal{R}_s(M_{*,i})}{\sum_i \Phi(M_{*,i})} \tag{3.9}$$

where $\Phi(M_{*,i})$ are computed using the Schechter function (Schechter, 1976):

$$\Phi(M_*) = \Phi_c \left(\frac{M_*}{M_c}\right)^{-\alpha_c} e^{-M_*/M_c}, \tag{3.10}$$

with $\Phi_c$ a normalising constant, $M_c = 10^{11.14} M_\odot$ and $\alpha_c = 1.43$, using parameter estimates from EAGLE cosmological simulations (Furlong et al., 2015). Combining equation 3.9 and data from

[10] the number of WTDEs per year per Gpc$^3$

table 3.1 yields a local rate density from GCs of $\sim$ 30 $\mathrm{Gpc^{-3}yr^{-1}}$ (17 and 13 $\mathrm{Gpc^{-3}yr^{-1}}$ from elliptical and spiral galaxies respectively).

We then assume the dwarf galaxy contribution to $\mathcal{R}$ has the form:

$$\mathcal{R}_{\mathrm{dg}} = f_{\mathrm{oc}} n_{\mathrm{dg}} \dot{N}_{\mathrm{dg}} \tag{3.11}$$

where $f_{\mathrm{oc}} \sim 0.1$ is the estimated occupation fraction in DGs of BHs with $M_{\mathrm{bh}} < 10^5 M_\odot$ (Alves Batista & Silk, 2017), although we note that Miller et al. (2015) estimate a lower limit on $f_{\mathrm{oc}}$ at 95% confidence to be $\sim$0.2 for galaxies with $M_* < 10^{10} M_\odot$. Our resulting estimated WTDE detection rates will likely be conservative with respect to $f_{\mathrm{oc}}$. $\dot{N}_{\mathrm{dg}} \sim 10^{-6}\,\mathrm{yr^{-1}}$ is the estimated rate of WTDEs per IMBH computed in MacLeod et al. (2014) and $n_{\mathrm{dg}}$ is the local number density of dwarf galaxies. The latter is estimated via integration of a double Schechter function with best fitting parameters from Blanton et al. (2005), obtained via their fitting of the luminosity function of extremely low luminosity galaxies observed with the SDSS (r-band), corrected for selection effects. Over the range of this integral (bounds defined below), it is assumed that only dwarf galaxies contribute to the number density.

There is currently no accepted, clear distinction in the literature between dwarf and non-dwarf galaxies. However for the purpose of estimating $n_{\mathrm{dg}}$, this is only a minor issue since $n_{\mathrm{dg}}$ is far less sensitive to the choice of upper bound compared with the lower. A coarse cut-off has previously been considered at $M_r \approx -18$ (also approximately the magnitude of the Large Magellanic Cloud; McConnachie 2012), we choose this as the upper bound of this integral. We then consider two different lower bounds herein. The first is at $M_r = -6$, which represents an extrapolation of the best fit Schechter function in Blanton et al. (2005) down to the lowest luminosity dwarf galaxies observed outside the Milky Way (see for example McConnachie 2012). This also roughly corresponds to requiring a minimum of $\sim 10^6 M_\odot$ contained within the half-light radius of the DG, based on Fig. 11 of Torrealba et al. (2019). The second choice, at $M_r = -12$, is a more conservative estimate, which is approximately the lowest value of the $M_r$ data used for model fitting in Blanton et al. (2005). These two different lower bounds yield estimated $n_{\mathrm{dg}}$ of 2.6 and $0.3h^3\mathrm{Mpc^{-3}}$ respectively, with the former being approximately an order of magnitude greater than a previous lower limit for $n_{\mathrm{dg}}$ of $0.12h^3\mathrm{Mpc^{-3}}$ estimated in Loveday (1997). However, the authors of that work suggest that their reported $n_{\mathrm{dg}}$ is likely to be significantly lower than the true value due to the various assumptions they make on galaxy clustering that lead to underestimation of $n_{\mathrm{dg}}$.

Adopting $h = 0.678$ from Planck Collaboration XIII (2016), this leads to our rate density estimates from BHs in DGs with $10^4 M_\odot < M_{\mathrm{bh}} < 10^5 M_\odot$ of:

$$\mathcal{R}_{\mathrm{dg}} \approx \begin{cases} 10\,\mathrm{Gpc^{-3}yr^{-1}} & \text{if } n_{\mathrm{dg}} = 0.3h^3\mathrm{Mpc^{-3}},\ f_{\mathrm{oc}} = 0.1 \\ 80\,\mathrm{Gpc^{-3}yr^{-1}} & \text{if } n_{\mathrm{dg}} = 2.6h^3\mathrm{Mpc^{-3}},\ f_{\mathrm{oc}} = 0.1. \end{cases}$$

The DG contribution to $\mathcal{R}_{\mathrm{dg}}$ is expressed conditional on our adopted $n_{\mathrm{dg}}$ and $f_{\mathrm{oc}}$ values; this highlights that if, for example, $f_{\mathrm{occ}}$ were to double, then we would expect that $\mathcal{R}_{\mathrm{dg}}$ would rise to 160 $\mathrm{Gpc^{-3}yr^{-1}}$ for the case where $n_{\mathrm{dg}} = 2.6h^3\mathrm{Mpc^{-3}}$.

Similar to previous studies, we find that dwarf galaxies are the dominant source of WTDEs, although even our highest $\mathcal{R}_{\mathrm{dg}}$ estimate is roughly a factor of 20 lower than in Shcherbakov et al.

(2013). The main cause of this disagreement is that we use $\dot{N}_{\rm dg} \sim 10^{-6}\,{\rm yr}^{-1}$, compared with their $0.15 \times 10^{-3}\,{\rm yr}^{-1}$ (they assume a rate of stellar tidal disruption of $10^{-3}\,{\rm yr}^{-1}$, and 15% of these events involve white dwarfs) which bears the problem of BH overgrowth; the IMBHs would rapidly grow above $10^5 M_\odot$ via WTDEs alone if those rates were sustained. Furthermore, they assume that all dwarf galaxies will host an IMBH with mass low enough for a valid WTDE, whereas we estimate 10% of BHs in DGs will have $M_{\rm bh} < 10^5 M_\odot$.

### 3.4.2 eROSITA detection rate

For eRASS1, we consider the set of WTDEs that flare with $-90{\rm d} < t_{\rm flare} < 180{\rm d}$ (relative to start of eRASS1), and within the spherical volume enclosed by $z < 0.24$, as potentially detectable events. This range is chosen based on the $\mathcal{D}$ curves in Fig. 3.3.2, where $\mathcal{D}(M_{\rm bh}, \mathcal{T} > 90{\rm d}, z > 0.24) \sim 0$. Using our WTDE rate density estimates ($\mathcal{R}$), this volume and period corresponds to $N_{\rm tot,gc} = 180$ and $N_{\rm tot,dg} = 470$ (40) events from globular clusters and dwarf galaxies respectively (bracketed number for the case of $n_{\rm dg} = 0.3h^3{\rm Mpc}^{-3}$). The number of WTDE detections during each eRASS is then estimated by drawing $N_{\rm tot,gc}$ and $N_{\rm tot,dg}$ random WTDE configurations and computing the number of these that are detected (further described below). We then repeat this 1000 times, replicating 1000 eRASS scans of each host class, to study the distribution of the expected number of detections.

We consider the sky to be divided into two parts which have exposure above and below the ~0.5ks exposure of an intermediate skyfield, occupying 90% and 10% of the sky each. These weights are obtained from the estimated exposure map predicted for eROSITA's All-Sky Survey (J. Robrade, private communication). eROSITA's detection sensitivity to WTDEs in each region is then modelled using the estimated $\mathcal{D}$ for the 2090 and 110135 skyfields respectively. The number of detections in eRASS1 from globular clusters, $N_{\rm gc}$, is estimated via:

$$N_{\rm gc} = N_{\rm gc,\,eq.} + N_{\rm gc,\,int.} \tag{3.12}$$

where $N_{\rm gc,\,eq.}$ ($N_{\rm gc,\,int.}$) is 90% (10%) of the estimated number of detections over the whole sky when assuming a detection sensitivity of the equatorial (intermediate) field. Similarly, the number of WTDE detections from dwarf galaxies is:

$$N_{\rm dg} = N_{\rm dg,\,eq.} + N_{\rm dg,\,int.}, \tag{3.13}$$

where definitions follow those for equation 3.12.

Our simulations yield $\mathcal{D}(M_{\rm bh}, \mathcal{T}, z)$ in each skyfield. To cast this into a form easier to estimate the rate of WTDE detection, we infer eROSITA's detection sensitivity to all WTDEs occurring up to $z_{\rm max}$ as a function of $\mathcal{T}$, via marginalising over $z$:

$$\mathcal{D}(M_{\rm bh}, \mathcal{T}) = \int \mathcal{D}(M_{\rm bh}, \mathcal{T}, z) p(z) {\rm d}z \tag{3.14}$$

where $p(z)$ is the expected $z$ distribution of WTDEs. To estimate this (due to lack of observational evidence of WTDEs), we split up the local universe into a series of nested spherical shells and assume no significant evolution in the TDE hosts within this volume. Setting each shell to have

the same event number density, $n$, in comoving space, then the number of events per shell, d$N$, is expected to be:

$$\mathrm{d}N = n\mathrm{d}V_{\mathrm{sh}}(z) \tag{3.15}$$

where $\mathrm{d}V_{\mathrm{sh}}(z)$ is the comoving volume of the shell, with probability of finding an event at redshift z being:

$$p(z) \approx \frac{V_{\mathrm{sh}}(z)}{V_{\mathrm{max}}} \sim \frac{4\pi z^2}{\frac{4}{3}\pi z_{\mathrm{max}}^3} = Az^2. \tag{3.16}$$

with $A$ is defined such that $\int p(z)\mathrm{d}z = 1$.

After marginalisation, only $M_{\mathrm{bh}}$ and $\mathcal{T}$ require drawing to evaluate $\mathcal{D}$ for a particular WTDE configuration. $M_{\mathrm{bh}}$ is randomly chosen from $[10^3, 10^{3.5}, 10^4]M_\odot$ for globular clusters, and $[10^4, 10^{4.5}, 10^5]M_\odot$ for dwarf galaxies (effectively crudely sampled from a log-uniform distribution). A $\mathcal{T}$ sample for each WTDE is obtained by combining randomly drawn $t_{\mathrm{flare}}$ and $t_{\mathrm{first}}$ values, where $t_{\mathrm{flare}}$ is drawn from $\mathcal{U}(-90\mathrm{d}, 180\mathrm{d})$. For drawing samples of $t_{\mathrm{first}}$, we generate a random sky position and compute the time a source located there first enters eROSITA's FoV[11] in eRASS1 (using the SIXTE task `ero_vis` and a spacecraft attitude file). Random source positions are drawn by assuming WTDEs are equally likely to occur at all points on the sky, and are drawn randomly from the surface of a sphere via normalising a set of three random numbers generated from $\mathcal{U}$ $(0, 1)$ (Marsaglia, 1972). Finally, we draw $M_{\mathrm{bh}}$ from the range $10^3 - 10^4 M_\odot$ and $10^4 - 10^5 M_\odot$ for WTDEs in globular clusters and dwarf galaxies, respectively.

For each drawn $(M_{\mathrm{bh}}, \mathcal{T})$, we estimate the probability of detecting its associated WTDE, $\mathcal{D}$, via linear interpolation of equation 3.14. We generate a random number, $x$, from $\mathcal{U}(0, 1)$. If $x < \mathcal{D}$, then we classify the event as detected; otherwise it is a non-detection. We do this for each of the $N_{\mathrm{tot,gc}}$ and $N_{\mathrm{tot,dg}}$ detectable WTDEs in an eRASS, and repeat this process 1000 times. From this, we construct estimates for the number of WTDE detections per eRASS for the globular cluster and dwarf galaxy populations, with the distributions for the number of detections per eRASS presented in Fig. 3.3.

Scaling estimates for eRASS1 up to eRASS8, we estimate that over its 4 year all-sky survey, eROSITA may detect ∼2 WTDEs from dwarf galaxies (for $n_{\mathrm{dg}} = 2.6h^3\mathrm{Mpc}^{-3}$) and ∼1 from globular clusters. On the other hand, if the more conservative estimate for $n_{\mathrm{dg}}$ is more accurate, then we expect no detections from dwarf galaxies.

As a byproduct of this set of Monte Carlo simulations, we can also infer the $\mathcal{T}$ distributions of WTDEs classified as detected for different black hole masses and different sky positions (Fig. 3.4). For $M_{\mathrm{bh}} > 10^4 M_\odot$, nearly all detected events are within 5 days of their flaring, whereas $M_{\mathrm{bh}} < 10^4 M_\odot$ events have much broader $\mathcal{T}$ distributions. Transitioning to fields with deeper exposure shifts the peak of the distribution towards lower $\mathcal{T}$.

## 3.5 Discussion

We have simulated eROSITA observations of WTDEs and estimated the number of detections during eROSITA's four year all-sky survey. However, a detection is different to identification

[11] 1.02 deg diameter

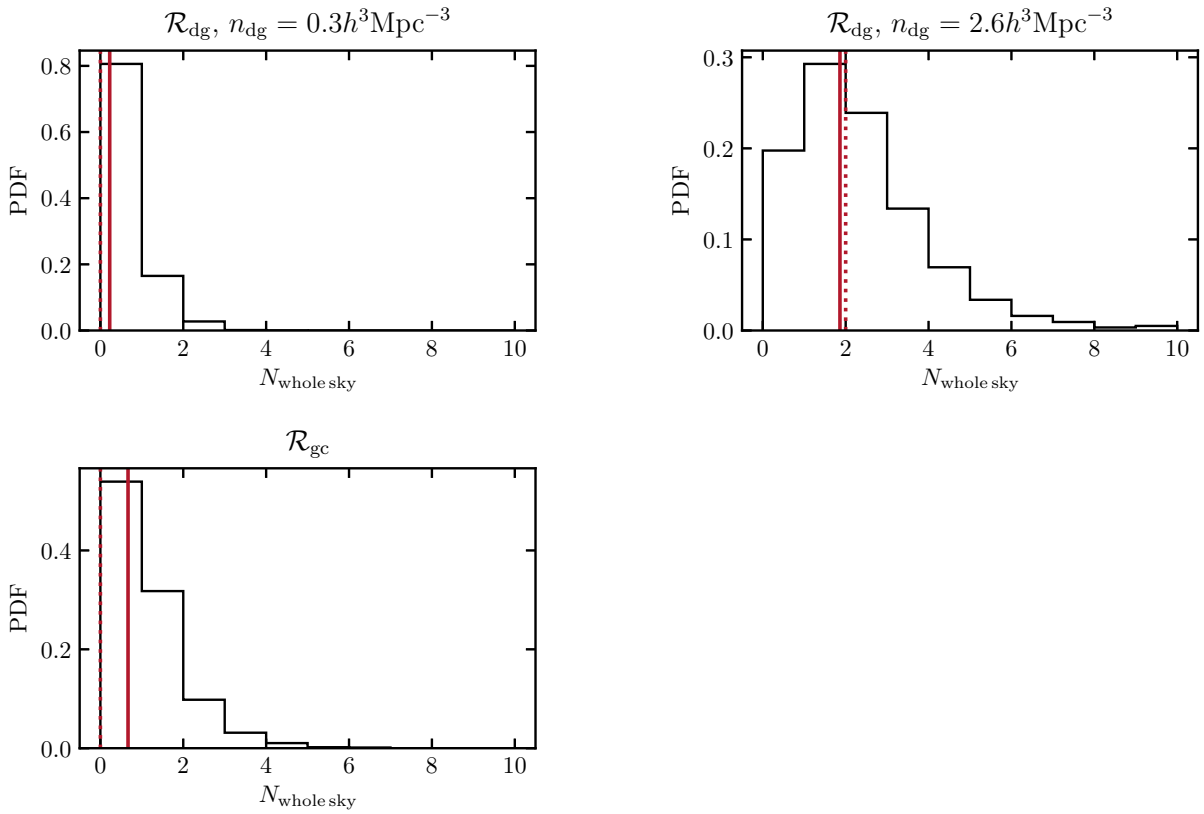


Figure 3.3: Distributions of $N_{\rm whole\ sky}$ – the expected number of detections up to eRASS8 over the whole sky, from dwarf galaxies (top two panels) and globular clusters (bottom). Dotted and solid vertical lines mark the median and mean of samples from the distribution.

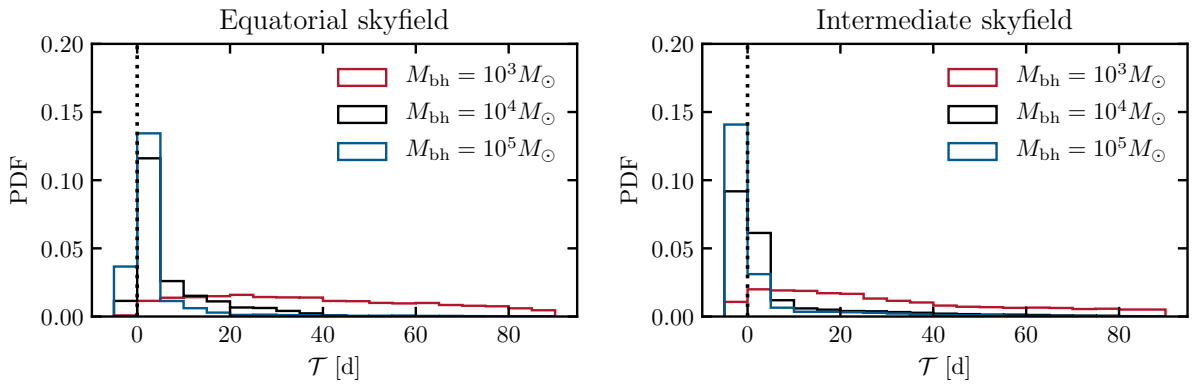


Figure 3.4: Inferred $\mathcal{T}$ distribution for detected WTDEs for different black hole masses, with top and bottom panels for the equatorial and intermediate skyfields respectively. Events detected with $\mathcal{T} < 0$ (left of the black dotted vertical line) will flare after eROSITA has visited it for the first time during an eRASS. For the deeper exposure fields, the source will be scanned multiple times over consecutive days. In agreement with inferences drawn from Fig. 3.3.2, as $M_{\rm bh}$ increases, $\mathcal{D}$ drops off more rapidly with $\mathcal{T}$.

– for the latter, we are confident in our classification of the object type. Prior to building up a sample of observed WTDEs, one must be able to pick out and identify the WTDEs hidden in eROSITA's source catalogues. In the following section, we first consider predicted multi-wavelength signatures for these events that could be used to assist mining the large eROSITA datasets. We then briefly discuss the caveats of our assumptions made in inferring eROSITA's detection efficiency that may alter our detection rate estimates.

### 3.5.1 Identification and multi-$\lambda$ signatures

WTDEs will be non-detectable at X-ray wavelengths in the subsequent eRASS sky pass 6 months after first detection (eg. Fig. 3.3.2) – for most eROSITA detections we may only have an X-ray light curve constructed from its ≈40s visits. From this sampling alone, it will be challenging to accurately classify a transient as a WTDE, or identify a candidate which requires further multi-wavelength followup. To overcome this and greatly reduce the number of false-positive WTDE identifications, it is vital to consider potential observational signatures of their hosts that already exist in the wealth of multi-wavelength sky surveys available.

The majority of new X-ray variable sources detected per eRASS will likely be flaring stars and AGN. To reduce this set of newly detected flaring sources to a smaller group of WTDE candidates, an initial cross-match with multi-wavelength surveys should be performed. For instance, one could discard transients associated with centres of known massive galaxies as their BHs will exceed the mass limit for successfully disrupting a WD. Similarly, stellar counterparts with significant proper motion measured by *Gaia* could be rejected. We note that based on the extent of theoretical uncertainties and lack of significant observational evidence surrounding WTDEs, we anticipate selection of these events will be particularly difficult during eRASS1. However by

eRASS2, we will have a much better understanding of the variable X-ray sky, thus allowing for a cleaner selection of the new flaring sources in following eRASS. Of the new X-ray transients that remain, additional information will be needed to support classification of the event as a WTDE. For instance, optical follow-up of a candidate may allow us to ascertain the nature of its host or detect an associated optical transient, whilst obtaining an estimated $z$ for the event would allow us to infer its luminosity and rule out certain classes of variable X-ray sources.

An estimate of the mass of the black hole involved will provide strong support for classification of a candidate transient as a WTDE (providing the LC shows anticipated WTDE-like behaviour). Recently, a tidal disruption event module (Mockler et al., 2019) has been developed for MOSFiT (Guillochon et al., 2018), an open source Python-based code for Bayesian parameter estimation of astronomical transients based on semi-analytical fitting of their multi-wavelength light curves. This extension module provides estimates of $M_{\rm bh}$ (amongst other parameters) involved in MS TDEs. However, a module based on MS TDEs may not be appropriate to use for WTDEs, since the uncertainty surrounding their X-ray emission (in particular, the extent of super-Eddington accretion these events go through at flaring) may significantly bias any parameter estimates constructed from inferred posterior distributions. A further MOSFiT module specialised to WTDEs may require development (potentially in light of eROSITA's discoveries). We note our ability to constrain $M_{\rm bh}$ from the X-ray light curve data will also be limited by eROSITA's time sampling; the most accurate parameter estimates will likely follow from candidates that we can track for consecutive days in eROSITA's polar fields, or for which follow-up X-ray coverage with other missions can be compiled.

An alternate way to estimate $M_{\rm bh}$ takes advantage of the majority of WTDE hosts expected to be DGs. For DGs with bulge mass estimates available, we may also crudely estimate $M_{\rm bh}$ from extrapolation of the $M_{\rm bh} - M_*$ relation (McConnell & Ma, 2013). Since DGs are typically low luminosity (and hard to identify in surveys), and assuming that the majority of DGs are satellites of more massive galaxies, then WTDEs will be observed as off-galactic centre transients. If a new, flaring X-ray source originating from a DG is detected, deeper follow-up observations with either *Chandra* or *XMM-Newton* would provide a higher quality X-ray spectrum than eROSITA and be very useful for constraining the X-ray emission processes for the event (for example, are there components to the X-ray spectrum other than a black-body accretion disc?). Such follow-up would also provide better sampling of the X-ray light curve, allowing us to study the time evolution of $\dot{M}_{\rm acc}$ for these systems and better understand the physical processes involved in the WTDE itself.

MS TDEs will significantly outnumber WTDEs across the DG population. However, it should be possible to distinguish between these two TDE classes as their LCs are expected to differ significantly within the IMBH mass range. As a basic comparison of LC behaviour, we consider a TDE for two cases: i) a $1M_\odot$ WD with $R_{\rm wd} = 0.01R_\odot$, and ii) a $1M_\odot$ MS star with $R_{\rm ms} = 1R_\odot$. For a configuration where $10^5 M_\odot$ BH with $\beta = 1$, $t_{\rm fb}$ scales $\propto R_*^{3/2}$ (equation 4.2) such that the decay timescale for the MS TDE is a factor of 1000 greater than for the WTDE (and similarly also for $t_{\rm peak}$). The slowing of MS TDEs from BHs $< 10^6 M_\odot$ is also found in Guillochon & Ramirez-Ruiz (2015), whereby these events typically evolve over several years due to highly inefficient debris circularisation and large viscous timescales for these systems. This is in agreement with a recent potential MS TDE identification from an IMBH (Lin et al., 2018), which

was found to decay over a decade. We also highlight here that eROSITA should identify a set of MS TDE candidates from IMBHs that will require subsequent X-ray follow-up observations to track LC decay after its all-sky survey finishes.

Approximately one in every six WD TDEs could result in a thermonuclear transient (MacLeod et al., 2016). Their optical light curves are anticipated to be Ia-like, but may differ by possessing shorter rise times (10-12 days), maximum luminosities $\sim 8 \times 10^{42}$erg s$^{-1}$ (dimmer than Ia) and single-peaked in the near-infrared bands, unlike standard Ias (eg. Kasen 2006). These events may already have been observed as the calcium-rich gap transients (Sell et al., 2015), though no significant X-ray counterpart has been found for these. eROSITA's All-Sky Survey should coincide with wide area sky scans such as the Zwicky Transient Facility (Bellm, 2014); finding a reported Ia-like supernova as a counterpart to an observed X-ray flare will support WTDE classification and help determine an origin of these Ca transients. MacLeod et al. (2016) predict optical spectra for these Ia-like events with expected signatures: Doppler shifts $\sim 10^4$kms$^{-1}$, P-Cygni lines, intermediate mass element production and a high dependence on viewing angle. For WTDEs that do not trigger thermonuclear burning, we anticipate optical spectra (based on MS TDEs observations eg. Gezari et al. 2012; Arcavi et al. 2014) to be mainly black body with different broadened emission lines depending on the composition of the WD being disrupted. For example, for WDs below $0.5M_{\odot}$ we would expect broadened He emission lines to dominate the optical spectrum.

### 3.5.2 Uncertainties in WTDE rate estimates

#### TDE modeling

As discussed in Section 3.2.1, mass fallback rates at early times post-disruption are predicted to be highly super-Eddington. We modelled emission during this period as Eddington-limited, such that the luminosity of the event will be determined by $M_{\rm bh}$ and approximately constant until sub-Eddington accretion rates are achieved. Several other classes of variable/ transient sources have been observed to be undergoing super-Eddington accretion (eg. ultra-luminous X-ray sources, Walton et al. 2013; and AGN, Middleton et al. 2011). However, the main factors that allow accretion to be super-Eddington are currently unclear. Evans et al. (2015) model accretion of the debris of a main sequence star onto an IMBH post-disruption. An initial phase of hyper-accretion is expected from accretion of portions of the star that falls directly onto the BH, followed by an approximately constant accretion rate until the debris is fully consumed[12]. If WTDEs are capable of attaining super-Eddington accretion rates, it is likely they will be observed with eROSITA as a set of fast decaying transients detectable to higher $z_{\rm max}$ anticipated in this work (if eROSITA can catch them very close to peak flaring).

In addition both WTDEs and MS TDEs are predicted to launch relativistic jets, with the potential to reach much higher luminosities. Furthermore, their light curves may fully trace the fallback rate over time, as opposed to Eddington limited phases such as in Fig. 3.1. These should be detectable to much higher $z$ than their non-jetted counterparts and boost the detection rate of

[12]An instance of this hyper-accretion behaviour could be in Jonker et al. (2013), where the light curve of the candidate WTDE shows very similar behaviour to the accretion rate curves plotted in Fig. 2 of Evans et al. (2015).

WTDEs. We anticipate jetted WTDEs may be distinguished from their non-jetted counterparts via their harder X-ray spectra from Comptonisation of the quasi-blackbody radiation from the jet's photosphere (although our ability to distinguish between these two spectral models may be limited in cases of low source count rates) (Shcherbakov et al., 2013).

On the other hand, the WTDE detection rate may be lowered by reprocessing of X-rays by optically thick outflows, predicted to be launched for cases of super Eddington fallback rates. Dai et al. (2018) propose a unified model of TDEs, whereby the viewing angle of the TDE affects whether one observes optical or X-ray dominated emission due to the angular dependence of the outflow's density. When viewed face on, one probes X-ray emission from the inner accretion disc, whereas a larger amount of X-ray radiation is reprocessed (due to increasing photoelectric absorption in the outflow density) with increasing viewing angle. Whilst the authors study super-Eddington accretion in MS TDEs, it is reasonable to assume that a similar viewing-angle dependence may also be present for WTDEs.

For simplicity, we have neglected black hole spin in this work. Introducing spin modifies the innermost stable orbit (ISCO) around the BH; larger spins lead to smaller ISCO radii such that the tidal disruption radius lies outside the ISCO and tidal disruption flares are possible (as opposed to the star being swallowed whole). Thus for a given WD, the maximum $M_{\rm bh}$ that could tidally disrupt it is greater, allowing black holes up to $10^6 M_\odot$ to disrupt WDs. This could represent an increase in the number of BHs with potential to fully disrupt WDs, depending on the net alignment of black hole spins in the local universe.

### Simulating eROSITA observations

The simulated source populations in this work have only included WTDEs and AGN. Since unresolved AGN populations are expected to contribute most to the CXB (eg. Lehmer et al. 2012), realistically modeling the AGN component is necessary for accurate estimates of source detection efficiency. We do not expect exclusion of other X-ray source populations, such as stars and galaxy clusters, to significantly affect our detection efficiency estimates. Furthermore, galaxy clusters in the redshift range for which we are WTDE-sensitive should be clearly identified as extended sources. We also note that galaxy clusters may provide a good place to search for WTDEs due to their large assembly of stellar mass, including GCs and DGs, in their dark matter potentials.

In addition, we have not included any other sources of X-ray variability that will be observed with eROSITA. Whilst inclusion of variability might allow for a rough quantification of the false positive rate for WTDE detection, no robust models currently exist for simulating synthetic X-ray light curves of all classes of variable X-ray sources. Even if these were available, it would still be a highly non-trivial task to robustly quantify the false positive rate. Our approach allows us to isolate the issue of how frequently eROSITA will detect WTDEs, from the broader problem of how many WTDEs can be identified amongst eROSITA's variable source population (since eROSITA's sparse time sampling introduces further complication to variable source classification). We have reported estimated detection rates of WTDEs where detection is counted as the source being detected by eSASS, but note the number of identified WTDEs will differ from the estimated detection rate.

## 3.6 Conclusions

We have simulated an extensive set of eROSITA observations during its all-sky survey of WTDEs, incorporating a realistic, non-variable X-ray background consisting of a CXB, SXB and a particle background. eROSITA's detection sensitivity to WTDEs as a function of black hole mass, redshift and $\mathcal{T}$ was then inferred. We estimated a novel rate density for WTDEs from BHs with masses between $10^3$ and $10^5 M_\odot$, which was then combined with the detection sensitivities to estimate the rate of eROSITA detecting WTDEs. By the end of its 4 year all-sky survey, eROSITA should be able to probe a sample of $\sim$ 3 quiescent intermediate mass black holes involved in WTDEs, if the estimated luminosity function of low luminosity galaxies can be extended down to the lowest luminosity dwarf galaxies observed outside the Milky Way. Due to the higher WTDE rate densities anticipated for dwarf galaxies, we expect eROSITA to be more sensitive to detecting disruptions involving black holes with masses $> 10^4 M_\odot$. Most detected WTDEs will be found within a few days of flaring, and detectable up to $z < 0.24$. In addition, we have explored and demonstrated the usefulness of the SIXTE simulator in forecasting transient detection abilities of future X-ray surveys.

# Chapter 4

# AT 2019avd: a novel addition to the diverse population of nuclear transients

In this chapter, we report on *SRG*/eROSITA, ZTF, ASAS-SN, Las Cumbres, NEOWISE-R, and *Swift* XRT/UVOT observations of the unique ongoing event AT 2019avd, located in the nucleus of a previously inactive galaxy at $z = 0.029$. eROSITA first observed AT 2019avd on 2020-04-28 during its first all sky survey, when it was detected as an ultra-soft X-ray source ($kT \sim 85$ eV) that was $\gtrsim 90$ times brighter in the $0.2 - 2$ keV band than a previous $3\sigma$ upper flux detection limit (with no archival X-ray detection at this position). The ZTF optical light curve in the $\sim 450$ days preceding the eROSITA detection is double peaked, and the eROSITA detection coincides with the rise of the second peak. Follow-up optical spectroscopy shows the emergence of a Bowen fluorescence feature and high-ionisation coronal lines ([Fe X] 6375 Å, [Fe XIV] 5303 Å), along with persistent broad Balmer emission lines (FWHM$\sim$ 1400 km s$^{-1}$). Whilst the X-ray properties make AT 2019avd a promising tidal disruption event (TDE) candidate, the optical properties are atypical for optically selected TDEs. We discuss potential alternative origins that could explain the observed properties of AT 2019avd, such as a stellar binary TDE candidate, or a TDE involving a super massive black hole binary.

The contents of this chapter were originally published in Malyali et al. (2021).

## 4.1 Introduction

Actively accreting supermassive black holes (SMBHs) have long been known to exhibit large amplitude flaring behaviour (e.g. Tohline & Osterbrock 1976; Antonucci & Cohen 1983; Penston & Pérez 1984; Shappee et al. 2014; Storchi-Bergmann et al. 2017; Frederick et al. 2019), whereby multi-epoch observations of galaxy nuclei, over year-long timescales, have revealed drastic changes in their luminosity. The physical mechanisms responsible for producing extreme accretion rate changes are still unclear, although various models have been suggested, such as state transitions in the inner disc (Noda & Done, 2018; Ross et al., 2018), radiation pressure instabilities in the disc (Śniegowska & Czerny, 2019), or tidal disruption events (TDEs; Merloni et al. 2015; Chan et al. 2019).

Whilst the sample of ignition events in galactic nuclei was previously limited to only a few objects, the advance of wide-field, high-cadence surveys over the last decade has facilitated the discovery of an increasing number of extreme state changes. This has resulted in tighter constraints on the timescales of flaring events for these systems. For example, Trakhtenbrot et al. (2019a) recently reported a new class of SMBH accretion event that sees a large amplitude rise in the optical/UV luminosity over timescales of months.

In addition to triggering drastic changes in the accretion rate in AGNs, TDEs can also cause quiescent black holes to transition into short-lived active phases. In a TDE, a star that passes too close to a BH is torn apart by strong tidal forces, with a fraction of the bound stellar debris then being accreted onto the BH (Hills, 1975; Young et al., 1977; Gurzadian & Ozernoi, 1981; Lacy et al., 1982; Rees, 1988; Phinney, 1989). Early TDE candidates were first identified through detection of large-amplitude (at least a factor of 20), ultra-soft X-ray flares (black-body temperatures between 40 and 100 eV) from quiescent galaxies during the ROSAT survey (Bade et al., 1996; Komossa & Bade, 1999; Komossa & Greiner, 1999; Grupe & Leighly, 1999; Greiner et al., 2000). Since then, the vast majority of TDE candidates have been optically selected, such as through the Sloan Digital Sky Survey (SDSS; e.g. van Velzen et al. 2011; Merloni et al. 2015), the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS; e.g. Gezari et al. 2012; Holoien et al. 2019a), the Palomar Transient Factory (PTF; e.g. Arcavi et al. 2014), the Intermediate Palomar Transient Factory (iPTF; e.g. Blagorodnova et al. 2017; Hung et al. 2017), the All Sky Automated Survey for SuperNovae (ASAS-SN; e.g. Holoien et al. 2014, 2016; Wevers et al. 2019; Holoien et al. 2019b), and the Zwicky Transient Facility (ZTF; e.g. van Velzen et al. 2019, 2021). Optically selected TDEs are characterised as blue nuclear transients with light curves showing longer/ shorter rise and decay timescales relative to supernovae (SNe)/ AGN[1], and a relatively smooth power-law decline. Optical spectroscopic follow-up of these events post-peak reveals blue continua (blackbody temperatures $\sim 10^4$K) with various broad emission lines (full width at half maximum, FWHM $\lesssim 10^4$ km s$^{-1}$); a recent characterisation of the different TDE spectroscopic classes was presented by van Velzen et al. (2021). Although a number of TDE candidates have also been found through UV selection (*GALEX*, Gezari et al. 2008, 2009), and X-ray selection (*XMM-Newton* Slew, Esquej et al. 2007, 2008b; Saxton et al. 2012, 2017), most of our understanding of TDEs is currently biased towards this set of observed properties of optically-selected TDEs.

Whilst most previous TDE searches focused on identifying TDEs in quiescent galaxies, an increasing number of candidates for TDEs in AGNs are being proposed in the literature (Merloni et al., 2015; Blanchard et al., 2017; Trakhtenbrot et al., 2019b; Liu et al., 2020; Ricci et al., 2020). In certain cases, the distinction between TDE and non-TDE-induced SMBH accretion state changes is becoming increasingly blurred (see also Neustadt et al. 2020). Variants of TDEs have also been proposed to explain more exotic phenomena, such as the recently observed quasi-periodic eruptions (QPEs) in a few galactic nuclei (Miniutti et al., 2019; Giustini et al., 2020; King, 2020), and periodic flaring seen in an AGN (Payne et al., 2021). Other origins for extreme nuclear transients involve SNe in the AGN accretion disc (Rozyczka et al., 1995), or interaction

[1]For large, well-defined AGN flares similar to those seen in Frederick et al. (2019), as opposed to stochastic AGN variability.

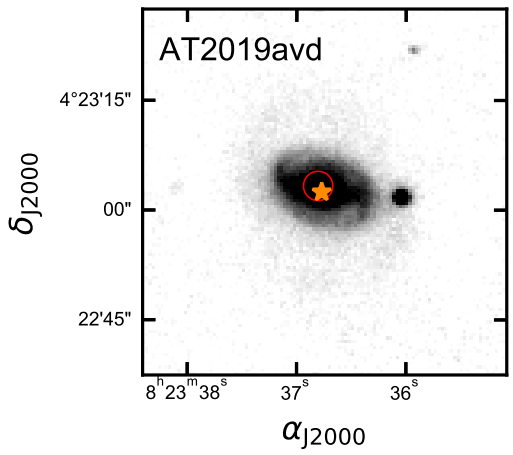


Figure 4.1: Pan-STARRS *g*-band image centred on the host galaxy of AT 2019avd. The dark orange star and red circle mark the ZTF position and eROSITA localisation respectively, where the radius of the circle is set to the 2″ uncertainty on the eROSITA source position.

of SMBH binaries (SMBHB) with an accretion disc (Kim et al., 2018). It is clear that such different physical origins may result in a diverse range of observed variability behaviours.

In this paper, we report on the ongoing extreme event AT 2019avd, which is a novel addition to the already diverse population of nuclear transients. AT 2019avd is associated to the previously inactive galaxy 2MASX J08233674+0423027 at $z = 0.029$ (see Fig. 4.1), and was first reported as ZTF19aaiqmgl at the Transient Name Server (TNS[2]) following its discovery by ZTF on 2019-02-09 UT[3] (Nordin et al., 2019). The transient was independently detected more than a year later on 2020-04-28 as a new ultra-soft nuclear X-ray source (Malyali et al., 2020) during the first all-sky survey of the eROSITA instrument (Predehl et al., in press) on-board the Russian/German *Spectrum-Roentgen-Gamma* (*SRG*) mission.

This work presents X-ray (*SRG*/eROSITA, *Swift*/XRT), optical/UV/mid-infrared (MIR) photometric (ZTF, ASAS-SN, NEOWISE-R, *Swift*/UVOT), and optical spectroscopic (NOT/ALFOSC, Las Cumbres Floyds, ANU/WiFeS) observations of AT 2019avd. In Section 4.2, we report our X-ray observations and analysis of AT 2019avd, whilst the photometric evolution and host galaxy properties are presented in Section 4.3. We then present details of our optical spectroscopic follow-up campaign in Section 4.4, before discussing possible origins for AT 2019avd in Section 5.5, and conclude in Section 4.6. We adopt a flat ΛCDM cosmology throughout this paper, with $H_0 = 67.7\,\mathrm{km\,s^{-1}Mpc^{-1}}$, $\Omega_\mathrm{m} = 0.309$ (Planck Collaboration XIII, 2016); $z = 0.029$ thus corresponds to a luminosity distance of 130 Mpc. All magnitudes will be reported in the AB system, unless otherwise stated.

[2]`https://wis-tns.weizmann.ac.il/`

[3]all dates in this paper will be reported in UT format.

## 4.2 X-ray observations

### 4.2.1 eROSITA discovery

AT 2019avd was discovered in a dedicated search for candidate TDEs in the first eROSITA all-sky survey (eRASS1). Here, the eROSITA source catalogue (version 945 of the source detection pipeline of the eROSITA Science Analysis Software, eSASS, Brunner et al. in prep.) was systematically examined for new soft X-ray sources associated with the nuclei of galaxies that showed no prior indication of being an AGN.

The eROSITA data for AT 2019avd are composed of four consecutive scans with gaps of 4 hr each and a midtime of 2020-04-28. The total on-source exposure amounts to 140 s (see Table 4.1). The source was localised to ($RA_{J2000}$, $Dec_{J2000}$)=(08h23m37s, 04°23′03″), with a $1\sigma$ positional uncertainty of 2″, which is consistent with the nucleus of the galaxy 2MASX J08233674+0423027.

Photons were extracted using the eSASS task SRCTOOL (version 945) choosing a circular aperture of radius 36″ centred on the above position (84 counts were detected within this region). Background counts were selected from a circular annulus of inner and outer radii 72″ and 144″, respectively. Using the best-fit spectral model (see Section 4.2.3), we derived a $0.2-2$ keV flux of $(1.4\pm0.2)\times10^{-12}$ erg cm$^{-2}$s$^{-1}$ ($1\sigma$).

No X-ray source has previously been detected at the location of AT 2019avd. Using both the Upper Limit Server[4] and webPIMMS[5], and assuming an absorbed black-body spectral model with $kT = 80$ eV, and Galactic neutral hydrogen column density (see also Section 4.2.3), $N_{\rm H} = 2.42\times10^{20}$ cm$^{-2}$, we infer an $0.2-2$ keV $3\sigma$ upper limit of $1.7\times10^{-14}$ erg cm$^{-2}$s$^{-1}$ for a serendipitous 7 ks *XMM-Newton* pointed observation obtained on 2015-04-08[6]. Earlier constraints can be derived from *ROSAT* observations obtained on 1990-10-14, 1996-11-13, and 1997-04-11 with $3\sigma$ upper limits of $4.2\times10^{-13}$, $4.0\times10^{-13}$ , and $1.2\times10^{-13}$ erg cm$^{-2}$s$^{-1}$, respectively.

eROSITA thus first observed AT 2019avd in a state where it had brightened by at least a factor of 90 in the $0.2-2$ keV band relative to the deepest archival X-ray observation (luminosity history presented in Fig. 4.2).

### 4.2.2 Swift follow-up

Triggered by the eROSITA detection, a series of follow-up observations were performed with the *Neil Gehrels Swift Observatory* (P.I.s: A. Malyali & B. Trakhtenbrot). Observations were obtained roughly every 7 days, until the source was no longer visible due to Sun angle constraints; a further *Swift* observation was then obtained $\sim$ 3 months later. A log of the observations can be found in Table 4.1. The XRT observations were performed in photon counting mode. The data were reduced using the XRTPIPELINE task included in version 6.25 of the HEASOFT package. Spectra for each of the five epochs were extracted using the XRTPRODUCTS task. Source counts

[4] `http://xmmuls.esac.esa.int/upperlimitserver/`

[5] `https://heasarc.gsfc.nasa.gov/cgi-bin/Tools/w3pimms/w3pimms.pl`

[6] *XMM-Newton* OBSID=0741580501

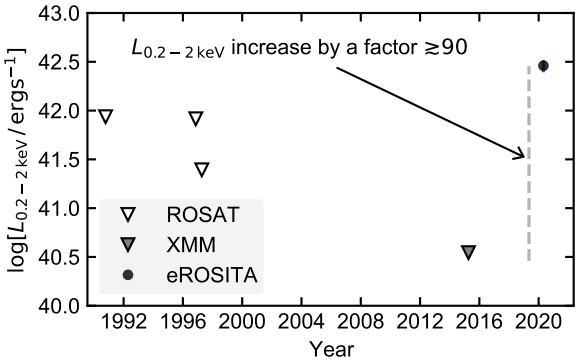


Figure 4.2: Long-term X-ray light curve in the 0.2–2 keV energy band of AT 2019avd up until the first eROSITA observation. Triangles denote $3\sigma$ upper limits for *ROSAT*/PSPC and *XMM-Newton*/EPIC-pn, whilst the black circle marks the *SRG*/eROSITA discovery, where AT 2019avd is at least 90 times brighter than the *XMM-Newton* $3\sigma$ upper limit. The error bar on the eROSITA marker encloses the 68% credible region on the observed luminosity.

Table 4.1: Log of *SRG*/eROSITA and *Swift*/XRT observations of AT 2019avd until 2020-09-16. For eROSITA, the mid-date of the coverage in eRASS1 is given.

| Date | MJD | Telescope | ObsID | Exp. [s] |
|---|---|---|---|---|
| 2020-04-28 | 58967.7 | *SRG*/eROSITA | - | 140 |
| 2020-05-13 | 58982.4 | *Swift*/XRT | 00013495001 | 1617 |
| 2020-05-19 | 58988.3 | *Swift*/XRT | 00013495002 | 1966 |
| 2020-05-25 | 58994.0 | *Swift*/XRT | 00013495003 | 1982 |
| 2020-06-03 | 59003.3 | *Swift*/XRT | 00013495004 | 494 |
| 2020-06-10 | 59010.6 | *Swift*/XRT | 00013495005 | 1739 |
| 2020-09-16 | 59108.4 | *Swift*/XRT | 00013495006 | 2967 |

were extracted from a circular aperture of radius 47″ and background counts extracted from a circular annulus of inner and outer radii 70″ and 250″, respectively[7].

Observations with the Ultraviolet and Optical Telescope (UVOT; Roming et al. 2005) were obtained simultaneously with the XRT observations. Imaging was performed at three epochs (00013495001, ..004, ..005) using the UVW1 filter with exposures of 1.36, 1.95, and 1.93 ks, respectively. The remaining three observations utilised all six UVOT filters (UVW2, UVM2, UVW1, U, B, V) with accordingly shorter exposure times.

The UVOT flux was extracted with the UVOTSOURCE task using a 9″ radius aperture centred on the optical position of AT 2019avd, whilst a nearby circular region with 15″ radius was used for background subtraction. The photometry was extracted from each unique *Swift* observation ID, and is presented in Table 4.2 (we note that this photometry includes both AGN and host galaxy emission in order to be consistent with the SED fitting in Section 4.3.3). Relative to

[7]eROSITA and XRT have different PSFs and instrument backgrounds, thus the radii of the extraction regions were chosen based on each instrument and differ here.

Table 4.2: *Swift* UV photometry (corrected for Galactic extinction using the UVOT correction factors in Table 5 of Kataoka et al. 2008). The model magnitudes (for the host galaxy) were obtained by convolving the best-fit SED model (Section 4.3.3) with the UVOT transmission curves. A hyphen denotes that the given filter was not used on that observation date.

| Date | UVW1 | UVM2 | UVW2 |
|---|---|---|---|
| Model | 18.88 | 19.16 | 19.26 |
| 2020-05-13 | 18.01 ± 0.04 | - | - |
| 2020-05-19 | 18.23 ± 0.15 | 18.28 ± 0.11 | 18.27 ± 0.10 |
| 2020-05-25 | 17.85 ± 0.07 | 18.30 ± 0.07 | 18.31 ± 0.06 |
| 2020-06-03 | 17.89 ± 0.04 | - | - |
| 2020-06-10 | 17.80 ± 0.04 | - | - |
| 2020-09-16 | 17.78 ± 0.05 | 18.17 ± 0.06 | 18.23 ± 0.05 |

Table 4.3: Summary of priors adopted in the BXA analysis of the eROSITA and XRT spectra. For each fit, a log-uniform prior on $N_{\rm H}$ between $(0.8N_{\rm H}, 1.2N_{\rm H})$ was defined, where $N_{\rm H} = 2.42 \times 10^{20}\,{\rm cm}^{-2}$ (see Section 4.2.3). $\Gamma$ denotes the slope of a power law, $kT$ the black-body temperature, $A$ the normalisation. The prior over $A$ is in units $1.05{\times}10^{-6}{\rm erg\,cm^{-2}s^{-1}}$.

| Model | Priors |
|---|---|
| `tbabs*bbody` | $\log[kT/{\rm keV}] \sim \mathcal{U}(-2, 1)$, $\log[A] \sim \mathcal{U}(-10, 10)$ |
| `tbabs*powerlaw` | $\Gamma \sim \mathcal{U}(0, 8)$, $\log[A] \sim \mathcal{U}(-10, 10)$ |

UV photometry obtained prior to the initial optical outburst (see Section 4.3.3 and Fig. 4.7), AT 2019avd has brightened by $\sim 1$ mag in the UVW1, UVM2, and UVW2 bands, and brightens only by $\sim 0.1 - 0.2$ mag over *Swift* observations between 2020-05-13 and 2020-09-16.

### 4.2.3 X-ray spectral fitting

X-ray spectra were analysed using the Bayesian X-ray Analysis software (BXA, Buchner et al. 2014), which connects the nested sampling algorithm MultiNest (Feroz & Hobson, 2008) with the fitting environment CIAO/Sherpa (Freeman et al., 2001) and XSPEC (Arnaud, 1996). The spectra were fitted unbinned using the C-statistic (Cash, 1976), and the eROSITA and XRT backgrounds were both modelled using the principal component analysis (PCA) technique described in Simmonds et al. (2018). For each set of eROSITA and XRT spectra, a joint fit on both the source and background spectra was run. Two different models for the source spectra were used: (i) an absorbed black body (`tbabs*blackbody`), and (ii) an absorbed power law (`tbabs*powerlaw`). The equivalent Galactic neutral hydrogen column density, $N_{\rm H}$, was allowed to vary by 20% from its tabulated value in the HI4PI survey of $2.42 \times 10^{20}\,{\rm cm}^{-2}$ (HI4PI Collaboration et al., 2016) during fitting. The complete set of priors adopted under each model is listed in Table 4.3, whilst an example of the BXA fit to the eROSITA spectrum is shown in Fig. 4.3, and spectral fit results are presented in Table 4.4.

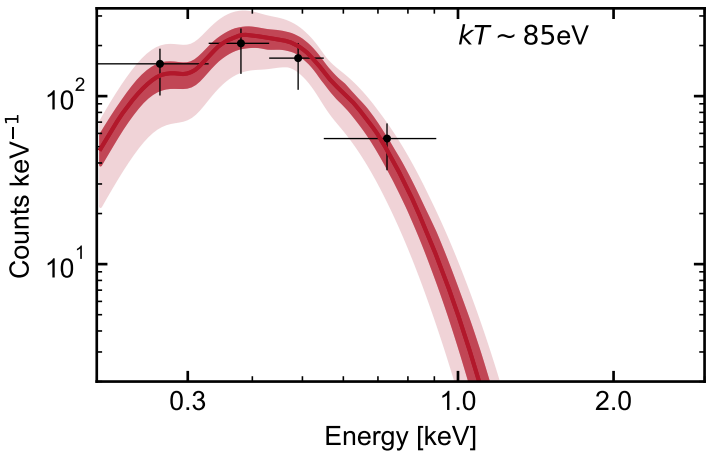


Figure 4.3: BXA fit to the eROSITA eRASS1 spectrum. Black markers are the binned observed data, whilst the red represents the fitted convolved model for `tbabs*blackbody` (darker and light red bands enclose the 68 % and 95 % posterior uncertainty on the model at each energy). Both the black-body and power-law fits to the (low count) eRASS1 spectrum suggest that the source is ultra-soft (see Table 4.4).

Table 4.4: X-ray spectral fit results from applying BXA to the extracted eROSITA and XRT spectra, with uncertainties enclosing 68% of the posterior for each parameter. $F_{0.2-2\mathrm{keV}}$ is the inferred observed (unabsorbed) flux under each model.

| OBSID | `tbabs*blackbody` | | | `tbabs*powerlaw` | | |
|---|---|---|---|---|---|---|
| | $N_H$ $[\times 10^{20}\mathrm{cm}^{-2}]$ | kT [eV] | $F_{0.2-2\mathrm{keV}}$ $[\times 10^{-12}\,\mathrm{erg\,cm^{-2}\,s^{-1}}]$ | $N_H$ $[\times 10^{20}\mathrm{cm}^{-2}]$ | $\Gamma$ | $F_{0.2-2\mathrm{keV}}$ $[\times 10^{-12}\,\mathrm{erg\,cm^{-2}\,s^{-1}}]$ |
| eRASS1 | $2.3^{+0.3}_{-0.3}$ | $85^{+6}_{-5}$ | $1.4^{+0.2}_{-0.2}$ | $2.5^{+0.3}_{-0.3}$ | $4.2^{+0.3}_{-0.3}$ | $1.6^{+0.2}_{-0.2}$ |
| 00013495001 | $2.4^{+0.4}_{-0.3}$ | $72^{+8}_{-8}$ | $1.4^{+0.2}_{-0.2}$ | $2.4^{+0.3}_{-0.3}$ | $5.3^{+0.4}_{-0.4}$ | $2.5^{+0.5}_{-0.5}$ |
| 00013495002 | $2.4^{+0.3}_{-0.3}$ | $83^{+12}_{-11}$ | $1.4^{+0.4}_{-0.4}$ | $2.4^{+0.3}_{-0.3}$ | $5.2^{+0.7}_{-0.6}$ | $2.6^{+0.8}_{-0.8}$ |
| 00013495003 | $2.4^{+0.3}_{-0.3}$ | $132^{+10}_{-10}$ | $1.0^{+0.1}_{-0.1}$ | $2.5^{+0.3}_{-0.3}$ | $3.7^{+0.2}_{-0.3}$ | $1.4^{+0.2}_{-0.2}$ |
| 00013495004 | $2.4^{+0.3}_{-0.3}$ | $107^{+10}_{-10}$ | $1.0^{+0.2}_{-0.2}$ | $2.4^{+0.3}_{-0.3}$ | $4.2^{+0.3}_{-0.3}$ | $1.6^{+0.3}_{-0.3}$ |
| 00013495005 | $2.4^{+0.3}_{-0.3}$ | $91^{+6}_{-6}$ | $1.5^{+0.2}_{-0.2}$ | $2.5^{+0.3}_{-0.3}$ | $4.9^{+0.3}_{-0.3}$ | $2.6^{+0.4}_{-0.4}$ |
| 00013495006 | $2.4^{+0.3}_{-0.3}$ | $115^{+3}_{-3}$ | $9.7^{+0.4}_{-0.4}$ | $2.8^{+0.1}_{-0.1}$ | $4.1^{+0.1}_{-0.1}$ | $14.0^{+0.7}_{-0.7}$ |

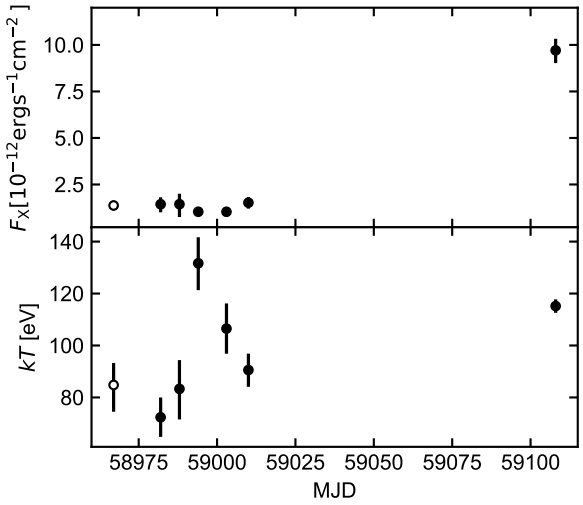


Figure 4.4: X-ray evolution of AT 2019avd. The empty and filled black markers represent the eROSITA and XRT observations respectively; error bars enclose 95% of the posterior.

Over the course of the six weeks following the initial eROSITA detection, there was no major variability in the $0.2-2$ keV flux between the eROSITA and XRT observations (Table 4.4 and Fig. 4.4). However, the $0.2-2$ keV flux in the last *Swift* epoch increased by a factor of about six relative to the previous observation.

AT 2019avd remained in an ultra-soft state during the *Swift* monitoring campaign, although there is variability in the inferred black-body temperatures ($kT$ ranges between minimum and maximum values of $72\pm8$ eV and $132\pm10$ eV, respectively). The inferred black-body temperatures are similar to those measured in the X-ray emission of previously observed thermal TDEs ($45 \lesssim kT \lesssim 130$ eV, e.g. van Velzen et al. 2021), and are also consistent with the temperatures of the soft excess shown in AGN (e.g. Table A1 in Gliozzi & Williams 2020).

## 4.3 Photometric evolution and host galaxy properties

### 4.3.1 Optical evolution

The region around the position of AT 2019avd has been monitored by ZTF (Bellm et al., 2019; Graham et al., 2019) in the *r* and *g* bands from 2019-01-12 until the time of writing. On 2019-02-09 (over a year before the eROSITA detection), ZTF first detected the transient ZTF19aaiqmgl with an inferred separation from the galaxy centre of $0\farcs04$ [8], and *r*-band magnitude $17.64 \pm 0.07$ (reference subtracted, Fig. 4.1).

For MJD<58855 (2020-01-07), we obtained a forced photometry ZTF light curve for AT 2019avd (Masci et al., 2019). For MJD>58855, we downloaded the ZTF light curve of AT 2019avd using

[8] https://lasair.roe.ac.uk/object/ZTF19aaiqmgl/

the Lasair alert broker (Smith et al., 2019), which processes and reports to the community on transients detected within the large ZTF data streams. Both of these light curves are constructed from PSF-fit photometry measurements run on ZTF difference images. We also obtained additional photometric observations with the Spectral Energy Distribution Machine (SEDM; Blagorodnova et al. 2018) on the Palomar 60-inch telescope. The SEDM photometry was host-subtracted using SDSS reference images, as described in Fremling et al. (2016). These two light curves, and the host-subtracted SEDM photometry, were then combined for subsequent analysis, and are shown in Fig. 4.5.

After the initial detection on 2019-02-09, AT 2019avd continued to brighten until reaching its maximum observed brightness of $r \sim 16.8$ mag on 2019-02-20. Between 2019-02-24 and 2020-01-01, the $g$-band magnitude of the host nucleus decayed nearly monotonically from 17.13±0.09 mag to 20.08±0.20 mag, followed by a re-brightening to 18.58±0.13 mag on 2020-05-03. The late time SEDM photometry around 2020-09-19 revealed a further brightening to $r$ and $g$-band magnitudes of $\sim$ 17.6 mag and $\sim$ 18.4 mag respectively. The first eROSITA observation occurred during the rise of the second major peak of the ZTF light curve (Fig. 4.5).

The location of AT 2019avd has also been monitored in the $V$-band by ASAS-SN (Shappee et al., 2014; Kochanek et al., 2017) from February 2012 to November 2018, and in the $g$-band from October 2017 to September 2020 (the time of writing). No major optical outbursts were seen in the ASAS-SN light curve prior to the ZTF detection (Fig. A.1); given the joint ASAS-SN and ZTF light curves, it is likely that the system 'ignited' around MJD = 58510 (2019-01-27).

### Rise and decay timescales in the light curve

In the following, we fit the light-curve model presented in equation 1 of van Velzen et al. (2019), which models the rise with a half-Gaussian function, and an exponential function for the decay, to the first and second peaks of the ZTF light curve, using UltraNest[9] (Buchner, 2016, 2019) as our sampler. Whilst such a model is not physically motivated, it enables a comparison of the timescales involved in the light curve of AT 2019avd with those of the population of ZTF nuclear transients presented in van Velzen et al. (2019).

While fitting the first peak, we first filter out observations outside of the MJD period between 58450 and 58650, and we then run a joint fit of the $g$ and $r$ band observations in flux space. Our model has seven free parameters, defined following the notation of van Velzen et al. (2019): $\sigma_r$ and $\sigma_g$, the rise timescale of the light curve in the $r$ and $g$ bands respectively; $\tau_r$ and $\tau_g$, the decay timescale of the light curve in $r$ and $g$ bands; $F_{\mathrm{peak},r}$ and $F_{\mathrm{peak},g}$, the peak flux in $r$ and $g$ bands; $t_{\mathrm{peak}}$, the time of the peak of the light curve (to enable a comparison with van Velzen et al. 2019, we assume that the light-curve model peaks at the same time in both of these bands). For the second peak, we filter out observations outside of the MJD period 58840 and 59115 (the late-time SEDM datapoints are used in the fitting), and because we do not sample the decay of this peak, we only model the rise here. The model for the second peak has five free parameters, with $\tau_r$ and $\tau_g$ now being omitted. We list our priors in Table A.1, and present the fits in Fig. 4.6.

From the posterior means, we infer $\sigma_r = 7.9 \pm 0.3$, $\sigma_g = 7.2 \pm 0.2$, $\tau_r = 58.2 \pm 0.5$ and

[9] https://github.com/JohannesBuchner/UltraNest

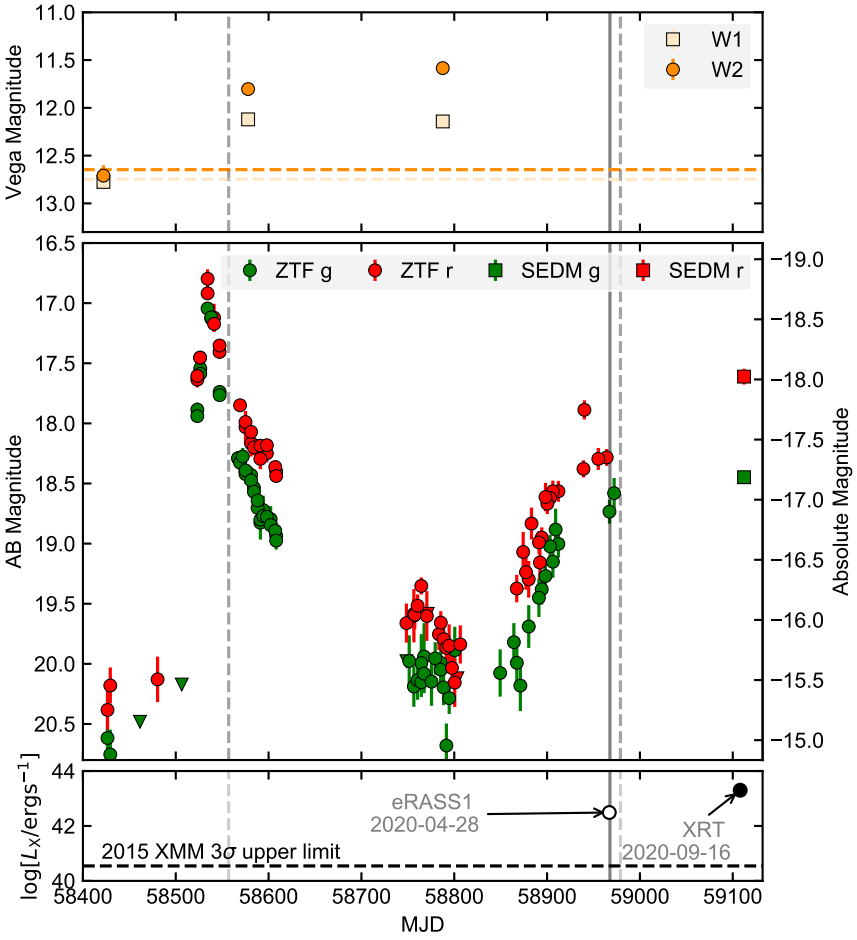


Figure 4.5: NEOWISE-R (non-host subtracted, top) and ZTF/ SEDM (middle) light curves of AT 2019avd, with the immediate 0.2 – 2 keV X-ray history shown in the bottom panel. The eROSITA eRASS1 detection and the *Swift* observation from 2020-09-16 are the empty and filled black markers, respectively. The solid grey vertical line marks the MJD of the eRASS1 observation, whilst grey dashed lines mark the times of the NOT and the first FLOYDS spectrum (Table 4.5). No significant variability before the initial 2019 outburst is observed in the host nucleus of AT 2019avd with archival NEOWISE-R and ASAS-SN observations (Fig. A.1). The NEOWISE-R observations pre-outburst are observed with mean W1, W2 marked out in the top panel by the cream and orange dashed lines respectively. For plotting clarity, we omit the high-cadence ZTF Partnership observations obtained between MJD 58820 and 58860, and we rebin the $\sim$ 3 SEDM observations in each filter into a single data point.

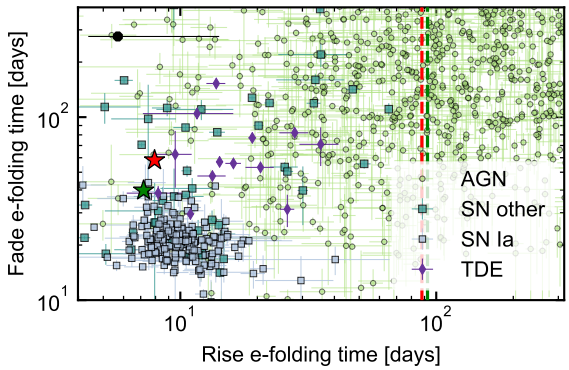


Figure 4.6: AT 2019avd variability compared with previously classified ZTF nuclear transients (non-AT 2019avd data presented originally in van Velzen et al. 2021), with red and green stars computed from the fitted model components for each respective filter. The red and green vertical lines mark the e-folding rise time of the second optical peak in the $r$ and $g$ bands, respectively. We also plot the rise and decay e-fold timescales inferred from the ASAS-SN $V$-band light curve of the nuclear transient AT 2017bgt (Trakhtenbrot et al. 2019a; see also Section 4.5.1) with a black marker. Not only is the double-peaked light curve of AT 2019avd clearly distinct from the other light curves of sources in the AT 2017bgt nuclear transient class, but the first peak of AT 2019avd decays much faster than the AT 2017bgt flare, whilst the second peak rises much slower than the AT 2017bgt flare.

$\tau_g = 39.8 \pm 0.4$ days for the first optical peak (68% credible intervals). Whilst the rise timescales in each filter are consistent with each other to within $2\sigma$, the decay timescales in each filter significantly differ. With $\tau_r > \tau_g$, the first peak shows a potential cooling signature during its decay phase, although we are unable to constrain the temperature evolution during this because of a lack of contemporaneous observations in other wavelength bands. Relative to the population of nuclear transients in van Velzen et al. (2019), one sees that these are short rise and decay timescales relative to those of AGN flares, and are thus more similar to those in the van Velzen et al. (2019) sample of TDEs and SNe (Fig. 4.6). As expected from Fig. 4.5, the inferred rise times for the second peak are longer and more AGN-like, with $\tau_r \sim 88$ days and $\tau_g \sim 93$ days.

### 4.3.2 Mid-infrared variability

The location of AT 2019avd was observed in the $W1$ (3.4 $\mu$m) and $W2$ (4.6 $\mu$m) bands by the Wide-Field Infrared Survey Explorer mission (WISE, Wright et al. 2010) in 2010, Near-Earth Object WISE (NEOWISE; Mainzer et al. 2011) in late 2010 and 2011, and from December 2013 until now, twice per year as part of the NEOWISE reactivation mission (NEOWISE-R; Mainzer et al. 2014). The NEOWISE-R light curve was obtained from the NASA/IPAC Infrared Science Archive[10] by compiling all source detections within 5″ of the ZTF transient position. Individual flux measurements were rebinned to one data point per NEOWISE-R all-sky scan

[10] `https://irsa.ipac.caltech.edu/frontpage/`

(using a weighted mean) and converted into magnitudes. The resulting light curve is shown in Fig 4.5.

The MIR light curve was observed to be flat prior to the initial ZTF outburst, but showed significant brightening in the first NEOWISE-R epoch obtained thereafter. Observations obtained $\sim 6$ months later found the source to still be in the bright state despite having faded by $\sim 3$ mag in the optical. The MIR brightening was also accompanied by a significant reddening, evolving from $W1 - W2 \sim 0.08$ mag in AllWISE, to a more AGN-like $W1 - W2 \sim 0.6$ mag during flaring. The $W1 - W2$ colour before the outburst is much lower than the suggested cuts ($W1 - W2 \gtrsim 0.7$ mag) for identifying AGNs in previous MIR classification schemes (Stern et al., 2012; Assef et al., 2013, 2018), further supporting the hypothesis that there was no strong recent AGN activity in AT 2019avd at that time (although the use of WISE colours for selecting AGNs is less effective at lower AGN luminosities; see discussion in Padovani et al. 2017).

### 4.3.3 Host-galaxy properties

The spectral energy distribution (SED) of the host galaxy of AT 2019avd was compiled from archival[11] UV to MIR photometry from *GALEX* (FUV, NUV), SDSS DR12 ($g$, $r$, $i$, $z$), UKIDSS ($y$, $J$, $H$, $K$), and AllWISE (W1, W2). The SED was modelled using CIGALE (Burgarella et al., 2005; Boquien et al., 2019), which allows the estimation of the physical parameters of a galaxy by fitting composite stellar populations combined with recipes describing the star formation history and attenuation. The best-fitting model (see Fig. 4.7) is that of a galaxy with a stellar mass of $(1.6 \pm 0.8) \times 10^{10} M_{\odot}$, a star formation rate (SFR) of $0.17 \pm 0.05\ M_{\odot}\mathrm{yr}^{-1}$, and little attenuation, $\mathrm{E(B-V)} = 0.03 \pm 0.02$ mag, which experienced a burst of star formation $3.7 \pm 0.2$ Gyr ago. The inferred stellar mass and SFR place the host galaxy of AT 2019avd in the 'green valley' between the star-forming main sequence and quenched elliptical galaxies (adopting the green valley definition presented in Law-Smith et al. 2017b).

The SED fit suggests that the host galaxy did not show strong signs of nuclear activity prior to the detection of AT 2019avd. This is further supported by the absence of a radio counterpart in the FIRST catalogue (Becker et al., 1995) within 30″ of AT 2019avd, with a catalogue upper detection limit at this position of 0.96 mJy/beam[12].

## 4.4 Optical spectral analysis

### 4.4.1 Spectroscopic observations

On 2019-03-15, ~33 days after the first observed peak in the ZTF light curve, an optical spectrum of AT 2019avd was obtained by Gezari et al. (2020) with the Alhambra Faint Object Spectrograph and Camera (ALFOSC)[13] on the 2.56 m Nordic Optical Telescope (NOT). The spectrum was obtained with a 1.″0 wide slit, grism #4 (covering the wavelength region from 3650-9200 Å),

[11] 'Archival' is defined here by photometry taken prior to the initial ZTF optical outburst.

[12] `http://sundog.stsci.edu/cgi-bin/searchfirst`.

[13] `http://www.not.iac.es/instruments/alfosc`

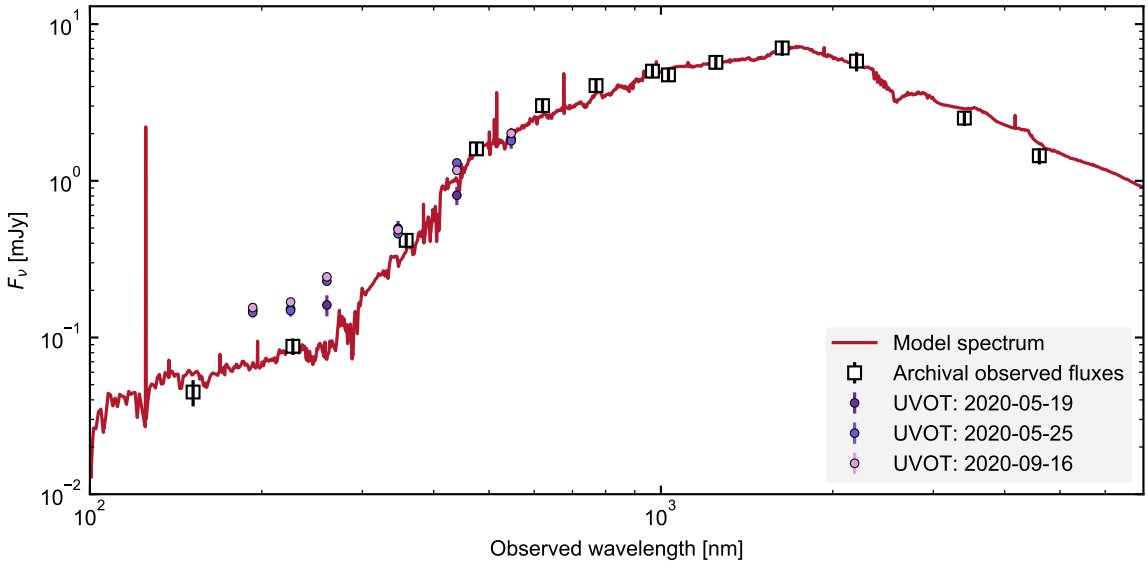


Figure 4.7: Spectral energy distribution of the host galaxy of AT 2019avd compiled from archival GALEX, SDSS, UKIDSS, and ALLWISE photometry, with the best-fit model shown as a red solid line. The three epochs of *Swift* UVOT photometry where all filters were used are also plotted. AT 2019avd shows a $\sim$ 1 mag rise in the UVW1, UVM2, and UVW1 bands relative to the best fit model to the archival photometry.

and the slit was positioned along the parallactic angle at the beginning of the 1800s exposure. Reductions were performed in a standard way using mainly `iraf` based software, including bias corrections, flat fielding, wavelength calibration using HeNe arc lamps imaged immediately after the target and flux calibrations using observations of a spectrophotometric standard star.

No further spectra were taken until after eROSITA had detected the large-amplitude soft-X-ray flare from AT 2019avd in late April 2020, which triggered a further five epochs of spectroscopy (dates listed in Table 4.5) using the FLOYDS spectrographs (Brown et al., 2013) mounted on the Las Cumbres Observatory 2m telescopes at Haleakala, Hawaii, and Siding Spring, Australia. Each spectrum was taken with a 3.6ks exposure, using the 'red/blu' grism and a slit width of 2″. The spectra were reduced using PyRAF tasks as described in Valenti et al. (2014). FLOYDS covers the entire 3500-10000 Å range in a single exposure by capturing two spectral orders (one red and one blue) simultaneously, yielding $R \sim 400$. The different orders are usually merged into a single spectrum using the region between 4900 and 5700 Å, which is present in both the red and blue orders. However, in this case, in order to avoid erroneous wavelength shifts at the blue edge of the red order (where there are fewer arclines), all FLOYDS spectra were merged using a reduced stitching region of 5400 to 5500 Å[14]. This stitching was done manually in Python, by replacing fluxes in that wavelength range with an average of the linear interpolations of the two orders.

In addition, a higher resolution spectrum (R∼3000) was obtained on 2020-05-29 with the Wide Field Spectrograph (WiFeS; Dopita et al., 2007, 2010) mounted on the 2.3m ANU telescope at Siding Spring Observatory. We employed the R3000 and B3000 gratings, and obtained an arc lamp exposure after each target exposure. The total spectral range from the two gratings is 3500 to 7000 Å. The data were reduced using the PyWiFeS reduction pipeline (Childress et al., 2014), which produces three-dimensional data (data cubes). These spectra are bias subtracted, flat-fielded, wavelength and flux calibrated, and corrected for telluric absorption. We then extracted the spectra from the slitlets that captured AT 2019avd using the IRAF (Tody, 1986) task APALL which allowed for background subtraction.

A comparison of the NOT and WiFeS spectra is presented in Fig. 4.8, and the spectral evolution in the FLOYDS spectra is shown in Fig. 4.9. A log of the spectroscopic observations of AT 2019avd is presented in Table 4.5. We note that we have not found any archival optical spectra of the host galaxy that were obtained prior to the initial 2019 outburst discovered by ZTF.

### 4.4.2 Summary of the main observed features of the optical spectra

The NOT spectrum from 2019-03-15 appears similar to broad line AGN spectra, showing a relatively flat continuum (in terms of $F_\lambda$) and broad Balmer emission lines (H$\alpha$, H$\beta$, H$\gamma$, H$\delta$; Fig. 4.8). However, the strong Fe II complex that is frequently seen in some AGNs is not present. The H$\alpha$ profile is asymmetric due to the blending of unresolved H$\alpha$ and narrow [N II] 6549, 6583 Å lines, whilst the asymmetry of the H$\gamma$ line is likely due to blending of H$\gamma$ and [O III] 4363 Å emission. The other notable features are the [S II] doublet at 6717 and 6731 Å (again blended, but later resolved in the WiFeS spectrum), and the weak He I emission at 5876 Å. As no archival

[14]The most extreme arcline used to calibrate each order is at ∼5460 Å.

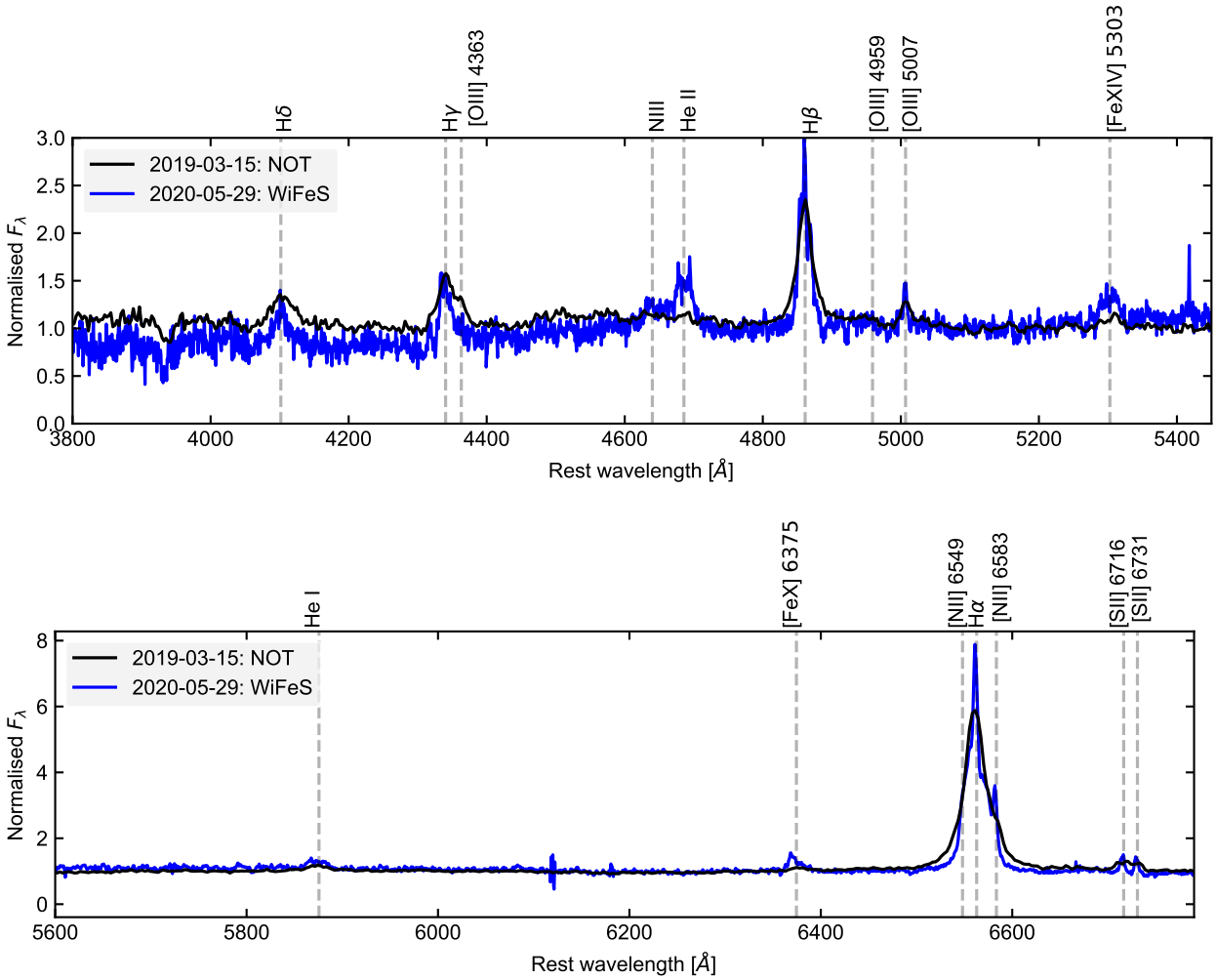


Figure 4.8: Comparison of NOT and WiFeS spectra (black and blue respectively). The top panel shows the wavelength range 3800-5450 Å, while the bottom panel shows the 5600-6800 Å range. The most notable changes are (a) the emergence of the broad emission feature around rest-frame wavelength 4686 Å and (b) an increase in intensity of the high-ionisation coronal Fe lines (~5300 and 6370 Å). The WiFeS spectrum is of much higher resolution relative to the NOT spectrum, and therefore is able to better resolve narrow emission lines, such as the [S II] doublet at 6716 and 6731 Å. Neither are shown corrected for Galactic extinction. The NOT spectrum was normalised by its continuum flux in the 5100-5200 Å range (rest frame), whilst the blue and red arms of the WiFeS spectra were normalised in the 5100-5200 Å and 6400-6450 Å ranges respectively (rest frame).

Table 4.5: Spectroscopic observations of AT 2019avd.

| UT Date | Tel. | Instrument | Exp. [ks] | Airmass |
|---|---|---|---|---|
| 2019-03-15 | NOT | ALFOSC | 1.8 | 1.5 |
| 2020-05-10 | FTS | FLOYDS-S | 3.6 | 1.4 |
| 2020-05-12 | FTS | FLOYDS-S | 3.6 | 1.6 |
| 2020-05-18 | FTN | FLOYDS-N | 3.6 | 1.6 |
| 2020-05-29 | ANU | WiFeS | 1.8 | 1.5 |
| 2020-05-31 | FTS | FLOYDS-S | 3.6 | 1.7 |
| 2020-06-06 | FTS | FLOYDS-S | 3.6 | 1.9 |

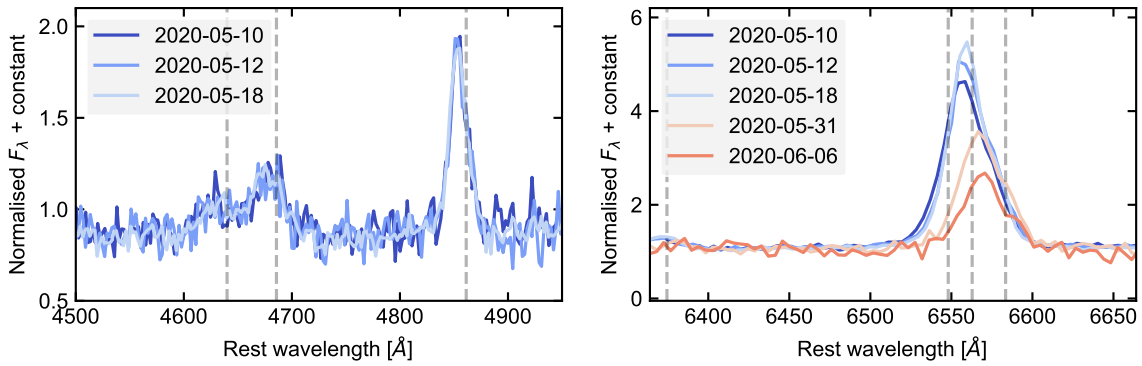


Figure 4.9: Evolution of the Bowen+H$\beta$ (top) and H$\alpha$ (bottom) Balmer emission lines observed through the five epochs of FLOYDS spectroscopy. Grey dashed lines match those in Fig. 4.8. Epochs 2020-05-31 and 2020-06-06 were of low S/N in the blue wavelength range, and thus are omitted from the plot here. The minor evolution of the H$\alpha$ peak position over the FLOYDS spectra was deemed to be most likely due to aperture-related effects during observations.

Table 4.6: Emission line ratios relative to [O III] 5007Å, where the inferred [O III] 5007Å flux in each spectrum is $1.34 \pm 0.09 \times 10^{-15}$ erg cm$^{-2}$s$^{-1}$ and $4.8 \pm 0.7 \times 10^{-16}$ erg cm$^{-2}$s$^{-1}$. The two spectra were obtained with different slit widths and orientations, and have not been calibrated with independent photometric measurements, hence the line ratios relative to [O III] 5007Å reported here. A dashed entry indicates that a given emission line was not clearly detected in the optical spectral fitting.

| Date | N III 4640 | He II 4686 | H$\beta$ | H$\alpha$ | [N II] (6549 + 6583) | [S II] (6716 + 6731) |
|---|---|---|---|---|---|---|
| 2019-03-15 | - | - | $10^{+3}_{-3}$ | $38^{+3}_{-3}$ | $7^{+1}_{-1}$ | $1.7^{+0.1}_{-0.1}$ |
| 2020-05-29 | $4^{+5}_{-3}$ | $6^{+7}_{-5}$ | $11^{+2}_{-2}$ | $41^{+6}_{-6}$ | $5^{+1}_{-1}$ | $1.7^{+0.2}_{-0.2}$ |

spectrum of the host galaxy is available, we are unable to judge whether or not the main observed emission features appeared at the onset of the extreme optical variability. The WiFeS spectrum from 2020-05-29 (Fig. 4.8) shows the same emission features as the NOT spectrum, with the addition of a broad emission feature around 4680 Å and an apparent increase in intensity of a set of high-ionisation coronal lines ([Fe XIV] 5303 Å and [Fe X] 6375 Å, with ionisation potentials of 392 and 262eV respectively). We assume that the [Fe X] is not blended with the [O I] 6364 Å emission feature, because the latter is expected to be a third of the intensity of the [O I] 6300 Å emission (e.g. Pelat et al. 1987), which is not detected.

The FLOYDS spectra (Fig. 4.9) show no major evolution in the Balmer emission line profiles, and show the broad emission feature around 4680 Å from 2020-05-10 (for epochs with sufficiently high S/N ratios in the blue wavelength range), which was reported to the TNS (and first identified) in Trakhtenbrot et al. (2020).

### 4.4.3 Optical spectrum modelling

For the two higher resolution spectra (NOT and WiFeS), the region around the main observed emission lines is fitted separately (H$\gamma$, 4240Å $< \lambda <$ 4440Å; He II, 4500Å $< \lambda <$ 4800Å; H$\beta$, 4700Å $< \lambda <$ 5000Å; H$\alpha$, 6364Å $< \lambda <$ 6764Å; [S II] doublet, 6650Å $< \lambda <$ 6800Å; and ±100Å of the line centre for [O III] 5007 Å, [Fe X] 6375 Å). Each emission line complex is modelled with multiple Gaussians (an overview of these is presented in Table A.2), and each complex is fitted independently of the others. For all spectral fits, we assume a flat continuum component during the fitting process, and run our model fitting using the region slice sampler option within UltraNest. Spectral fits for the NOT and WiFeS spectra are shown in Figs. 4.10 and 4.11, whilst the spectral fit results are listed in Tables 4.6, 4.7, and 4.8.

### 4.4.4 Emission line diagnostics

#### Balmer emission

From the best-fitting spectral models, we infer a broad Balmer decrement, $F(\mathrm{H}\alpha^{\mathrm{b}})/F(\mathrm{H}\beta^{\mathrm{b}})$, of 3.4 in the WiFeS spectrum (we use superscripts 'b' and 'n' to refer to the broad and narrow components of a given emission line when such are clearly detected). Such a decrement is

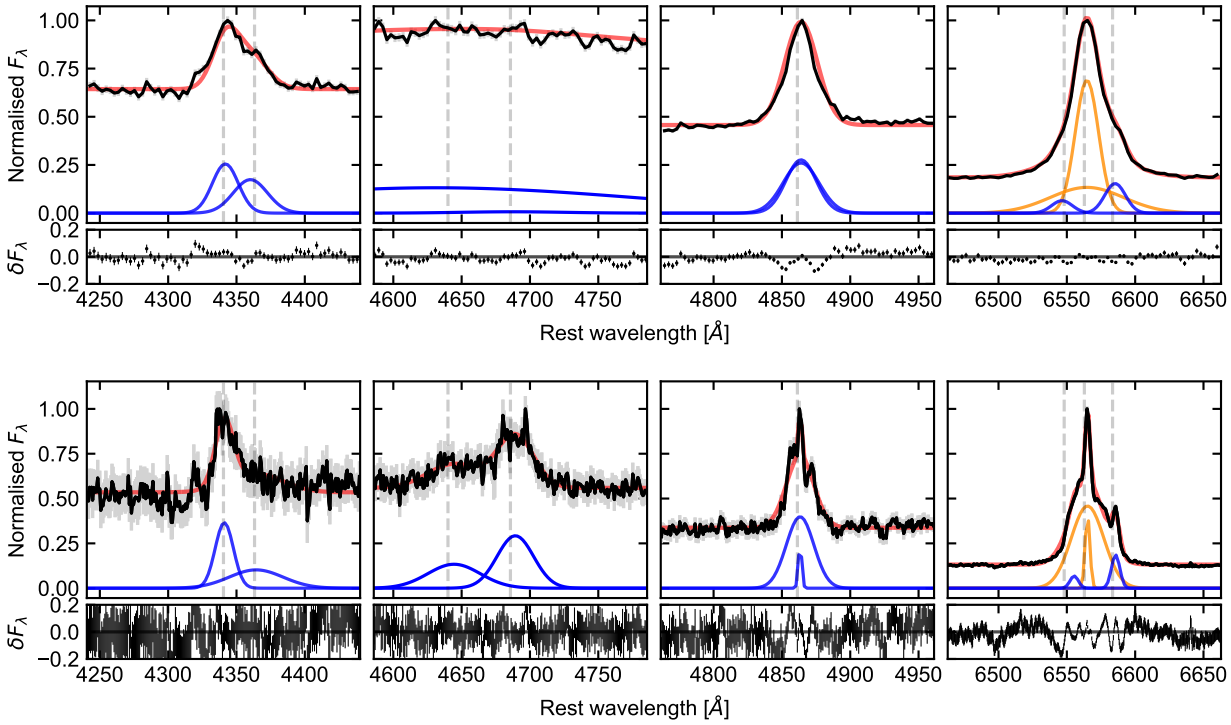


Figure 4.10: Zoomed-in plots of the main emission lines observed in both the NOT and WiFeS spectra (top and bottom panels respectively). The black line is the observed flux density, and the grey error bars are the associated uncertainties. We plot our fitted spectral model to the data for each region in red (including background component), whilst the blue and orange lines along the bottom represent the contribution of each source component to the fit (further described in Table A.2). The lower plots in each panel show the residuals in the spectral fitting, where $\delta F_\lambda$ is the difference between the observed $F_\lambda$ and the model $F_\lambda$, normalised by the model $F_\lambda$. We note that the double peaked appearance of the He II emission line in the WiFeS spectrum is most likely non-physical and due to the noisy optical spectrum, as no other broad lines show such similar line profiles.

Table 4.7: Emission line ratios from the WiFeS spectrum, where the narrow components were resolved. The superscript 'b' and 'n' denote the broad and narrow components, respectively.

| Line 1, Line 2 | F(Line 1)/ F(Line 2) |
|---|---|
| H$\alpha^{\rm n}$, H$\beta^{\rm n}$ | 5.8 ± 0.8 |
| H$\alpha^{\rm b}$, H$\beta^{\rm b}$ | 3.4 ± 0.1 |
| He II 4686, H$\beta^{\rm b}$ | 0.6 ± 0.1 |
| N III 4640, H$\beta^{\rm b}$ | 0.4 ± 0.1 |
| [Fe X], [O III] 5007 | 2.4 ± 0.3 |
| [Fe XIV], [O III] 5007 | 3.0 ± 0.5 |

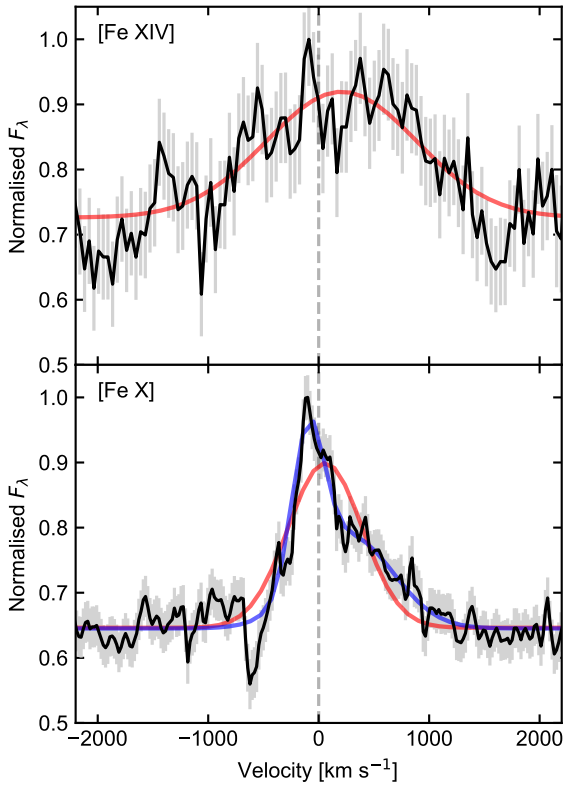


Figure 4.11: Best-fit single Gaussians (red) to the transient [Fe XIV] 5303 Å (top) and [Fe X] 6375 Å (bottom) coronal lines observed in the WiFeS spectrum. The lower ionisation line of the pair, [Fe X] 6375 Å, is more asymmetric, its broad base appears slightly blueshifted, and can also be fitted by a pair of Gaussians of FWHMs 330 ± 40 km s$^{-1}$ and 900 ± 100 km s$^{-1}$ (blue line), with $F$([Fe X] 6375)/$F$([O III] 5007) ∼ 2.6.

Table 4.8: Line widths inferred from the WiFeS spectrum.

| Line | FWHM [km s$^{-1}$] |
|---|---|
| N III 4640 | 2813 ± 648 |
| He II 4686 | 1959 ± 172 |
| H$\beta^{\rm n}$ | 173 ± 20 |
| H$\beta^{\rm b}$ | 1422 ± 11 |
| [O III] 5007 | 384 ± 80 |
| [Fe XIV] 5303 | 1558 ± 144 |
| [Fe X] 6375 | 768 ± 35 |
| H$\alpha^{\rm n}$ | 182 ± 3 |
| H$\alpha^{\rm b}$ | 1252 ± 9 |
| [N II] 6549 | 319 ± 12 |

consistent with what is observed in AGNs (e.g. Dong et al. 2005, 2007a; Baron et al. 2016), and is slightly higher than the predicted value of around 2.74-2.86[15] for case B recombination (Baker & Menzel, 1938) and thus a photoionisation origin. Whilst it was originally thought that the observed distribution in the Balmer decrements above 2.86 may have been due to a mix of collisional excitation and dust reddening in the centre of the host galaxy, several papers have suggested that the fundamental driver for this variance is the reddening (e.g. Dong et al. 2007a; Baron et al. 2016; Gaskell 2017). Dong et al. (2007a) find that after accounting for reddening, the intrinsic distribution of Balmer decrements in AGNs is well described by a log Gaussian of mean 3.06, with a 0.03 dex standard deviation, whilst a recent work by Gaskell (2017) find the intrinsic distribution is 2.72±0.04, and thus consistent with case B recombination.

Using these results, and by working on the assumption that the intrinsic Balmer decrement is set by Case B recombination to 2.86, we infer an $E(B-V) \sim 0.17$ and 0.65 mag from the broad and narrow Balmer emission lines respectively (using the Calzetti et al. 2000 extinction law)[16]. We note that the E(B-V) inferred from the Balmer decrement is larger than that inferred from SED fitting, which was performed on photometry that included light emitted from a larger region in the host galaxy than that probed by the Balmer decrement analysis.

### Bowen feature around 4680 Å

Both the FLOYDS and the WiFeS spectra show the emergence of a broad emission feature around 4680 Å, which is likely a blend of He II 4686 Å and N III 4640 Å. Although this feature overlaps with the 4400-4700 Å region, which can often show prominent Fe II emission in AGNs, we disfavour an Fe II origin here on the basis of no strong Fe II bump being observed from the strongest Fe II transitions in the 4500-4600 Å or ∼5150-5350 Å ranges (e.g. Kovačević et al. 2010). When comparing the WiFeS AT 2019avd spectrum to the composite SDSS quasar spectrum presented in Fig. 2 of Trakhtenbrot et al. (2019a), which was constructed from about 1000 SDSS quasars with broad Balmer lines of FWHM $\sim 2000\,\mathrm{km\,s^{-1}}$, the He II emission in AT 2019avd is much stronger relative to the Balmer emission in the AGN composite.

The N III 4640 Å emission suggests the presence of Bowen fluorescence (Bowen, 1928). He II Ly$\alpha$ photons at 303.783 Å are produced after recombination of $He^{++}$ [17], and can then either escape, ionise neutral H or He, or, because of the wavelength coincidence of O III 303.799 Å and 303.693 Å, be absorbed by O III. If the latter happens, then the later decay of the excited O III can produce a cascade of emission lines escaping the region (e.g. 3047, 3133, 3312, 3341, 3444, and 3760 Å[18]), and eventually a FUV O III 374.436 Å photon. The 374.436 Å can then be absorbed by ground-state N III, which further triggers a cascade of emission lines (N III 4100, 4640 Å). Bowen fluorescence typically requires a high flux of FUV/ soft-X-ray photons in order to produce the He II Ly$\alpha$ photons.

We measure relative line intensities of $F(\mathrm{He\,II})/F(\mathrm{H}\beta^{\mathrm{b}}) \sim 0.57$, $F(\mathrm{N\,III\,4640})/F(\mathrm{He\,II}) \sim$

[15]The predicted value is dependent on the assumed gas density and temperature.

[16]Alternatively, the inferred $E(B-V)$ values are 0.10 and 0.59 if we assume that the intrinsic Balmer decrement is 3.06 as in Dong et al. (2007a).

[17]The He II ionisation potential is 54.4 eV.

[18]Unfortunately, our spectra do not cover the 3000-4000 Å range to detect the other O III Bowen lines.

0.65 and $F(\mathrm{N\,III}\,4640)/F(\mathrm{H}\beta^{\mathrm{b}}) \sim 0.37$. Netzer et al. (1985) predicted the relative Bowen line intensities in AGNs under a range of different metal gas densities and abundances, where they found that to produce the high $F(\mathrm{He\,II})/F(\mathrm{H}\beta^{b})$ ratios seen in AT 2019avd as well as the high observed $F(\mathrm{N\,III}\,4640)/F(\mathrm{H}\beta^{b})$ ratio, the gas producing the Bowen fluorescence must have very high density ($n_{\mathrm{H}} > 10^{9.5}\mathrm{cm}^{-3}$) and high N and O abundances relative to cosmic abundances.

### Coronal lines

From the line fitting seen on the WiFeS spectrum in Fig. 4.11, we infer the luminosities of the [Fe X] 6375 Å and [Fe XIV] 5303 Å emission lines to be $\sim 2{\times}10^{39}$ and $\sim 3{\times}10^{39}$ erg s$^{-1}$. We also infer relative intensities of $F([\mathrm{Fe\,X}]\,6375)/F([\mathrm{O\,III}]\,5007) \sim 2.4$ and $F([\mathrm{Fe\,XIV}]\,6375)/F([\mathrm{O\,III}]\,5007) \sim 3$. Based on the coronal line ratio definitions proposed in Wang et al. (2012), AT 2019avd is classified as an extreme coronal line emitter (ECLE), where extreme is defined relative to the line ratios seen in coronal line AGNs (e.g. Nagao et al. 2000 report a maximum line ratio for $F([\mathrm{Fe\,X}]\,6375)/F([\mathrm{O\,III}]\,5007)$ of 0.24 over a sample of 124 Seyferts). Also, given the non-detected set of [Fe VII] emission lines in AT 2019avd which are seen in some ECLEs, and relatively weak [O III] 5007 Å emission, AT 2019avd belongs to the subset of ECLEs that were designated as TDEs in Wang et al. (2012).

The Fe coronal lines are narrower relative to the He II and N III 4640 Å emission lines (Table 4.8), with FWHM for the [Fe XIV] 5303 Å and [Fe X] 6375 Å of 1560 ± 140 and 770 ± 40 km s$^{-1}$ respectively. Under the assumption that the line widths are set by the virial motion of the gas, this suggests that the coronal lines are produced further away from the BH than the Bowen lines, and also with the higher ionisation coronal lines being produced closer to the BH than the lower ionisation lines. The width of [Fe XIV] 5303 Å is comparable to the observed Balmer emission. We also note the differing line profiles of the [Fe XIV] 5303 Å and [Fe X] 6375 Å emission, with the latter showing a stronger blue asymmetry (Fig. 4.11).

As discussed in Wang et al. (2012), the weakness of [Fe VII] emission relative to [Fe X] and [Fe XIV] may be explained through the coronal line gas either being overionised under a high X-ray flux, or due to collisional de-excitation of [Fe VII], because it has a lower critical density ($\sim 10^{7}$ cm$^{-3}$) compared with the higher ionisation lines ($\sim 10^{10}$ cm$^{-3}$, Korista & Ferland 1989).

### Black hole mass estimate

We assume that the gas that produces the broad H$\beta$ emission is virialised around the SMBH at the centre of the galaxy, and use the 'single epoch' mass-estimation technique (e.g. Vestergaard & Peterson 2006) to infer the black hole mass using the following scaling relation from Assef et al. (2011):

$$\log\left(\frac{M_{\mathrm{BH}}}{M_{\odot}}\right) = A + B\log\left(\frac{\lambda L_{\lambda}}{10^{44}\,\mathrm{erg\,s^{-1}}}\right) + C\log\left(\frac{\mathrm{FWHM}}{\mathrm{km\,s^{-1}}}\right), \tag{4.1}$$

with $A = 0.895$, $B = 0.52$ and $C = 2$. From the measured FWHM of the broad $H\beta$ component 1420 km s$^{-1}$ and $L_{5100} = \lambda L_{\lambda}(5000\text{Å}) \sim 2 \times 10^{42}$ erg s$^{-1}$ from the WiFeS spectrum[19], we then

[19] $L_{\lambda}(5100\text{Å})$ is computed from the mean of $L_{\lambda}$ between 5095 and 5105 Å

infer $\log[M_{\rm BH}/M_\odot] \sim 6.3$, albeit with a large uncertainty of ∼0.3 dex (Assef et al., 2011). We note that using this technique requires the correlations between continuum luminosity and radius of the broad line region (BLR; e.g. Kaspi et al. 2005) obtained in previous AGN reverberation mapping experiments to also hold for the BLR around the SMBH in AT 2019avd.

#### Baldwin, Phillips, and Terlevich line diagnostic

From the fitting of the WiFeS spectrum, we infer line flux ratios of $\log[[\text{N II}]\ 6583/\text{H}\alpha^{\rm n}] = -0.099^{+0.015}_{-0.016}$ and $\log[[\text{O III}]\ 5007/\text{H}\beta^{\rm n}] = 0.09^{+0.08}_{-0.10}$. According to a Baldwin, Phillips, and Terlevich (BPT) line diagnostic test (Baldwin et al., 1981), such line ratios suggest that a blend of star formation and AGN activity is responsible for producing the narrow line emission in the host galaxy of AT 2019avd (Kauffmann et al., 2003; Kewley et al., 2006). Without an archival spectrum though, it is unclear whether the [O III] 5007 Å and [N II] 6583 Å lines have increased in intensity since the initial ZTF outburst, or an AGN-like ionising source has always been present.

### 4.4.5 Mapping out the BLR

Assuming that each observed emission line is broadened due to its virial motion around the central BH, we can use the measured FWHMs to obtain rough estimates of the distances from the central ionising source at which each line is produced (Fig. 4.12). Similar to previous work (e.g. Korista et al. 1995; Kollatschny 2003; Bentz et al. 2010), we also find evidence for a stratified BLR, whereby the higher ionisation lines are produced in regions closer to the BH.

## 4.5 Discussion

Based purely on its X-ray luminosity evolution, AT 2019avd most likely involves an accreting SMBH at the centre of a galaxy. Whilst the large amplitude X-ray flaring (factor of $\gtrsim$600), soft X-ray spectrum, lack of previous strong (and sustained) AGN activity, and the implied unabsorbed X-ray peak luminosity in the 0.2 – 2 keV energy range of $2 \times 10^{43}$ erg s$^{-1}$ (using spectroscopic $z = 0.029$, see section 4.4.1) initially made the source a strong TDE candidate, this is clearly discordant with the double-peaked optical variability seen in the ZTF observations (it does not look like a prototypical, single-event TDE as observed elsewhere). In the following section, we discuss potential origins of the rich phenomenology seen in AT 2019avd.

### 4.5.1 AT 2019avd as non-TDE-induced AGN variability

If AT 2019avd is related to AGN activity that was not induced by a TDE (herein referred to simply as AGN 'activity' or 'variability'[20]), then the combination of its X-ray and optical light curves make it one of the most extreme cases of AGN variability observed to date.

[20] As a TDE may transform a quiescent BH into an AGN, the variability in BHs induced by TDEs is also just a subset of AGN variability.

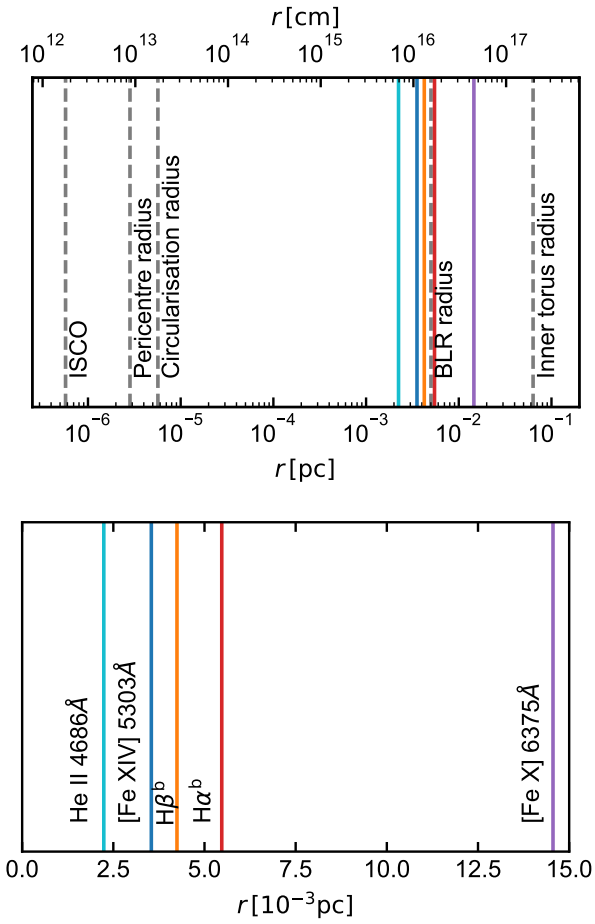


Figure 4.12: Estimated radii from the BH where different observed optical emission lines are produced in AT 2019avd, compared with various key physical length scales predicted in the literature (assuming $\log[M_{\rm BH}/M_{\odot}] = 6.3$). The pericentre and circularisation radii are computed assuming a Sun-like star incident on this BH with its closest approach at the tidal radius. Similarly to Kollatschny et al. (2014), we see evidence for a stratified BLR. The coloured lines represent length scales that were obtained based on observations of AT 2019avd, whilst the grey dashed lines are based on various scaling relations in the literature (BLR radius based on Kaspi et al. 2005, whilst the inner torus radius was computed using equation 1 of Nenkova et al. 2008, assuming a dust sublimation temperature of 1500K).

It is clear that the X-ray spectrum of AT 2019avd (section 4.2.3) is far softer than what is commonly seen in Seyfert 1s; for example, the power-law slope for *Swift* OBSID 00013495001 was $5.3^{+0.4}_{-0.4}$, whilst Nandra & Pounds (1994) model the observed power-law slope distribution with a Gaussian distribution of mean 1.95 and standard deviation 0.15. However, based on the measured FWHMs of the broad Balmer emission lines in the optical spectrum, it would be classified as a NLSy1, and softer spectral indices have also been observed in the NLSy1 population; a systematic *ROSAT* study of this by Boller et al. (1996) found power-law slopes of up to $\sim 5$. NLSy1s are also known to exhibit rapid, large-amplitude X-ray variability (e.g. Boller et al. 1996). As the X-ray variability of NLSy1s over longer timescales has not been extensively monitored before, how common AT 2019avd-like X-ray flares are within this population is currently unclear. For this reason, the X-ray properties alone cannot be used to state that the observed variability in AT 2019avd was induced by a TDE.

However, AT 2019avd shows a number of features in its optical spectrum that are infrequently seen in NLSy1s. First, NLSy1s commonly show strong Fe II emission (e.g. Rakshit et al. 2017), whereas this is not seen in the WiFeS spectrum, and only a weak Fe II complex is seen in the NOT spectrum in AT 2019avd. Instead, the most prominent Fe emission we observe are the transient, ECLE-like higher ionisation coronal lines of [Fe XIV] 5303 Å and [Fe X] 6375 Å in the WiFeS spectrum. During our spectroscopic follow-up campaign, we also observe the appearance of He II 4686 Å and N III 4640 Å emission lines (attributed to Bowen fluorescence). The optical spectrum at late times appears similar to the recently identified new class of flaring transients by Trakhtenbrot et al. (2019a), and we present a comparison of AT 2019avd with this class in Fig. 4.13. Whilst AT 2019avd shares the broad emission feature around 4680Å with the AT 2017bgt flare class, the optical spectrum of AT 2019avd is distinguishable from the other members based on its much weaker [O III] 5007Å emission line. A likely reason for this is that the host galaxies of the other flares had persistent, higher luminosity AGNs in them prior to the optical outburst, relative to AT 2019avd. In addition, AT 2019avd's large amplitude, ultra-soft X-ray flare, and its optical light-curve evolution make it unique amongst the AT 2017bgt flare class.

Finally, we stress that the double-peaked optical variability shown by AT 2019avd is unprecedented for a NLSy1, which when combined with its X-ray properties, make AT 2019avd clearly unique relative to all previous examples of AGN variability. Further examples of NLSy1 variability seen during the ZTF survey will be presented in a separate publication (Frederick et al., 2020).

### 4.5.2 An origin related to tidal disruption?

#### Canonical tidal disruption event

As AT 2019avd shows a very-large-amplitude, soft-X-ray flare from the nucleus of a galaxy that shows no strong signs of prior AGN activity, it appears similar to the predicted observational signatures for TDEs (e.g. Rees 1988) and most of the previous X-ray-selected thermal TDE candidates (Bade et al., 1996; Komossa & Bade, 1999; Komossa & Greiner, 1999; Grupe & Leighly, 1999; Greiner et al., 2000; Saxton et al., 2019). On the other hand, its optical spectrum

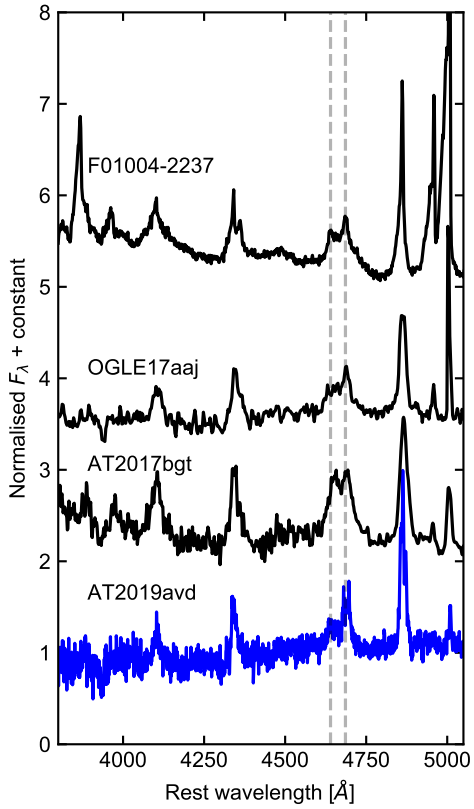


Figure 4.13: Comparison of the optical spectrum of AT 2019avd with those of the three nuclear transients recently identified as a new class of flares from accreting SMBHs in Trakhtenbrot et al. (2019a). The two dashed grey lines mark the positions of N III 4640 Å and He II 4686 Å. All objects share high $F$(He II 4686 Å)/$F$(H$\beta$), and at least one Bowen emission line (N III 4640Å).

shows a far weaker blue continuum component relative to that seen in optically selected TDEs, as well as narrower Balmer emission lines (for TDEs where these are detected); based on these two pieces of evidence, it would be straightforward to declare that AT 2019avd is not a TDE candidate, according to criteria for optical TDE selection in van Velzen et al. (2021).

The observed broad Balmer emission lines in AT 2019avd instead appear more like those commonly seen in the broad emission lines of Seyfert 1s. With such similarity, a mechanism analogous to the broad line emission in AGNs is likely operating in AT 2019avd, whereby the line widths of hydrogen recombination lines are set by the gas kinematics (whereas some TDEs may have line widths set by repeated non-coherent electron scattering; e.g. Roth & Kasen 2018), and the high densities in the BLR result in the line intensity responding effectively instantaneously to changes in the continuum flux. In the limit of a weak TDE-like reprocessing layer[21], the optical spectrum of a TDE may appear similar to that of an AGN, as has been previously suggested (e.g. Gaskell & Rojas Lobos 2014). The timescales for the evolution of the spectral features in such systems may be different from those observed in optically selected TDEs, as they originate from a region further away from the BH than the reprocessing layer.

The optical emission mechanism in TDEs is currently not well understood, although it is thought to arise either from shocks produced from stellar debris stream self-intersections (Shiokawa et al., 2015b; Piran et al., 2015), or from debris reprocessing the emission from an accretion disc (e.g. Loeb & Ulmer 1997; Ulmer et al. 1998; Roth et al. 2016; Roth & Kasen 2018). However, it is unclear how luminous the shocks are from stream self-intersections, whilst for the reprocessing scenario we still do not understand where the reprocessor is situated, where it forms, how large its covering angle would be from the BH, how efficiently it converts disc emission into the optical wavebands, or how all of these aspects are affected by the properties of the BH and those of the disrupted star. There is currently not a large enough sample of TDEs selected through *both* X-ray and optical surveys to test these various models of optical emission, and to properly assess the various complex underlying selection effects likely present in the existing TDE candidate population. A key example of these effects is the fact that only a small fraction of optically selected TDEs show transient X-ray emission ($\sim$ 25% of optically selected TDEs in van Velzen et al. 2021 were X-ray bright); Dai et al. (2018) suggested that the observed properties of a TDE may be dependent upon the viewing angle to the newly formed disc.

Given the above, and that there are also no X-ray selected, non-relativistic TDEs in the literature that have high-cadence optical photometric light curves available[22], we cannot rule out a TDE-related origin for AT 2019avd simply on the basis of a lack of optically selected TDE features in the optical spectrum. However, we do disfavour the canonical TDE interpretation (seen in optically selected TDEs) for this flare on the basis of the double-peaked optical light curve, which has not been observed in any of the TDEs identified by ZTF so far. Secondary maxima have previously been seen in the light curves of some TDE candidates (a compilation is presented in Fig. 8 of Wevers et al. 2019), though not at optical wavelengths and of far smaller amplitude increase compared with AT 2019avd (with the exception of the TDE in an AGN candidate in

[21] And likely a lack of optically-selected observed TDE features.

[22] Although the 4 X-ray bright TDEs in van Velzen et al. (2021) were monitored at a high cadence with ZTF and *Swift* UVOT, these were optically-selected TDEs.

Merloni et al. 2015).

**A more exotic variant of a tidal disruption event?**

A large fraction of stars may exist in binary systems (e.g. Lada 2006). Mandel & Levin (2015) studied the various outcomes of a binary star passing close to a SMBH from a nearly radial orbit. In $\sim 20\%$ of such approaches, a double tidal disruption event (DTDE) is produced, whereby both stars in the binary are disrupted in succession. These latter authors estimated that $\sim$10% of all stellar tidal disruptions could be associated with DTDEs, with such events expected to produce double-peaked light curves.

We can use the inferred rise-to-peak timescales from the ZTF light curves to test the feasibility of whether AT 2019avd may have been triggered by a DTDE, specifically for the case where each peak is associated with the rise to peak mass fallback of each successive disruption. Guillochon & Ramirez-Ruiz (2013) present the time taken for a single TDE to reach peak mass fallback rate (in their equation A2):

$$t_{\rm peak} = B_\gamma \left(\frac{M_{\rm BH}}{10^6 M_\odot}\right)^{1/2} \left(\frac{M_\star}{M_\odot}\right)^{-1} \left(\frac{R_\star}{R_\odot}\right)^{3/2} \text{ years}, \tag{4.2}$$

where $B_\gamma$ is a function of $\beta$, the ratio of the tidal radius of the BH to the pericentre of the orbit of the star, $\gamma$ is the polytropic index of the star[23], $M_{\rm BH}$ is the black hole mass, and $M_\star$ and $R_\star$ are the mass and radius of the star being disrupted.

Similarly to Merloni et al. (2015), we then generate a grid of $M_\star$ and $\beta$, log-uniformly between $(0.1 M_\odot, 100 M_\odot)$ and (0.5, 4), respectively, and compute $R_\star$ for each $M_\star$ using the mass–radius relationship for zero-age main sequence stars presented in Tout et al. (1996). For each possible combination of $M_\star$ and $\beta$, and for a black hole with $\log[M_{\rm BH}/M_\odot] \sim 6.3$, we check whether it can produce $t_{\rm peak}$ (using equation 4.2) within 20% of the observed peak timescales in the ZTF light curves ($\sim$ 24 days and $\sim$ 260 days for the first and second peak respectively). We also enforce the constraint that its tidal radius lies outside of the Schwarzschild radius for the system, so that it can produce a TDE with the star being swallowed whole by the black hole.

We plot the permitted regions of the $M_\star, \beta$ parameter space in red in Fig. 4.14, where we see that no main sequence binary star configuration can reproduce the observed rise times for both the first and second peaks. It would also be possible to obtain further constraints on the feasibility of this scenario based on the observed peak luminosities (similar to Merloni et al. 2015) and their ratio, as well as from the inferred properties of the binary itself, such as from the time between the two observed peaks (which could be used to constrain the semi-major axis) and the inferred mass ratio. However, the constraints provided from $t_{\rm peak}$ are perhaps the simplest to implement and are sufficient to highlight the caveats of a *simple* DTDE interpretation.

Bonnerot & Rossi (2019) recently suggested that following the disruption of a stellar binary, the two separate debris streams may collide prior to their fallback onto the black hole. These collisions then shock-heat the gas, and were predicted to produce an optical flare prior to the main flare of the disruption event. Such a model for a binary TDE could potentially explain the

[23] We use $\gamma = 4/3$ for $0.3 M_\odot < M_\star < 22 M_\odot$, and $\gamma = 5/3$ for $M_\star$ outside this range, as in Mockler et al. 2019.

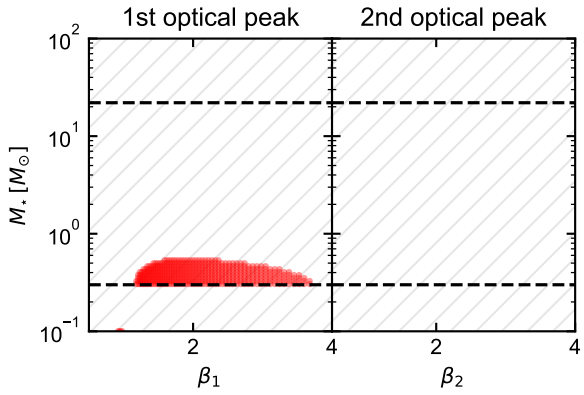


Figure 4.14: Constraints on the $M_\star$, $\beta$ parameter space, obtained for explaining the origin of AT 2019avd as a DTDE on SMBH. Red markers represent a permitted $M_\star$, $\beta$ configuration, whilst a region that contains grey hashing represents a configuration that is not able to reproduce the observed timescales for the given peak. Results were obtained for a black hole with $\log[M_{\rm BH}/M_\odot] \sim 6.3$. Since there are no red markers on the second optical peak plot, there is no permitted $M_\star$, $\beta$ pairing that can reproduce the observed peak timescale for the second optical peak. The black dashed lines bound $0.3M_\odot < M_\star < 22M_\odot$, where we adopt $\gamma = 4/3$.

observed double-peak light curve, and the observed emergence of the Bowen feature after the second peak (the soft X-rays can only be emitted once the accretion disc has formed). However, a caveat to this interpretation is that both a strong ionising flux and high gas densities are required for Bowen fluorescence to be produced, and we cannot confidently state here that the reason for not observing Bowen lines in the NOT spectrum is the absence of an X-ray-emitting accretion disc during that observation, because the absence of Bowen lines may also be due to insufficiently high gas densities (not all TDEs that are X-ray bright have displayed Bowen emission lines). We do not rule out this more complex DTDE scenario for AT 2019avd here, but do not perform a detailed comparison between the simulations in Bonnerot & Rossi (2019) and AT 2019avd in the present paper. Another alternative could be that AT 2019avd involved some type of TDE about a SMBH binary (e.g. Liu et al. 2009; Coughlin et al. 2017), where in such systems, the presence of the secondary BH can perturb the accretion flow onto the primary, leading to intermittent light curves.

### 4.5.3 Could AT 2019avd be supernova-related?

The spectra of Type IIn SNe can appear similar to those of AGNs (e.g. Filippenko 1989), as they can show broad and narrow emission lines, an absence of P-Cygni profiles, and higher luminosities and slower decay timescales relative to normal Type II SNe (Nyholm et al., 2020). Type IIn SNe typically also show the highest X-ray luminosities amongst all SNe. However, AT 2019avd has a $L_{0.2-2\rm keV}$ that is about an order of magnitude higher than what is seen in most X-ray-luminous Type IIn SNe, when considering the sample of IIn shown in Fig. 3 of Dwarkadas & Gruszko (2012). Furthermore, the X-ray emission from Type IIn SNe is predicted to be hard

(e.g. Ofek et al. 2013), whilst that of AT 2019avd is ultra-soft. Based on the X-ray emission alone, we disfavour the idea that both optical peaks in AT 2019avd are related to a single Type IIn supernova.

Given the observed peak and decay timescales (Fig. 4.6), the peak absolute magnitude of the optical light curve ($\sim -18.5$), the small amount of reddening seen in the ZTF light curve during the decay phase, and the NOT spectrum, the first optical peak may have been associated with a Type IIn SN. The second optical peak would then be associated with a 'turn on' event in the SMBH that sees a vast increase in the accretion rate and the luminosity of the BH. This scenario would then explain why the He II, Bowen, and coronal lines are not seen in the NOT spectrum, and only in the spectra taken after the second peak. However, the probability of observing both a Type IIn SN and an AGN 'turn on' event within just over a year of each other is extremely small given the apparent rarity of extreme 'turn-on' events in AGNs (especially those showing an AT 2019avd-like X-ray outburst) and the expected detection rates for Type IIn SNe (e.g. Feindt et al. 2019), and we therefore disfavour a scenario where AT 2019avd is the *chance* coincidence of a Type IIn SN and extreme AGN ignition event within roughly one year of each other.

## 4.6 Conclusions

This paper presents an overview of a set of multi-wavelength observations of an exceptional nuclear transient, AT 2019avd, whose main observed features are as follows:

1. eROSITA detected an ultra-soft ($kT \sim 85$ eV) X-ray brightening ($\gtrsim 90$ times brighter than a previous $3\sigma$ upper flux limit) from a previously X-ray-inactive galaxy (Section 4.2).

2. AT 2019avd was initially observed on a weekly basis with *Swift* XRT/UVOT for 6 weeks following the eROSITA detection. The host had brightened in all UVOT bands by $\sim 1$ mag relative to archival *GALEX* observations, and was observed with $0.2-2$ keV X-ray flux consistent with the eROSITA detection (Section 4.2). A further *Swift* observation ~5 months after the initial eROSITA detection revealed a brightening by a factor of approximately six in the $0.2-2$ keV band relative to the eROSITA detection. AT 2019avd therefore shows a net brightening in the $0.2-2$ keV band by a factor of *at least* 600 relative to the $3\sigma$ upper detection limit derived from an *XMM-Newton* pointing in 2015.

3. In the 450 days prior to the eROSITA detection, ZTF observed a double-peaked light curve (Section 4.3). The first optical peak shows rise and decay timescales akin to TDEs and SNe, whilst the rise time of the second peak is more similar to those seen in AGNs. No optical outbursts were detected during ASAS-SN observations over the seven years preceding the initial outburst seen by ZTF.

4. Optical spectroscopic follow-up finds transient He II emission, Bowen fluorescence lines, and high-ionisation coronal lines ([Fe X] 6375 Å, [Fe XIV] 5303 Å) in the spectra taken after the second optical peak, but not in the spectrum taken 30 days after the first peak. The presence of such a set of lines requires an intense source of soft X-ray emission and

extremely high densities. Broad Balmer emission lines were detected in spectra 30 days after the first peak in the ZTF light curve, as well as in all spectra taken in the weeks after the eROSITA detection with FWHM $\sim$ 1400km s$^{-1}$ (Section 4.4).

AT 2019avd thus shows a set of observed features which have never been observed together in the same nuclear transient before, and further complicates the non-trivial task of distinguishing the physical origin of large-amplitude variability seen in galactic nuclei. Whilst a discussion on the potential origins of this transient is presented in Section 5.5, it is still unclear what has triggered such exotic behaviour. Detailed simulations would be welcome to distinguish between the various possible scenarios. These will be well complimented with future planned observations (*Swift*, *NICER*, *XMM-Newton*) monitoring the late-time evolution of AT 2019avd. Finally, we note that during its eight successive all-sky surveys in the following years, eROSITA will systematically monitor the X-ray variability of AGNs and map out the population of nuclear transients. With this information, we will be able to better understand the extent of the X-ray variability shown by AT 2019avd, and make a more informed judgement on the origin of this transient.

# Chapter 5

# Classifying the Variable Sky of eROSITA

eROSITA detects a large number of transient and variable sources per day during its All-Sky Survey, originating from a diverse population of different astrophysical systems (see Chapter 1). When a flaring source is detected, it is highly desirable to quickly understand what potential source class has produced this flux increase, so that: i) one may promptly identify the less common source classes (e.g. cataclysmic variables) amongst a sea of contaminants (e.g. stellar coronal flares), ii) additional follow-up observations may be initiated whilst the source is still in outburst. If there is a relatively large delay between identification and classification, then there is a higher risk of observing the source after the outburst has finished, which may prevent a robust source classification, and also further studying the physics of the system. Furthermore, the manual vetting of all of eROSITA's variables, as well as 'doing science' on these alerts (e.g. triggering and organising follow-up campaigns; reducing observations; analysing, interpreting and reporting on interesting transients) would be extremely man-power demanding.

To address this problem, I developed a pipeline for the automatic identification and classification of transient and variable sources during the eROSITA All-Sky Survey, and present a detailed report of this pipeline in the following chapter. Adopting a supervised approach to classification, a random forest-based classifier was trained on a set of X-ray and multi-wavelength features (derived from the counterpart properties in the *Gaia* and *CatWISE* source catalogues). This trained classifier is then used to produce probabilistic classifications of X-ray variable sources discovered during the third eROSITA All-Sky Survey (eRASS3). The inclusion of multi-wavelength features boosts all-round classification performance, and are vital for an accurate and precise automated source classification. Finally, the deployment of this pipeline and its application during eRASS3 is discussed, and examples of new, interesting variable systems uncovered by the pipeline are presented.

## 5.1 Introduction

eROSITA (Predehl et al., 2021) is expected to detect several million X-ray sources over the course of its 4 year All-Sky Survey (eRASS), bringing X-ray astronomy into the 'big data' domain. For the large number of variable and transient sources within this population, it will be

important to classify each source promptly after detection, and identify those for which time-critical follow-up observations will be most beneficial.

Previous wide-field surveys generally successfully relied on source classification via human visual inspection, since coordinated teams of trained astronomers were usually able to accurately classify the detected source populations[1]. However as the size and complexity of the datasets to be analysed grows ever larger with newer surveys, the effectiveness and efficiency of human-based classification will drop, due to the increase in delay between source detection and identification potentially causing missed follow-up opportunities for the most 'interesting' sources. Optimising scientific exploitation of eROSITA in the time domain, and also the next generation of wide-field surveys such as the Vera Rubin Observatory Legacy Survey of Space and Time (LSST; Ivezić et al. 2019), thus requires the development of new, faster, accurate automated classification tools to be used. The popularity of machine learning algorithms amongst astronomers for enabling this has grown vastly in recent years (see Baron 2019 for a recent overview of the field).

ML techniques can be broadly divided into *supervised* and *unsupervised* learning methods (Fig. 5.1). In the former, a *training* dataset is used, along with a machine learning algorithm, to learn a function that maps from a given input to a known target output. This learnt function is then used to predict the output for each object in a previously unseen dataset (the *test* set, for which the output is unknown). The nature of the target output is dependent on the task performed; for classification it is typically a class label (or a probability of belonging to a given class), and an estimated value for regression problems. Unsupervised learning has no target outputs for each object in the dataset. Instead, it aims to discover patterns and structures present in the dataset, such as through dimensionality reduction (Boroson & Green, 1992; Krone-Martins & Moitinho, 2014), clustering (Baron et al., 2015; Garcia-Dias et al., 2018), data visualisation or anomaly detection (Protopapas et al., 2006; Baron & Poznanski, 2017; Reis et al., 2018).

The time variable emission from an object depends on the type of the underlying astrophysical system. With wide field surveys returning increasing numbers of light curves, efforts into astronomical time series classification have increased in recent years, with ML techniques found to have strong classification performance. This is largely due to their speed and ability to identify patterns in complex datasets, potentially imperceptible to humans. The majority of approaches use supervised classification, such as in variable star classification (Castro et al., 2017; Naul et al., 2018), general variable source classification (Lo et al., 2014; Farrell et al., 2015), variable versus transient source classification (Bloom et al., 2012), supernovae light curves (Lochner et al., 2016; Charnock & Moss, 2017; Revsbech et al., 2018; Möller & de Boissière, 2020), or general transient classification (Sooknunan et al., 2021; Muthukrishna et al., 2019).

In the following, I present a detailed report of a general ML-based pipeline for transient and variable source classification during eROSITA's All-Sky Survey[2]. In Section 5.2, the procedure for selecting the transients and variable sources within the source catalogues generated during the live eROSITA sky survey is described. In Section 5.3, the approach to producing auto-

[1] Although previous surveys may have used optical spectroscopy for source identification, the process of inspecting the spectra would ultimately have involved a human manually classifying the source.

[2] This pipeline is only applied to the German half of the eROSITA sky.

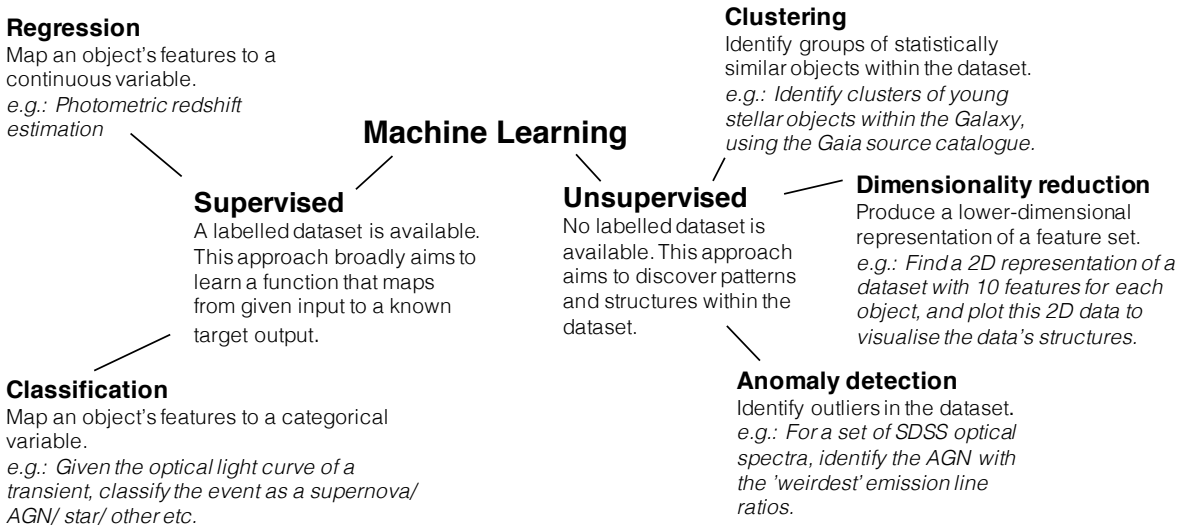


Figure 5.1: Overview of the main approaches of machine learning with example applications of each within astronomy.

matic classifications of the eROSITA variable sources is presented, including details on training data set generation, classifier training, and classifier performance. In Section 5.4, I discuss the deployment of this trained classifier during the eROSITA survey (to produce automated classifications of the variables and transients identified using the process described in Section 5.2), an overview of operations, limitations of the current classifier set-up, as well as examples of interesting sources which were successfully classified during eRASS3. Finally, I conclude in Section 5.5 with a summary of this work and a discussion on further steps.

## 5.2 Transient and variable source identification

During its All-Sky Survey, the SRG spacecraft is in contact with the ground stations only once per day over a ~4 hour window, where it transmits its telemetry collected over the previous ~24 hours. This telemetry is then received by the Russian Space Institute (IKI), is preprocessed, and is then sent on to the Max Planck Institute for Extraterrestial Physics for a standard data processing (using the *eROSITA Science Analysis Software*, eSASS; Brunner et al., in prep.), and also to the *Near Real Time Analysis* pipeline (NRTA; Kreykenbohm et al., in prep.), developed and operated by the group of J. Wilms in Bamberg, Germany. The main design motivations for the NRTA were to provide an interface for both eROSITA engineers and operators to monitor the health of the instrument based on the recently received telemetry, and also for *promptly* identifying the brightest X-ray transients seen by eROSITA that could need short reaction times between observation and follow-up. In this work, I focus on the transient and variable sky uncovered by the eSASS pipeline (as opposed to the NRTA) and an overview of the pipeline is presented in

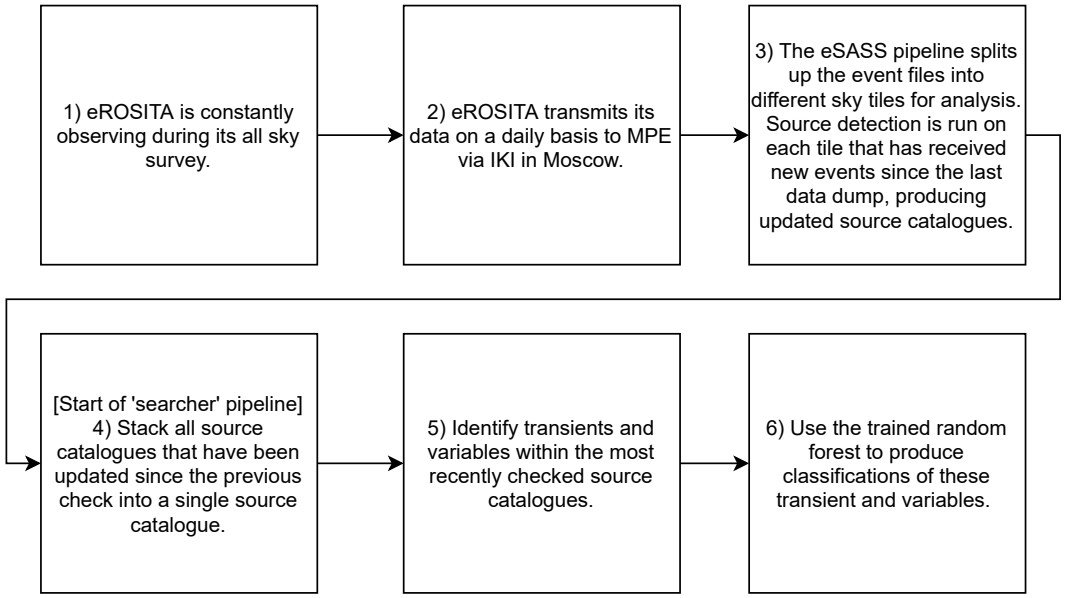


Figure 5.2: Overview of the transient and variable source identification pipeline. Further details on the process of identifying transients and variables can be found in Section 5.2, whilst details on the machine-learning used in this work are presented in Section 5.3.

Fig. 5.2.

The eSASS pipeline produces science data products (e.g. merged and calibrated event files, images, exposure maps, sensitivity maps, source catalogues, spectra, light curves) from the processed telemetry. For computational purposes, the observational data is split up into 4700 different sky tiles for analysis (each being $3.6° \times 3.6°$), with the eSASS pipeline being run on all events within a given eRASS for each tile. On roughly a daily basis, the most recent data dump is analysed by the eSASS pipeline, and the source catalogues for tiles which have received further observation since the previous dump are updated, with an example of this process for a single tile presented in Figure 5.3.

The `searcher` pipeline looking for transient and variable sources is currently run once per week[3]. The first step of the `searcher` is to produce a stacked catalogue of all source catalogues which have been updated since the previous `searcher` run (with duplicate sources lying in the overlap regions between neighbouring sky tiles cleaned). This stacked catalogue thus contains all source detections within the current eRASS, and is then cross-matched (using a $20''$ matching radius[4]) with a catalogue containing the eRASS1x2[5] sources (see Fig. 5.4 for an overview on creating this catalogue). The latest detected sources in eRASS3 are then split off into (potential) variables and transients, depending on whether they have or do not have a counterpart in the eRASS1x2 catalogue (i.e. have a previous detection by eROSITA or not). The above steps

[3]The pipeline is run autonomously using `crontab`.

[4]This value was chosen as it exceeds the typical point source uncertainty during an eRASS scan, but is still lower than the typical separation between a pair of nearest-neighbour point sources.

[5]We use the 'x' in catalogue names in this work to denote a cross-match between two catalogues.

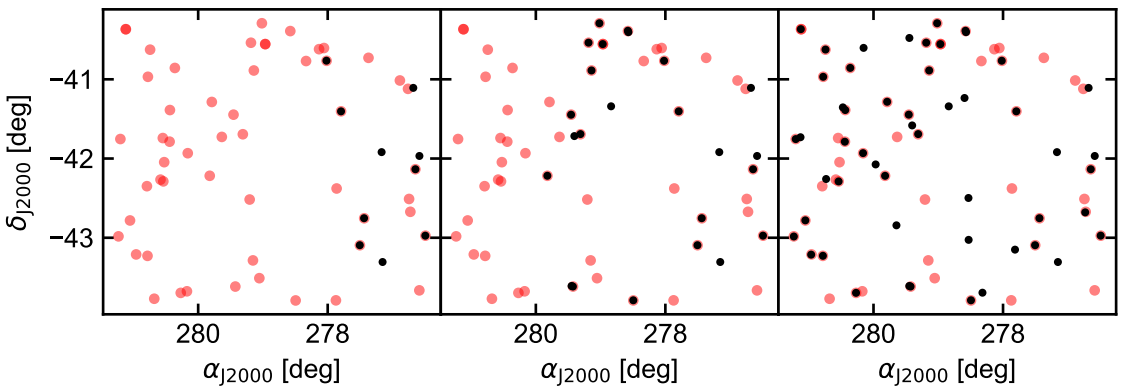


Figure 5.3: Example updating of the source catalogue for a single sky tile during the progression of eRASS3. Each black marker represents a detected source in the eSASS3 catalogue for this tile, and each panel shows the source detections for different versions of the catalogue of the tile (with the earliest and latest versions in the left and right panels respectively). The red markers show the detected sources in this sky region in eRASS2. Both eRASS2 and eRASS3 source detections have been filtered here to only include sources with `DET_LIKE` above 15 (for plotting clarity). In the left panel, the source detection is run at a time when eROSITA has only partially observed the tile, so sources are only detected at lower right ascension. As eRASS3 progresses, an increasing fraction of the tile is observed by eROSITA, such that an increasing number of sources are detected within the tile. The `searcher` pipeline monitors the updating of all tiles within the available eROSITA footprint during the survey, and is used to select out the most interesting transients and variables as soon as they are detected.

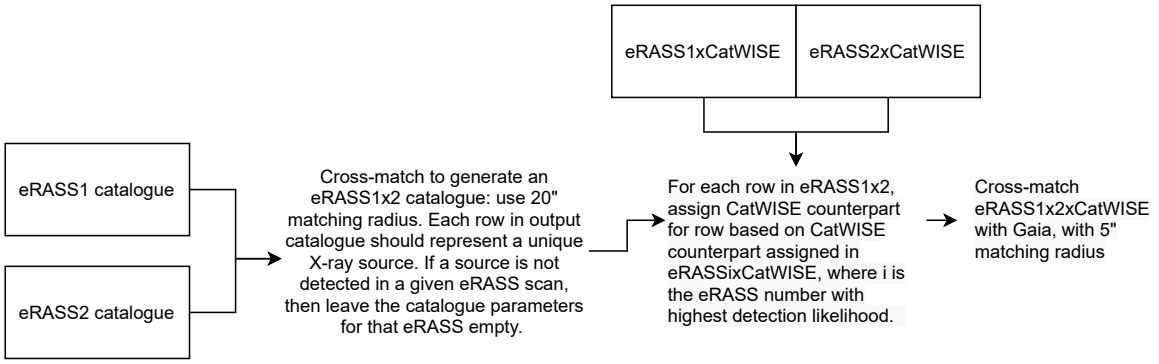


Figure 5.4: Flowchart for the generation of an eRASS1x2xCatWISExGaia catalogue. The cross-matching to produce the eRASS1x2 catalogue is performed using a positional cross-match here (although the eRASS1xCatWISE and eRASS2xCatWISE catalogues are generated through NWAY, see text for details), and utilises all eRASS1 and eRASS2 point sources without applying a cut in detection significance.

describe the procedure for transient and variable source identification in eRASS3 and will change over the course of the survey; in the $N$th eRASS, the stacked catalogue of newly detected sources will be cross-matched with a catalogue containing eRASS1x2x...($N-1$).

Given that the majority of the sources that eROSITA detects will be consistent with an approximately constant count rate between successive eRASS scans, a sub-selection on the potential variables is then performed to only include sources which have undergone 'large' flux changes between successive eRASS. This is done based on the variability flags described in Table 5.1, where variables are identified as sources which have at least one variability flag set to 1. In the following, I primarily focus on brightening events, largely because such systems are more rewarding to follow-up with other facilities, and the follow-up efforts may be more likely to catch the event while it is still in outburst. In future work, I will also implement variability flags that search for systems that show a significant fading of their X-ray flux over the course of the eRASS.

## 5.3 Automated classification

For eROSITA's variables and transients, a supervised approach to classification is adopted, where the *random forest* (RF) algorithm (Breiman, 2001) is used to learn a mapping from a set of input features to a class label for each source. Using this learnt mapping, the trained RF can then be used to make classifications of a previously unseen data set. Whilst a wealth of different learning algorithms exist, RFs often demonstrate very strong performance at classification tasks (Lochner et al., 2016), and in recent years have become widely used within astronomy (Goldstein et al., 2015; Lo et al., 2014; Revsbech et al., 2018; Decker French & Zabludoff, 2018; Ishida et al., 2019). In this work, a classifier is first trained on a training set derived from the eRASS1x2 catalogue, and this trained classifier is then applied to transients and variables identified in eRASS3.

Table 5.1: Variability flags computed from the eRASS1x2x3 source catalogue used to quantify whether a source detected in eRASS3 (the most recent eRASS) might show interesting variability behaviour that warrants a prompt automated classification. The value of each flag is set to 1 if the 'Criterion' is satisfied, and 0 otherwise. Fluxes in the 0.2-2.3 keV band for a given eRASS are computed from summing the source's eSASS-inferred maximum likelihood fluxes in the 0.2-0.6 keV and 0.6-2.3 keV bands from that eRASS (this also applies throughout this chapter, unless stated otherwise.)

| Variability Flag | Criterion |
|---|---|
| FLUX_3_over_2_BAND_SOFT | The source shows a flaring by a factor greater than 5 in the 0.2-2.3 keV energy band between eRASS3 and eRASS2 (where the flaring is computed using [eRASS3_ML_FLUX_1+eRASS3_ML_FLUX_2] / [eRASS2_ML_FLUX_1+eRASS2_ML_FLUX_2], with ML_FLUX_1 and ML_FLUX_2 being the estimated source flux in bands 0.2-0.6 keV and 0.6-2.3 keV). The source may or may not be detected in eRASS1. |
| FLUX_3_over_1_BAND_SOFT | The source shows a flaring by a factor greater than 5 in the 0.2-2.3 keV energy band between eRASS3 and eRASS1. The source may or may not be detected in eRASS2. |
| ON-OFF-ON | The source was detected in eRASS1 and eRASS3 with a 0.2-2.3 keV flux above $10^{-13}$ erg s$^{-1}$ cm$^2$, but was not detected in eRASS2. The eRASS1 and eRASS3 fluxes are much higher than the typical point source detection sensitivity in the 0.2-2.3 keV band of $5.4 \times 10^{-14}$ ergs$^{-1}$ cm$^{-2}$ for eRASS1 in an equatorial sky region (Predehl et al., 2021). Assuming a similar detection sensitivity in eRASS2 to eRASS1, then the source should also have been detected in eRASS2 if its eRASS1 and eRASS3 fluxes were above $10^{-13}$ erg s$^{-1}$ cm$^2$, unless it was highly variable (or happened to lie in a region of particularly lower exposure in eRASS2). |
| BRIGHTEN_LAST_3_ERASS | The source shows a brightening from eRASS1 to eRASS2 to eRASS3 in the 0.2-2.3 keV energy band (i.e. $F_{0.2-2.3,\,\mathrm{eRASS1}} < F_{0.2-2.3,\,\mathrm{eRASS2}} < F_{0.2-2.3,\,\mathrm{eRASS3}}$. |

In the following, I first briefly discuss the RF algorithm in section 5.3.1, before describing the generation of a training dataset and its properties in section 5.3.2. Details of the training process are then presented in section 5.3.3, and the metrics used for assessing classifier performance and choosing an optimal set of features in section 5.3.4.

### 5.3.1 The random forest algorithm

A RF is an ensemble of *decision trees* (DT). In the training process, *bootstrap sampling* is used to create randomly chosen subsets of the training data set which are then used to construct each DT. Each DT is built in a top-down manner - beginning with the the root node of the tree, the input dataset is split in two depending on whether the value of its $i$-th feature is above or below a given threshold for that feature. This feature threshold is set based upon finding an optimal splitting of the data set, where a search is performed through feature space to find a threshold that maximises the separation between the source classes according to the specified splitting criterion (typically the *Gini impurity* or the *information gain*). Objects with their $i$-th feature above (below) the threshold are sent to the right (left) child node; each node therefore splits the data set up into two purer subsets and is associated with a criterion for splitting the dataset. This splitting process is then applied in recursion to the child nodes until a specified stopping criterion is reached, such as when the node becomes a *leaf* and contains only objects of a single class.

Once trained, a new object can then be propagated through each tree in the forest (using the successive splitting condition on each node) until it reaches a leaf of the tree. The object is then classified based upon the class label of that given leaf. Final classification probabilities can be obtained for each source based upon the number of times each object was classified as belonging to a given class, divided by the total number of decision trees. If needed, the object can be given a single classification by choosing the class with highest probability (effectively a majority vote across all trees). The bootstrap sampling of the training data set during training decreases the correlation between the different trees, such that the RF is more robust to noise in the training data and offers far better generalisation performance over single DTs between the training and test data sets.

### 5.3.2 Generating an initial training data set

At its simplest, a training dataset was generated through cross-matching the eRASS1x2 source catalogue with multiple catalogues of known sources (Table 5.2), and stacking these into a single catalogue of labelled eROSITA sources. However, to reduce the number of mislabelled sources in the training dataset, and to improve the accuracy of the multi-wavelength information on each X-ray detected source (which is essential for effective source classification), this cross-matching was run as follows.

Individually for the eRASS1 and eRASS2 source catalogues (i.e. not the eRASS1x2 catalogue), the *CatWISE* (Eisenhardt et al., 2020) mid-infrared counterparts of each detected point source were identified using the Bayesian cross-matching tool `NWAY` (Salvato et al. 2018, Salvato et al., in prep.). For each source in the primary catalogue (the eRASS source catalogue here), `NWAY` returns one or more potential counterparts in the secondary catalogue (*CatWISE*),

Table 5.2: Overview of the catalogues of 'known' (or very strong candidate) sources used for constructing the initial training dataset. The class balance of the training dataset is presented in Fig. 5.5.

| Label | Description |
|---|---|
| AGN | Objects classified as Active Galactic Nuclei (AGN) in the 13th edition of the Véron-Cetty & Véron (2010) catalogue. |
| QSO | Objects classified as Quasi-Stellar Objects (QSO) in the 13th edition of the Véron-Cetty & Véron (2010) catalogue, where a QSO is distinguished from an AGN therein as a star-like nucleus with absolute magnitude $M_B = -22.25$ mag. |
| Star | The list of main-sequence stars reported in Smart et al. (2021), and sources reported in Marton et al. (2019) that have a probability of being a main-sequence star larger than 0.75 (further details on the computation of this probability are presented in Marton et al. 2019). In addition, the ultra-cool dwarf candidates listed in Table 2 of Reylé (2018) are included. |
| YSO | Sources reported in Marton et al. (2019) that have a probability of being a Young Stellar Objects (YSO) larger than 0.75 (further details on the computation of this probability are presented in Marton et al. 2019). |
| XRB | Catalogue of known Low-mass and High-mass X-ray binaries, provided via the eROSITA Compact Object Working Group. Low-mass XRBs are taken from the Ritter & Kolb (2003) and Liu et al. (2007) catalogues, whilst the High-mass XRBs were taken from Liu et al. (2006). |
| CV | Compilation of known CV systems (with and without orbital periods), provided via eROSITA Compact Object Working Group (Schwope, priv. comm.). |

and also allows for the possibility that no suitable counterpart in the secondary catalogue exists. The possible counterparts for each primary source are also graded based on the value of their match_flag parameter: each counterpart in the secondary catalogue with the most likely association has match_flag== 1 (match_flag also equals 1 for the cases where no good *CatWISE* counterpart exists[6]), and secondary sources with weaker associations have match_flag== 2. For simplicity in this work, we reduce the eRASS1xCatWISE and eRASS2xCatWISE catalogues to only include the best counterparts for each source (match_flag== 1). This means that the *CatWISE* information for each source is either of the best matching counterpart, or is missing for the case of there being no suitable counterpart. Of the eRASS sources which have a match in the *CatWISE* catalogue, $\sim$ 10% have at least one secondary counterpart in the *CatWISE* catalogue (in both eRASS1 and eRASS2).

Then, for each row in the eRASS1x2 catalogue, a *CatWISE* source is assigned, based upon the best counterpart selected in the eRASS with the highest detection likelihood[7], DET_LIKE, since i) in the case of a non-detection of a source in eRASS1 or eRASS2, the counterpart should be assigned in the eRASS1x2 catalogue using the NWAY counterpart of the source when it was detected (i.e. highest detection likelihood), and ii) a more robust association should be *ideally* obtained for such cases. For example, for a source detected in eRASS1 and eRASS2 with DET_LIKE 10 and 200 respectively, we would assign a counterpart for it in the eRASS1x2xCatWISE catalogue based on the best counterpart assigned in the eRASS2xCatWISE catalogue. Similarly, for a source detected in eRASS1 but not in eRASS2, we take the counterpart in the eRASS1x2xCatWISE catalogue based on the eRASS1xCatWISE catalogue. An eRASS1x2xCatWISExGaia catalogue is then created through cross-matching the eRASS1x2xCatWISE catalogue with *Gaia* DR3 using a 5″matching radius and the BEST_MATCH_RA and BEST_MATCH_DEC columns from the eRASS1x2xCatWISE catalogue. If there is a *CatWISE* counterpart assigned for a source, then the BEST_MATCH_RA and BEST_MATCH_DEC values are set to the RA and DEC of the *CatWISE* source position, whereas for sources without a *CatWISE* counterpart, BEST_MATCH_RA and BEST_MATCH_DEC are set to the estimated eSASS position of the source in the highest detection likelihood eRASS (only one possible *Gaia* counterpart may be assigned per unique X-ray source). Since the BEST_MATCH_RA and BEST_MATCH_DEC columns contain a mix of eROSITA and *CatWISE* source positions, the 5″ matching radius used for the *Gaia* cross-match was chosen as a trade-off between the matching radii that would be appropriate for the *CatWISE* and eROSITA source catalogue positions.

A labelled catalogue of eROSITA detected sources is then created through cross-matching the eRASS1x2xCatWISExGaia catalogue (using the BEST_MATCH_RA and BEST_MATCH_DEC columns and a 5″ matching radius) with each of the catalogues of known sources listed in Table 5.2. This initially produces a set of 50920 sources (61 XRB, 485 CV, 889 YSO, 6503 stars, 19042 AGN and 23940 QSOs). I then only include sources in the training set that have a *CatWISE* counter-

[6]These two cases can be distinguished between based on NWAY's ncat parameter, which is equal to 1 if no counterpart was identified in the secondary catalogue, but equal to 2 if a counterpart was identified in the secondary.

[7]The *CatWISE* counterparts have only been assigned so far with NWAY for the separate eRASS catalogues (this work is performed by the catalogue cross-matching team within the eROSITA consortium). For this reason, we resort to first matching the eRASS1 and eRASS2 catalogues, and then assigning counterparts in this eRASS1x2 catalogue based on the eRASS1xCatWISE and eRASS2xCatWISE catalogues described above.

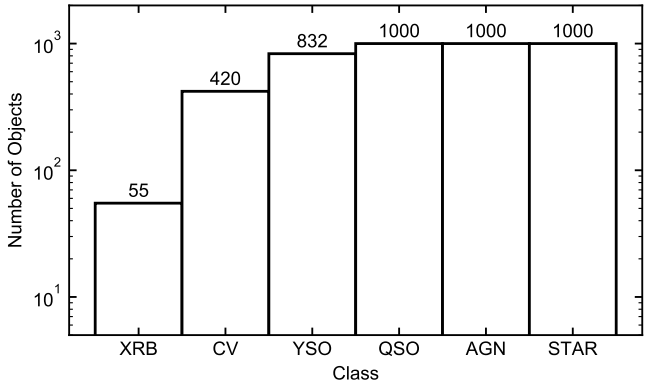


Figure 5.5: Class balance of the initial labelled training dataset used for training the random forest; the number of sources in each class is shown above each class's bar.

part with $p_i > 0.5$, where $p_i$ is the relative probability of an association being the true counterpart across all possible counterparts (provided by NWAY). Furthermore, all sources in the AGN and QSO catalogues which have a *Gaia* counterpart with significant parallax or proper motion (`PM_SIG` or `PLX_SIG`$>3$, see definitions in Table 5.3), are removed from the training dataset, as such measurements for these source classes will be spurious (i.e. these objects should be effectively stationary or too distant for a significant motion or parallax measurement). Finally, to reduce the significant class imbalance in the training set, we randomly drop out AGN, QSO and stars from the training set to obtain the final class balance plotted in Fig. 5.5. This undersampling of the AGN, QSO and stars is done as an *extreme* class imbalance in the training dataset can worsen classification performance on the minority class, if not accounted for (further discussion in Section 5.5.2). The number of sources in each undersampled class is still large enough for the classifier to learn the key characteristics of each source class, based on the classifier's performance reported in Section 5.3.4. The labelled source catalogue is subsequently used for training the random forest classifier. The training set described in this section is used during the initial phases of the eRASS survey, but will be updated and revised as the survey progresses and new, larger sets of labelled sources are identified by the eROSITA consortium.

For each source, we construct a set of features[8] derived from the eROSITA source catalogues, and a set of multi-wavelength features for each object based on the *CatWISE* and *Gaia* counterparts assigned to each source (all considered features are listed in Table 5.3). The main eROSITA features are the three 'X-ray colours' derived from the ratios of the eSASS inferred fluxes in the energy bands 0.2-0.6 keV, 0.6-2.3 keV and 2.3-5keV, to the flux in the 0.2-5 keV band[9], and the

[8] 'Features' refers to a set of properties of a source, using terminology from the machine learning literature here.

[9] Whenever a flux ratio that involves an eROSITA flux is computed for the training set, source parameters in the eRASS with the highest detection likelihood are used (assuming that the source will be in a brighter state when the detection likelihood is higher). For instance, if the detection likelihood is higher in eRASS2 than eRASS1 for a source, then the `log_ML_FLUX_0_over_W1_FLUX` ratio would be computed using the `ML_FLUX_0` value from the eRASS2 catalogue. The brightest eRASS flux is used, instead of an average flux across the two eRASS, because the classifier is currently applied to flaring X-ray sources (Table 5.1), and should be trained on flux ratios of the source when in outburst where possible.

absolute value of the source's galactic latitude. The multi-wavelength features used are a set of fluxes, flux ratios between different energy bands, *CatWISE* variability features, magnitudes, and the proper motion and parallax significance.

No X-ray variability information is used in this work for training the classifier, such as the flux variability amplitudes between the two successive eRASS scans, primarily due to issues associated with computing such amplitudes for cases where a source is not detected in a given eRASS. Without a detection, one can only compute the flare amplitude relative to a flux upper limit at the source's position. The major obstacle here was that such upper limits were not available for each source in the eRASS1x2 catalogue when a source was detected in only a single eRASS, such that no flare amplitudes could be computed. A second potential issue stems from eROSITA's scanning pattern during the eRASS, which produces a highly non-uniform exposure map over the whole sky (Predehl et al., 2021). This results in deeper upper limits on the source flux being obtained in sky regions with greater exposure times, such that the variability amplitude could strongly correlate with the source's position on the sky and not necessarily with the actual physical variability of the source. The flux variability shown by the source is thus currently only used for selecting interesting variables in the latest eRASS (section 5.2), and I show in section 5.3.4 that the classifier performs well even in the absence of X-ray variability features.

Fig. 5.6 shows the *CatWISE* W1-band apparent magnitude against the $F_{0.2-2.3\,\mathrm{keV}}$ flux for sources in the training data set. Also plotted is the relation between W1 and $F_{0.2-2.3\,\mathrm{keV}}$ suggested in Salvato et al. (2018) for separating AGN from stars; whilst this relation holds over $\sim$ 3 orders of magnitude in X-ray flux, additional features are still needed to break up the degeneracy between the different source classes. AGN, QSOs, CVs and XRBs are characterised by higher $F_\mathrm{X}/F_\mathrm{opt}$ ratios relative to stars and YSOs (Fig. 5.7). AGN are typically optically redder than QSOs at a given `W1-W2` (Fig. 5.8), and objects without a significant parallax in the `W1-W2` $< 0.5$ mag region are most likely lower luminosity AGN, with weak hot dust emission from an obscuring torus structure, as envisioned in various AGN unification schemes (see Padovani et al. 2017 for a recent review), relative to their host's emission in these wavebands. For a given object with an AGN-like $F_\mathrm{X}/F_\mathrm{opt}$ ratio in this `BP-RP` region, CVs generally occupy lower `W1-W2` values relative to QSOs (Fig. 5.8). Whilst both YSOs and stars generally show significant parallaxes, YSOs may be distinguished based on their redder mid-IR colours at a given $F_\mathrm{X}/F_\mathrm{W1}$ ratio (Fig. 5.9), stemming from YSOs residing in more dusty environments relative to stars, and tend to also show higher $F_\mathrm{X}/F_\mathrm{opt}$ values at a given $F_\mathrm{X}/F_\mathrm{W1}$ (Fig. 5.10). Visualising the training set in a Hertzsprung-Russell diagram (Fig. 5.11), one sees the CVs generally occupying the valley between the main-sequence and isolated white dwarf stars. Whilst a subset of CVs and XRBs may overlap with the main-sequence in this diagram, such systems may be distinguished based on their much higher $F_\mathrm{X}/F_\mathrm{opt}$ ratios.

### 5.3.3 Training the classifier and feature selection

To identify an optimal feature set for training the classifier, 36 different feature set versions were created from various subsets of the features listed in section B.1, with these sets including i) only-eROSITA based features, ii) eROSITA + *CatWISE* features, iii) eROSITA and *Gaia* features, and iv) eROSITA, *Gaia* and *CatWISE* features. For each feature set version, a 5-fold cross-validation

Table 5.3: Overview of the different features considered during the training of the classifier. ML_FLUX_*i* is the eSASS estimated source flux in band *i*, where band numbers 0, 1, 2, 3 refer to the 0.2-5, 0.2-0.6, 0.6-2.3, 2.3-5 keV energy bands respectively.

| Feature Name | Description |
|---|---|
| ABS_BII | Absolute value of the source's Galactic latitude. |
| log_X_RATIO_1 | log((ML_FLUX_0 - ML_FLUX_1) / ML_FLUX_0) |
| log_X_RATIO_2 | log((ML_FLUX_0 - ML_FLUX_2) / ML_FLUX_0) |
| log_X_RATIO_3 | log((ML_FLUX_0 - ML_FLUX_3) / ML_FLUX_0) |
| ML_CTS_3 | eSASS estimated number of source counts in the 2.3-5 keV energy band. |
| ML_EXP_3 | eSASS estimated effective exposure time for the source during an eRASS in the 2.3-5 keV energy band. |
| W1mag | *CatWISE* counterpart's W1 magnitude. |
| W2mag | *CatWISE* counterpart's W2 magnitude. |
| W1-W2 | W1mag - W2mag |
| log_ML_FLUX_0_over_W1_FLUX | Log of the flux ratio between the 0.2-5 keV flux and the *CatWISE* counterpart's W1-band flux. |
| log_ML_FLUX_0_over_W2_FLUX | Log of the flux ratio between the 0.2-5 keV flux and the *CatWISE* counterpart's W2-band flux. |
| log_ML_FLUX_1_over_W1_FLUX | Log of the flux ratio between the 0.2-0.6 keV flux and the *CatWISE* counterpart's W1-band flux. |
| log_ML_FLUX_1_over_W2_FLUX | Log of the flux ratio between the 0.2-0.6 keV flux and the *CatWISE* counterpart's W2-band flux. |
| log_ML_FLUX_2_over_W1_FLUX | Log of the flux ratio between the 0.6-2.3 keV flux and the *CatWISE* counterpart's W1-band flux. |
| log_ML_FLUX_2_over_W2_FLUX | Log of the flux ratio between the 0.6-2.3 keV flux and the *CatWISE* counterpart's W2-band flux. |
| log_ML_FLUX_3_over_W1_FLUX | Log of the flux ratio between the 2.3-5 keV flux and the *CatWISE* counterpart's W1-band flux. |
| log_ML_FLUX_3_over_W2_FLUX | Log of the flux ratio between the 2.3-5 keV flux and the *CatWISE* counterpart's W2-band flux. |
| W1mLQ | -log(*CatWISE* catalogue probability that the source is not at a constant flux in the W1 band.) |
| W2mLQ | -log(*CatWISE* catalogue probability that the source is not at a constant flux in the W2 band.) |

Table 5.4: Table 5.3 continued.

| Feature Name | Description |
|---|---|
| phot_G_mean_mag | *Gaia* counterpart's *G* band magnitude. |
| BP-RP | *Gaia* counterpart's BP band magnitude - RP band magnitude. |
| GAIA_ABS_MAG | phot_G_mean_mag - 5log(100/parallax), where parallax is the Gaia counterpart's parallax. |
| log_ML_FLUX_0_over_phot_G_mean_flux | Log of the flux ratio between the 0.2-5 keV flux and the *Gaia* counterpart's *G*-band flux. |
| log_ML_FLUX_0_over_phot_BP_mean_flux | Log of the flux ratio between the 0.2-5 keV flux and the *Gaia* counterpart's BP-band flux. |
| log_ML_FLUX_0_over_phot_RP_mean_flux | Log of the flux ratio between the 0.2-5 keV flux and the *Gaia* counterpart's RP-band flux. |
| log_ML_FLUX_1_over_phot_G_mean_flux | Log of the flux ratio between the 0.2-0.6 keV flux and the *Gaia* counterpart's *G*-band flux. |
| log_ML_FLUX_1_over_phot_BP_mean_flux | Log of the flux ratio between the 0.2-0.6 keV flux and the *Gaia* counterpart's BP-band flux. |
| log_ML_FLUX_1_over_phot_RP_mean_flux | Log of the flux ratio between the 0.2-0.6 keV flux and the *Gaia* counterpart's RP-band flux. |
| log_ML_FLUX_2_over_phot_G_mean_flux | Log of the flux ratio between the 0.6-2.3 keV flux and the *Gaia* counterpart's *G*-band flux. |
| log_ML_FLUX_2_over_phot_BP_mean_flux | Log of the flux ratio between the 0.6-2.3 keV flux and the *Gaia* counterpart's BP-band flux. |
| log_ML_FLUX_2_over_phot_RP_mean_flux | Log of the flux ratio between the 0.6-2.3 keV flux and the *Gaia* counterpart's RP-band flux. |
| log_ML_FLUX_3_over_phot_G_mean_flux | Log of the flux ratio between the 2.3-5 keV flux and the *Gaia* counterpart's *G*-band flux. |
| log_ML_FLUX_3_over_phot_BP_mean_flux | Log of the flux ratio between the 2.3-5 keV flux and the *Gaia* counterpart's BP-band flux. |
| log_ML_FLUX_3_over_phot_RP_mean_flux | Log of the flux ratio between the 2.3-5 keV flux and the *Gaia* counterpart's RP-band flux. |
| PM_SIG | Significance of the *Gaia* counterpart's proper motion, computed as: $\sqrt{(\text{pmra}/\text{pmra\_error})^2 + (\text{pmdec}/\text{pmdec\_error})^2}$ |
| PLX_SIG | Significance of the *Gaia* counterpart's parallax, computed as: parallax/parallax_error. |
| G-W1 | phot_G_mean_mag - W1mag |
| G-W2 | phot_G_mean_mag - W2mag |

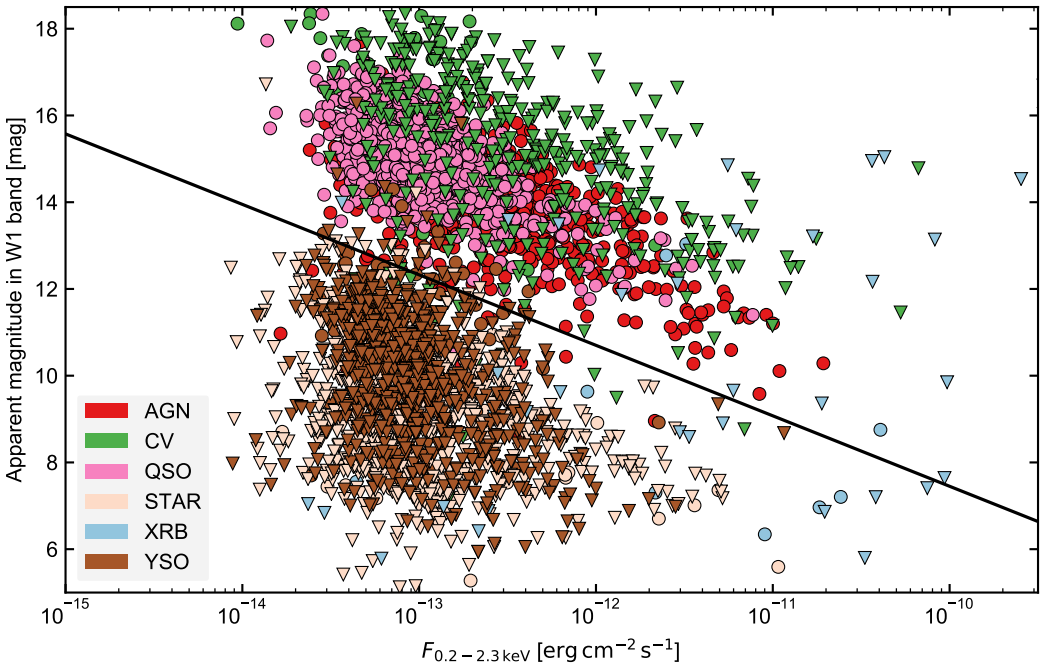


Figure 5.6: The training set properties shown in W1 and $F_{0.2-2.3\,\mathrm{keV}}$ space, with the black solid line equivalent to the relation presented in Salvato et al. (2018) for separating AGN and stars (W1 $= -1.625 \times \log[F_{0.5-2\,\mathrm{keV}}] - 8.8$). The circle and triangle markers denote objects with a non-significant and significant parallax measurement (`PLX_SIG` below and above 3 respectively; `PLX_SIG` defined in Table 5.3), thus typically corresponding to objects of extra-galactic and Galactic origin.

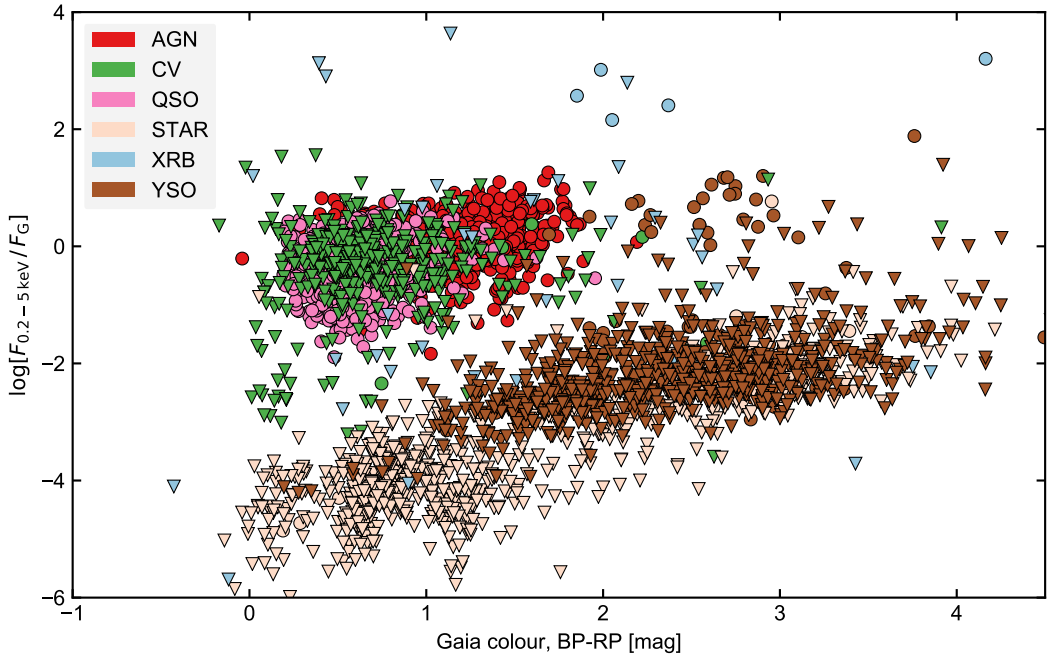


Figure 5.7: The ratio of the source fluxes in the 0.2-5 keV band to the *Gaia* G-band, against the colour of the source in *Gaia* photometry. As in Fig. 5.6, circle and triangle markers denote objects with PLX_SIG below and above 3 respectively.

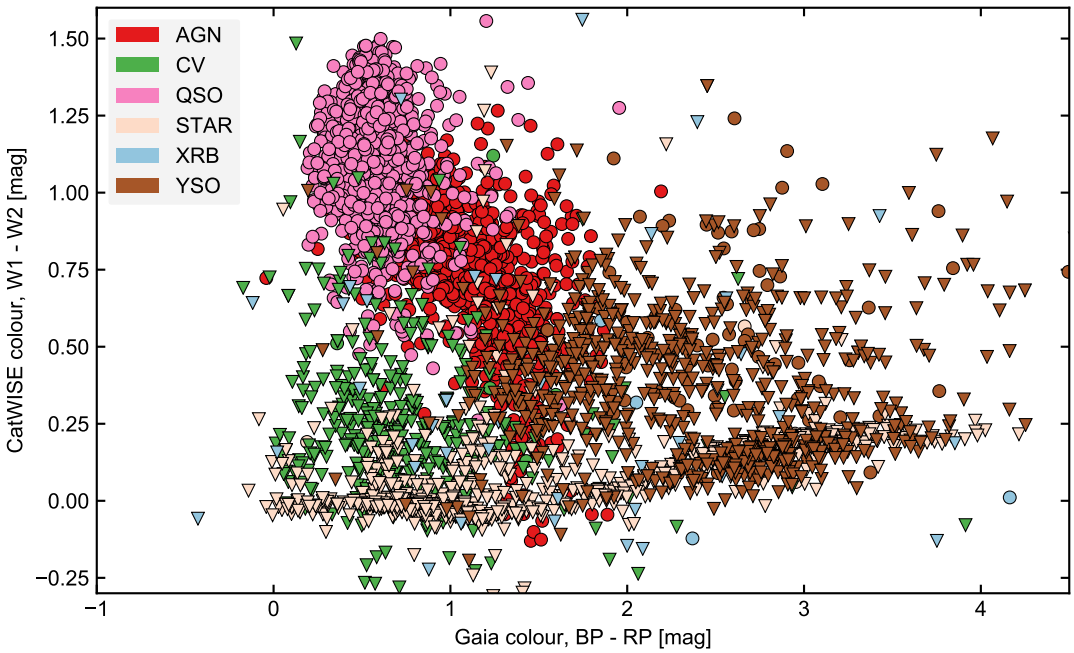


Figure 5.8: As in Fig. 5.6, circle and triangle markers denote objects with PLX_SIG below and above 3 respectively.

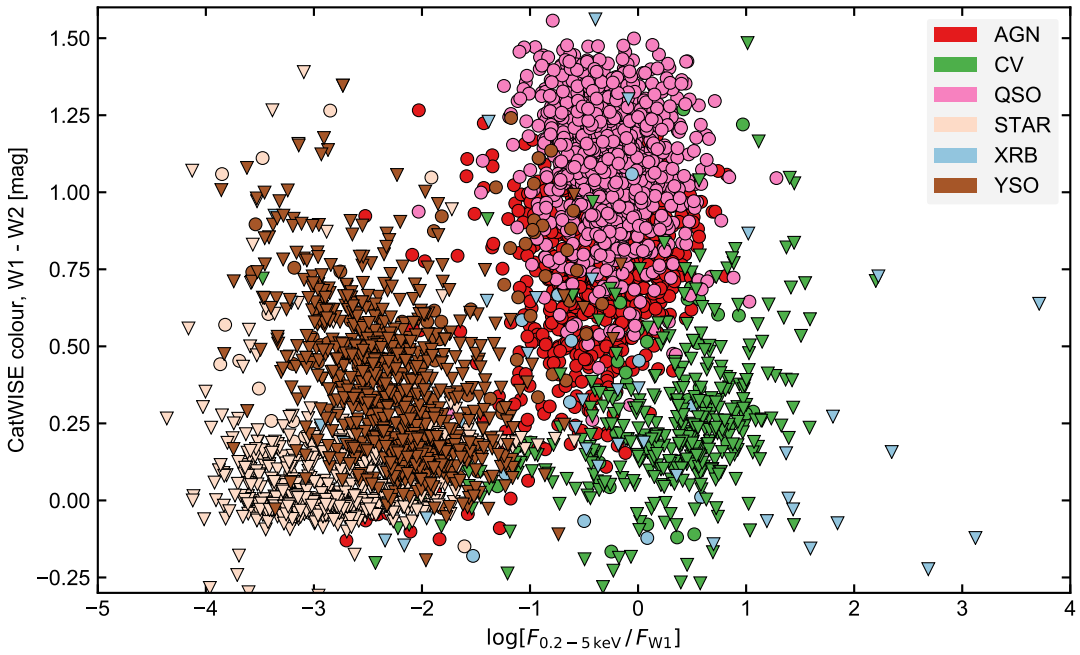


Figure 5.9: The mid-infrared *CatWISE* colour of the source, `W1-W2`, against the ratio of the source fluxes in the 0.2-5 keV band to the *CatWISE* W1-band. As in Fig. 5.6, circle and triangle markers denote objects with `PLX_SIG` below and above 3 respectively.

on the training dataset was run; this involves splitting the training dataset into 5 different random segments[10], training the RF using only objects from four folds, and applying this trained RF to predict the unseen data in the last fold. This was then repeated 5 times, using a different unseen fold each time.

During the above process, each random forest was initialised with 200 decision trees, the minimum samples for a node to be considered as a leaf node was set to 1, and the splitting criterion at each tree node based on the *Gini impurity*[11]. The number of features that could be searched through in each splitting step of the RF construction was set to the square root of the total number of features. To mitigate against the imbalanced training dataset, the 'balanced' mode was used to re-weight each class' contribution to the Gini impurity, with weights inversely proportional to each class' frequency in the training set. For each feature set, the same random number seed for splitting the training dataset during cross-validation was used, and also for initialising the random forest, such that any variations in the classifier performance are only due to the different feature sets used.

To assess the performance of the RF during cross-validation, I consider the classifier's *preci-*

[10] A stratified cross-validation is performed to ensure that the class balance in each fold is approximately constant.

[11] The Gini impurity, $G$, of a set of samples with $J$ different possible classes is given by: $G = \sum_{i=1}^{J} p_i(1 - p_i)$, where $p_i$ is the fraction of total samples in the set belonging to class $i$. If all samples in the set are of the same class, then the set is pure and $G = 0$.

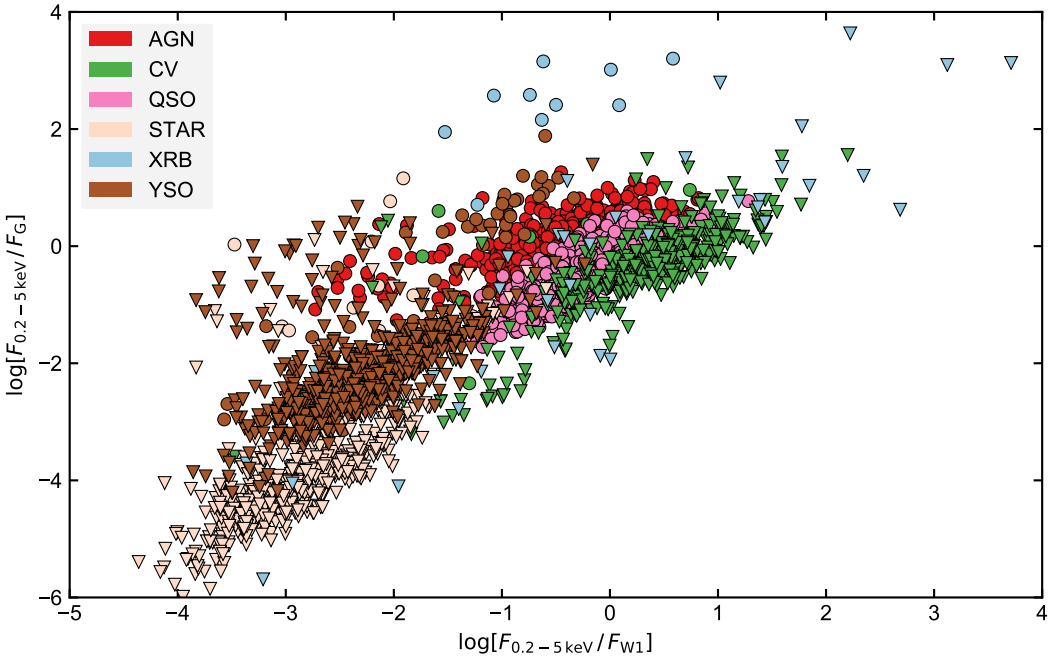


Figure 5.10: The ratio of the source fluxes in the 0.2-5 keV band to the *Gaia* G-band, against the ratio of the source fluxes in the 0.2-5 keV band to the *CatWISE* W1-band. As in Fig. 5.6, circle and triangle markers denote objects with `PLX_SIG` below and above 3 respectively.

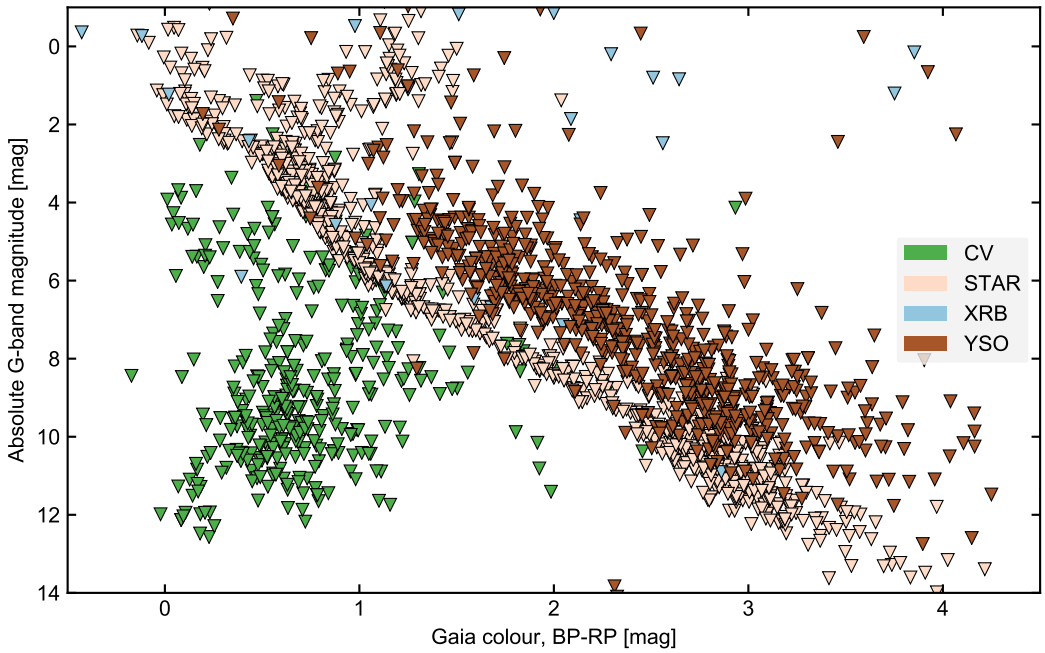


Figure 5.11: A *Gaia* Hertzsprung-Russell diagram for objects of Galactic origin in the training dataset. As in Fig. 5.6, circle and triangle markers denote objects with `PLX_SIG` below and above 3 respectively.

*sion*:

$$P = \frac{N_{\mathrm{TP}}}{N_{\mathrm{TP}} + N_{\mathrm{FP}}} \tag{5.1}$$

where $N_{\mathrm{TP}}$ and $N_{\mathrm{FP}}$ are the number of true and false positive classifications made on the unseen objects; and the classifier's *recall*:

$$R = \frac{N_{\mathrm{TP}}}{N_{\mathrm{TP}} + N_{\mathrm{FN}}} \tag{5.2}$$

where $N_{\mathrm{FN}}$ is the number of false negative classifications made on the unseen objects. In addition, the *average precision score*[12]:

$$AP = \sum_n (R_n - R_{n-1}) P_n, \tag{5.3}$$

is considered, where to compute $AP$, a set of probability thresholds are first generated between 0 and 1. Then, $R_n$ and $P_n$ are evaluated, which denote the recall and precision evaluated at the $n$-th probability threshold.

$P$, $N$ and $AP$ can be computed on a per-source class basis (for example, with what precision and recall does the RF classify CVs?), and/ or across all sources classes (a *micro* average). In this work, the mean of the micro-averaged precision score across all folds, $\overline{AP}_{\mathrm{micro}}$, is adopted as the main quantification of a trained classifier's performance.

[12] `https://scikit-learn.org/stable/modules/generated/sklearn.metrics.average_precision_score.html`

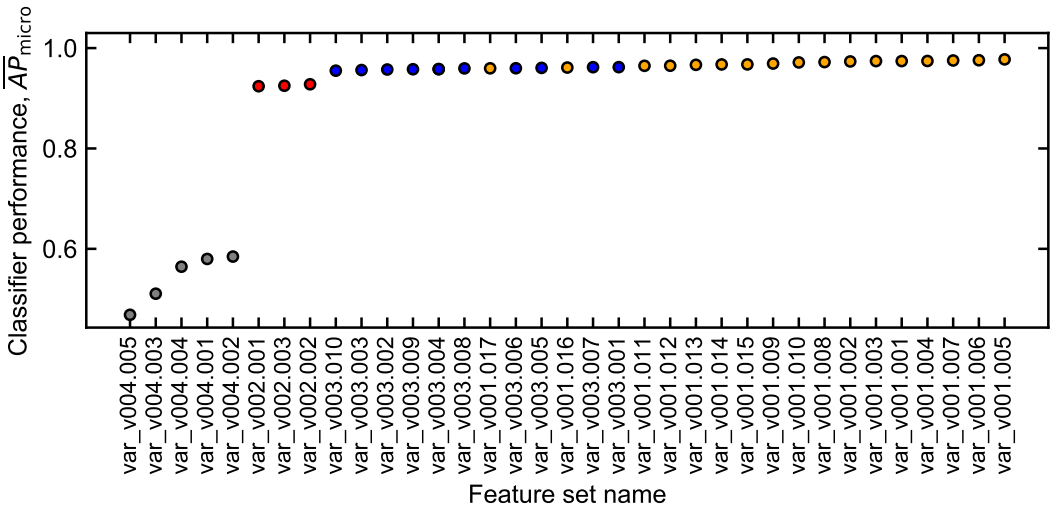


Figure 5.12: The mean of the micro-averaged precision score across all folds, $\overline{AP}_{\mathrm{micro}}$, obtained for each considered feature set (the x-axis tick labels refer to the names of each possible considered feature set; a full list of the features within each set are listed in Appendix B.1). Grey, red, blue and yellow markers denote feature sets derived from eROSITA only, eROSITA and *CatWISE* only, eROSITA and *Gaia* only, eROSITA, *Gaia* and *CatWISE* only features respectively. The inclusion of the multi-wavelength features produces a significant improvement of the classifier's performance.

In Fig. 5.12, the average precision score, $\overline{AP}_{\mathrm{micro}}$, is plotted for each of the 36 feature sets. When the classifier is trained only on features derived from the merged eRASS1xeRASS2 source cataloge, classification performance is extremely poor, with a maximum $\overline{AP}_{\mathrm{micro}} \sim 0.6$. This confirms that the classification of objects based purely on X-ray flux and X-ray colors, i.e. without a robust *Gaia*/ *CatWISE* counterpart association, should be treated with caution and be flagged as potentially spurious. Inclusion of the *CatWISE*, or the *Gaia* features, produces a large (∼0.3 and ∼0.35) increase in $\overline{AP}_{\mathrm{micro}}$, with the strongest classifier performances being produced for feature sets with a mix of eROSITA, *Gaia* and *CatWISE* features. For feature sets with $\overline{AP}_{\mathrm{micro}}$ above that of `var_v001.011` (i.e. those plotted to the right of `var_v001.011` in Fig. 5.12), the $\overline{AP}_{\mathrm{micro}}$ scores are consistent with each other when considering the range of $AP_{\mathrm{micro}}$ values across the 5 folds. To reduce the risk of the classifier being overfitted on the training set, the feature set with the fewest number of features within this group is chosen, `var_v001.014` ($\overline{AP}_{\mathrm{micro}} = 0.97$, with features `ABS_BII`, `W1_W2`, `log_ML_FLUX_0_over_W1_FLUX`, `BP-RP`, `PM_SIG`, `log_ML_FLUX_0_over_phot_G_mean_flux`, `G_W1`; definitions in Table 5.3), as the 'best' feature set, and used for training the classifier.

After having identified a final set of features to be used for training the classifier, further experiments with tweaking the hyper-parameters of the RF (which had previously been fixed during the search for the optimal feature set) were performed. Here, the number of trees used in the RF was varied between 50, 100, 200 and 300, and computed $\overline{AP}_{\mathrm{micro}}$ for each trained model

as before, but it was found that the original choice of 200 trees led to the highest $\overline{AP}_{\text{micro}}$ score; the 200 tree RF was therefore used.

### 5.3.4 Classification metrics

Fig. 5.13 shows a confusion matrix which provides a visual assessment of the classifier's performance. To construct this, the trained classifier produces probabilistic classifications of the previously unseen data, and the class with the highest prediction probability is selected as the classified label for a given object. Each row in the confusion matrix corresponds to objects of a given source class, and the value in a given cell is the number of objects that were classified with predicted label of that cell's column, normalised by the sum of all objects in the row. For example, in the first row of Fig. 5.13, 89% of objects with true label AGN were correctly classified as AGN, whilst 11% get misclassified as QSOs.

In Fig. 5.13, one sees that objects of extragalactic origin (AGN and QSOs) are typically correctly classified as such, with 10% of AGN being misclassified as QSOs (and vice-versa). The RF also correctly classifies 94% of CVs, stars and YSOs, with CV misclassifications being evenly spread over the other galactic source classes, whilst stars and YSOs are most often confused with each other. The weakest performance is for the XRB class, with only $\sim$ 45% of these systems being correctly classified, and generally being confused with CVs and YSOs.

Fig. 5.14 shows the precision-recall curves for the classifier's performance. For a set of probability thresholds between 0 and 1, the precision and recall are plotted (computed using equations 5.1 and 5.2), where a class label for an object is assigned if the probability of the source belonging to that source class is equal to or above the probability threshold. A perfectly performing classifier (all classifications are correct and all objects are recovered) would be represented by a point at (1, 1), whilst weaker classifiers would be characterised by curves that trace out low recall and precision regions.

The classifier is clearly significantly weaker for XRBs relative to the other source classes (Fig. 5.13 and 5.14). This likely stems from a combination of there only being $\sim$80 XRBs in the training dataset (Fig. 5.5), and that the XRBs are predominantly found within regions of high source densities (e.g. the Galactic plane), where identifying the counterpart to an eROSITA source is more challenging. This is due to the classifier needing multi-wavelength features for a robust classification (Fig. 5.12), which relies on the correct counterpart to be identified for each source. Future work to improve the classifier performance on XRBs is discussed in section 5.5.2.

The usage of the RF for classification also provides an estimate of the *relative importance* of each feature for distinguishing between the different source classes. For a given decision tree, one can compute the decrease in the Gini impurity associated with each feature, weighted by the number of objects that were split using this feature. The final relative importance for each feature can then be computed from averaging this value across all the decision trees in the random forest, and then normalising by the sum of this quantity over all features considered. We plot the relative importance of the features used in the trained classifier in Fig. 5.15, which suggests that the ratio of the X-ray/W1 band flux, the $G - W1$ colour, the absolute Galactic latitude, the significance of the *Gaia* counterpart's proper motion, and the $W1 - W2$ colour produce the strongest discriminatory power between the different source classes.

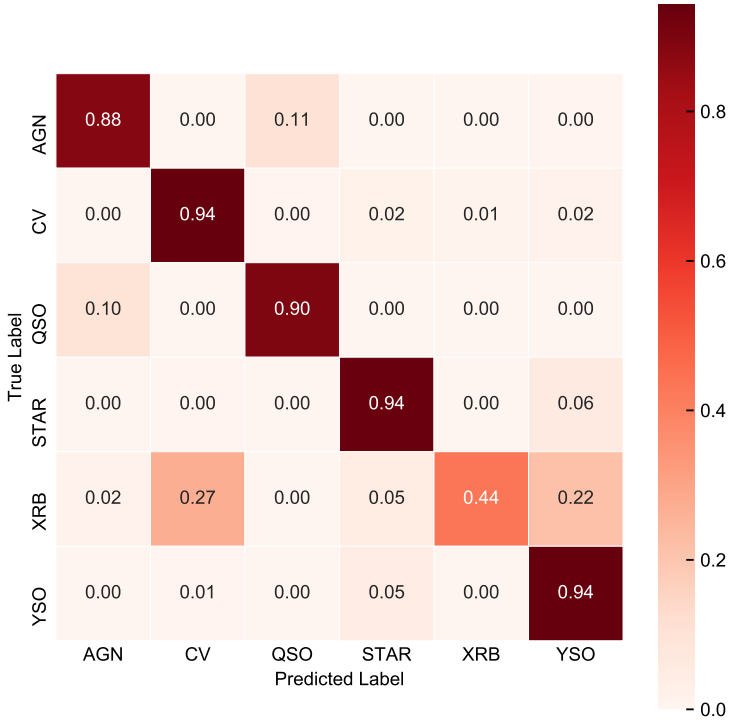


Figure 5.13: Row-normalised confusion matrix for the RF classifier trained using the optimised feature set `var_v001.014`, which uses the features: i) the flux ratio between the 0.2-5 keV and W1 band, ii) the flux ratio between the 0.2-5 keV and G band, iii) the mid-infrared colour, iv) the *Gaia* colour, v) the G-W1 colour, vi) the *Gaia* proper motion significance, and vii) the source's absolute Galactic latitude. The colour bar encodes the fraction of objects that are successfully classified within each row. A strong classification performance is obtained for all classes except for XRBs, which are typically misclassified as CVs or YSOs. The classifier also generally confuses AGN as QSOs and vice-versa, which is partially expected given the relatively arbitrary definition for QSOs in Véron-Cetty & Véron (2010). However, for a classifier use case where one is interested in looking for large amplitude flux changes from supermassive black holes in galactic nuclei (e.g. changing-look AGN), then this QSO/ AGN confusion has minimal impact on the 'usefulness' of the classifier, since one could simply just retain all variables classified as an AGN or QSO.

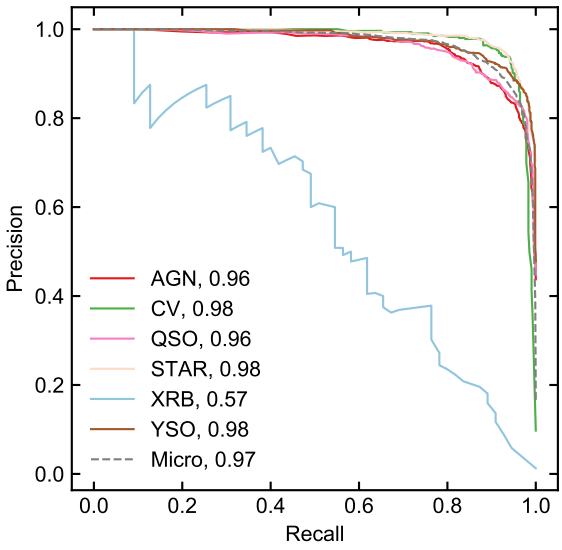


Figure 5.14: Precision-recall curves for the RF classifier trained with the optimised feature set `var_v001.014`, which uses the features: i) the flux ratio between the 0.2-5 keV and W1 band, ii) the flux ratio between the 0.2-5 keV and G band, iii) the mid-infrared colour, iv) the *Gaia* colour, v) the G-W1 colour, vi) the *Gaia* proper motion significance, and vii) the source's absolute Galactic latitude. The average precision score (equation 5.3) is shown next to each class name in the legend.

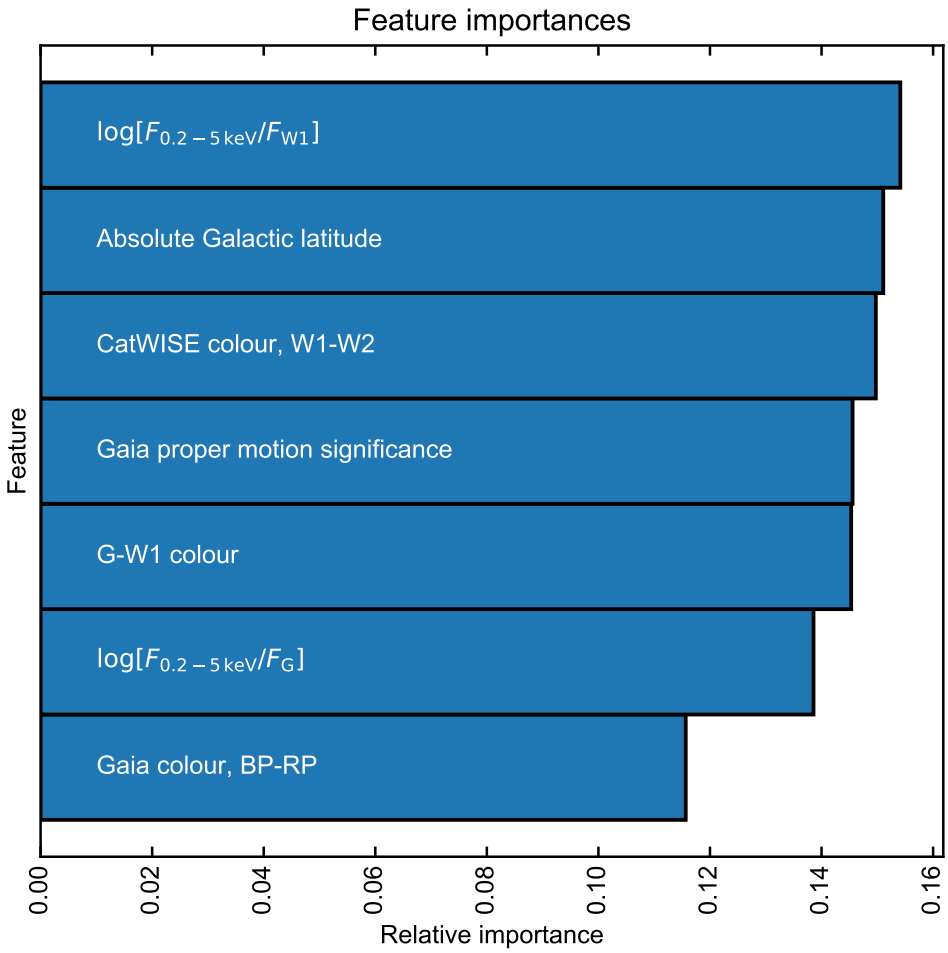


Figure 5.15: Relative importance of features used in a classification model that uses a combination of eROSITA, *Gaia* and *CatWISE* features (the model name is `var_v001.014`, as was also used for Fig. 5.14 and 5.13). Further details on each feature are presented in Table 5.3.

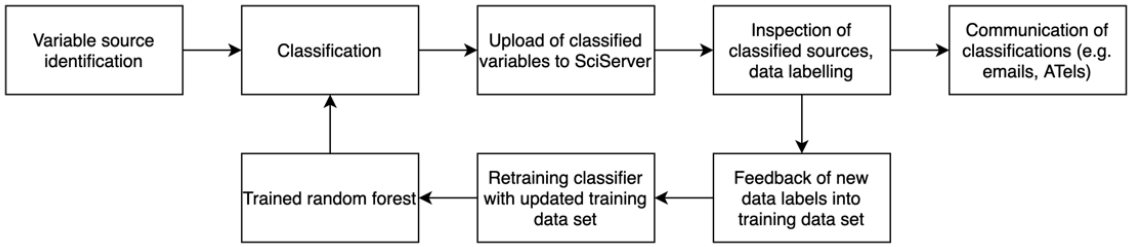


Figure 5.16: Flowchart of operation for variable source selection and classification during the live survey.

## 5.4 Application during eRASS

### 5.4.1 Operations

On a daily basis during weekdays, the searcher pipeline (section 5.2) identifies a set of new transients and variables within the latest eROSITA source catalogues. For each of these variable sources, the multi-wavelength features that the deployed classifier was trained on are computed, and the trained classifier is then used to produce probabilistic classifications of each of these sources. The classified sources are then uploaded onto the MPE SciServer[13] for storage, where members of the German eROSITA consortium are able to access and view these events. A diagram summarising these operations during the live eRASS survey is presented in Fig. 5.16.

To search efficiently through the classifications, a set of Jupyter notebooks[14] were created and made available on the SciServer. These provide an interface for a user to primarily: i) apply various filters on the classified sources and view them (e.g. show details of the 5 largest amplitude flaring sources between eRASS3 and eRASS1), ii) look at a summary set of information for each source, iii) label a source. An example of the summary set of information presented to the user is presented in Fig. 5.17, where this information was chosen so that it would aid assessing and verifying each automated classification. This therefore includes the counterpart assignment probability properties, the multi-wavelength features, and the historic X-ray variability of the source over all eROSITA sky surveys. Inclusion of the NWAY match probabilities also allows the user to assess the strength of the classification.

Through the data labelling functionality provided by the notebook, the user is also able to efficiently label the variable sources, such as AGN/ STAR/. . . , or UNKNOWN, for cases where there is a weak counterpart association, and the source class is not immediately clear. The sources labelled through this process are stored on the SciServer, and these are eventually added to the training dataset every few weeks, when the classifier gets retrained on the expanded training dataset.

[13] `https://www.sciserver.org/`
[14] `https://jupyter.org/`

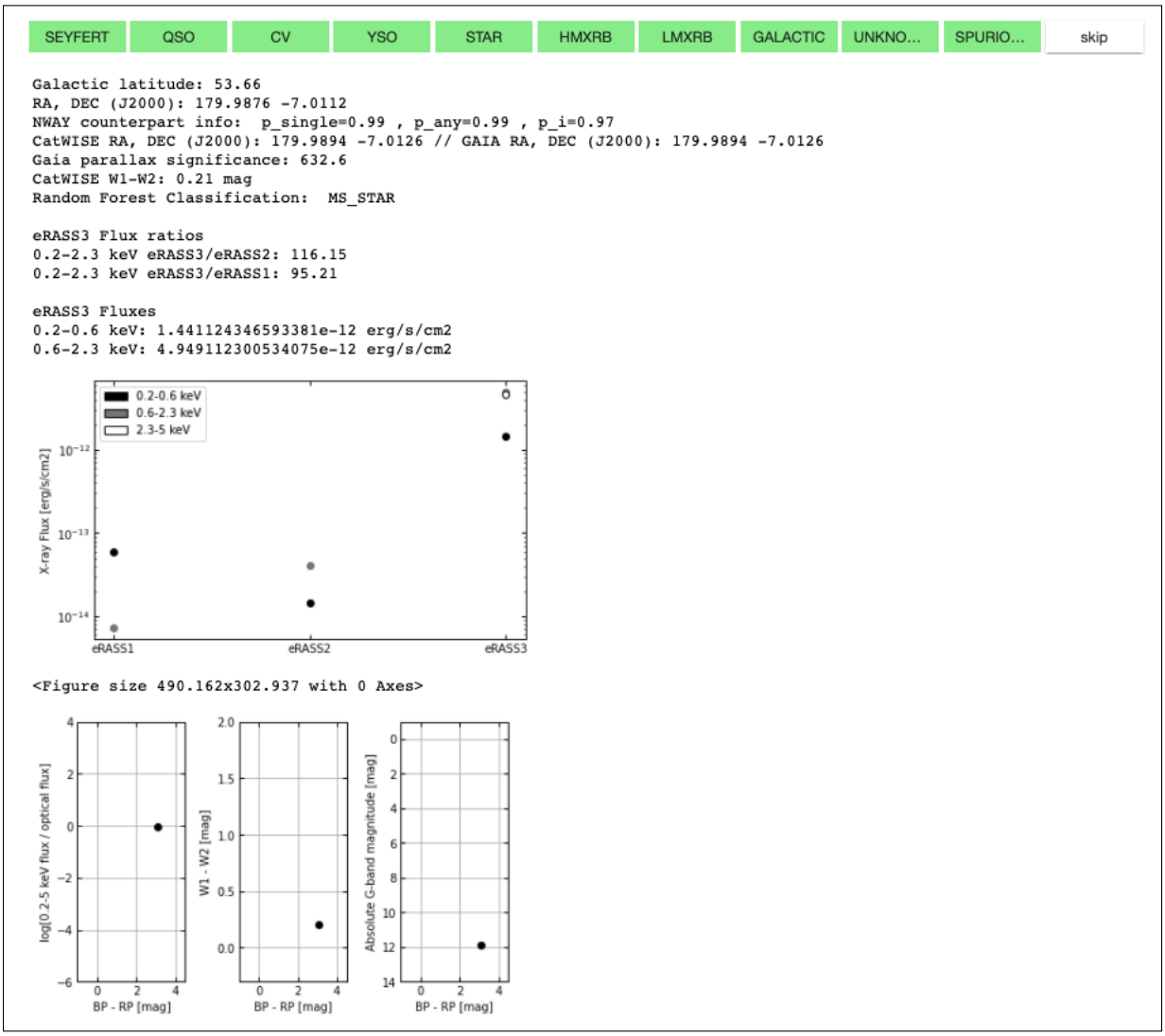


Figure 5.17: Screenshot of the interface presented to a user for searching through and labelling the most recently identified variable sources. The green buttons at the top are used for data labelling, where the user can click on a given button to label an object, with the label being stored on the MPE SciServer. Below this, a brief summary of each source is presented, followed by the eRASS light curve history, and the source's multiwavelength properties.

Table 5.5: Summary of the variability flags of 9060 sources identified by the pipeline between 2021-04-05 and 2021-04-29. The definitions of the variability flags are presented in Table 5.1.

| Variability flags | $N_{\rm src}$ |
|---|---|
| FLUX_3_over_2_BAND_SOFT | 197 |
| FLUX_3_over_1_BAND_SOFT | 113 |
| BRIGHTEN_LAST_3_ERASS | 6942 |
| VAR_ONOFFON | 1586 |
| VAR_ONOFFON and FLUX_3_over_1_BAND_SOFT | 4 |
| BRIGHTEN_LAST_3_ERASS and FLUX_3_over_1_BAND_SOFT | 158 |
| FLUX_3_over_1_BAND_SOFT and FLUX_3_over_2_BAND_SOFT | 34 |
| BRIGHTEN_LAST_3_ERASS and FLUX_3_over_1_BAND_SOFT and FLUX_3_over_2_BAND_SOFT | 26 |

### 5.4.2 Classifier performance during eRASS3

A full assessment of the classifier's performance will only be possible with extensive multi-wavelength spectroscopic follow-up resources (e.g. 4MOST, de Jong et al. 2019; SDSS-V, Kollmeier et al. 2017), since these would enable the true class of each source to be identified for the vast majority of objects. In the following, we instead consider the classifier's performance on variables identified between 2021-04-05 and 2021-04-29 by the pipeline based on the procedure for variable source selection described in section 5.2. This yields a total of 9060 sources, and a breakdown of their variability properties is presented in Table 5.5. Of these 9060 sources, ~91% were assigned a *CatWISE* counterpart using NWAY with $p_i > 0.5$ (i.e. the counterpart chosen is more likely to be the correct counterpart of all possible *CatWISE* counterparts)[15]. Since it was shown in section 5.3.3 that the classifier performs poorly in the absence of multi-wavelength features, automated classifications are thus typically less reliable for the remaining ~9%.

### 5.4.3 Example sources promptly picked out using the pipeline

Since being deployed during eRASS3, the variable source identification and classification pipeline has been highly effective at picking out interesting systems; the following section briefly presents some highlights of sources which have been selected using this approach.

#### Large amplitude flaring between successive eRASS

Through searching for large amplitude flux increases between eRASS3 and 2 using the FLUX_3_over_2_BAND_SOFT flag, the pipeline rediscovered CAL 83, a prototype of the super-soft X-ray source class initially discovered with the *Einstein Observatory* (Long et al., 1981). CAL 83 is a close binary system, consisting of a white dwarf accreting from a more massive

[15]These properties are based on preliminary versions of the counterpart catalogues available at the time of writing, and computed based on the counterparts inferred for the variable sources, and not the entire eROSITA source catalogues.

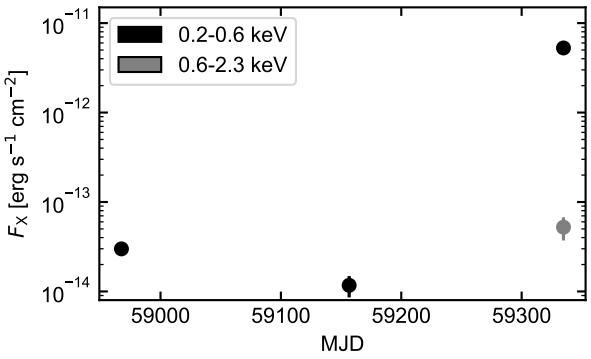


Figure 5.18: eRASS1 to eRASS3 light curve of CAL 83, a prototype of the super-soft X-ray source class. Between eRASS2 and eRASS3, the source brightened by a factor of ∼400 in the 0.2-0.6 keV band, and was non-detected in each eRASS in the 2.3-5 keV band. Error bars on the flux measurements are smaller than markers.

companion star, with the soft X-rays being produced from thermonuclear burning on the white dwarf's surface (van den Heuvel et al., 1992). CAL 83 has been previously observed to cycle through X-ray on and off states (Alcock et al., 1997; Greiner & Di Stefano, 2002), with the off states potentially occurring after large decreases in the accretion rate onto the white dwarf (e.g. Alcock et al. 1997). During eRASS1 and eRASS2, the source was detected in an off state, before flaring by a factor ∼400 in the 0.2-0.6 keV band in eRASS3 (Fig. 5.18). This brightening was promptly picked up by the `searcher` pipeline and classified as a CV (with classification probabilities of 0.885 and 0.115 for CV and XRB respectively, and all other classes 0 probability). The finding was communicated to the eROSITA Compact Object Working Group.

### Brightening over each successive eRASS

The `BRIGHTEN_LAST_3_ERASS` variability flag is generally useful for selecting slower evolving flares, such as AGN ignition events. During operations, the most interesting net successive brighteners are selected through sorting all sources with `BRIGHTEN_LAST_3_ERASS`=1 in order of descending `FLUX_3_over_1_BAND_SOFT`, and predominantly focusing on the largest flux increase events due to time limitations. One such example discovered using the pipeline presented in this work, is eRASSt J092719+230113, which brightened by a factor of ∼27 in the 0.2-2.3 keV band between eRASS1 and eRASS3 (although only a factor of ∼5 between eRASS2 and eRASS3 in the same band). eRASSt J092719+230113 is associated with a Seyfert I at $z = 0.026$, and was initially correctly classified as an AGN by the classifier. The Seyfert I is also extremely variable in the optical bands, with the joint eROSITA and ATLAS forced photometry light curve presented in Fig. 5.19.

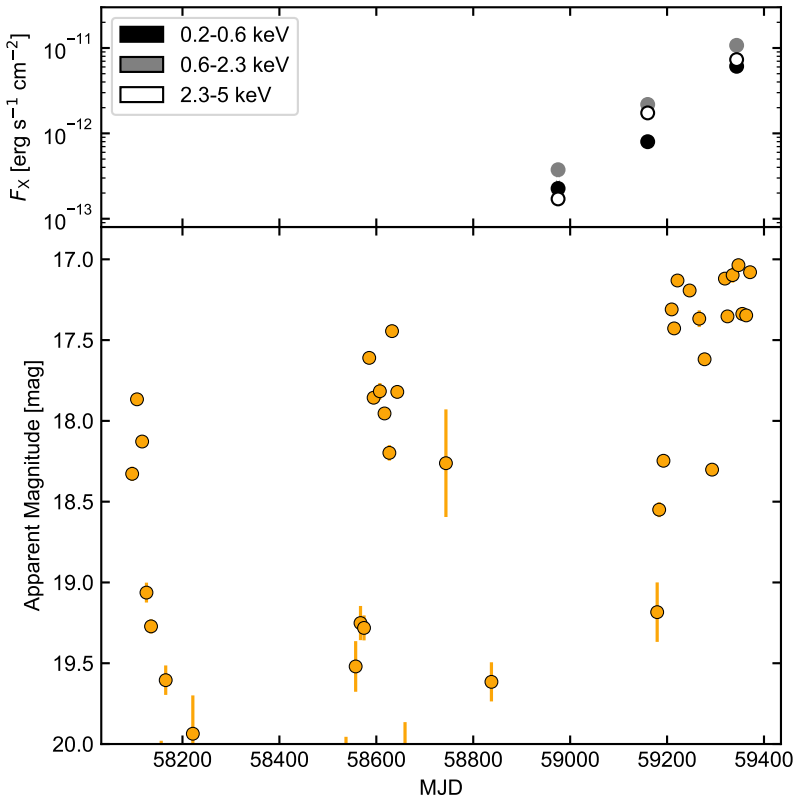


Figure 5.19: Joint eROSITA and ATLAS *o*-band (5600-8200 Å) light curve of a highly variable AGN, eRASSt J092719+230113, which was initially identified and classified as an AGN through the pipeline described in this work. The source shows a net brightening in the 0.2-2.3 keV band from eRASS1 to eRASS3 by a factor of ~27. Error bars on the eROSITA flux measurements are smaller than the markers.

**On-off-on events**

To swiftly identify the most extreme on-off-events (sources which were bright in eRASS1 and eRASS3, but not detected in eRASS2) during daily operations, we sort sources with `VAR_ON_OFF_ON` flag equal to 1 in descending 0.2-2.3 keV flux, based on the inferred fluxes in the eRASS1 and eRASS3 source catalogues. This sorting is performed in order to produce a cleaner selection of on-off-on sources, because the brightest sources with `VAR_ON_OFF_ON` ==1 are much less likely to be non-detected in eRASS2 due to a lower exposure relative to the eRASS1 or eRASS3 scans (for example, a source with eRASS1 and eRASS3 fluxes of $10^{-12}$ erg s$^{-1}$ cm$^{-2}$, but not detected in eRASS2, is more likely to be truly on-off-on compared with a source with eRASS1 and eRASS3 fluxes of $10^{-13}$ erg s$^{-1}$ cm$^2$ and not detected in eRASS2, where the eRASS2 non-detection could easily arise if the exposure in a given sky region is lower during eRASS2). This approach enabled the prompt identification of eRASSt J192932.9-560346, a novel eclipsing magnetic CV (the CV classification probability was 0.98 for this source), which was detected in eRASS1 and eRASS3 with 0.2-2.3 keV flux $\sim 1.2 \times 10^{-12}$ erg s$^{-1}$ cm$^{-2}$ and $\sim 1.5 \times 10^{-12}$ erg s$^{-1}$ cm$^{-2}$ respectively, but non-detected in eRASS2 (Fig. 5.20). The *Gaia* optical counterpart of eRASSt J192932.9-560346 was also separately reported as a potential CV candidate in the *Gaia* alerts system[16] based on its large amplitude optical variability (a joint eROSITA-Gaia light curve is presented in Fig. 5.20). Follow-up spectroscopy and photopolarimetry revealed the white dwarf to be accreting matter from its companion star via accretion columns on to both of its poles; further details and analysis of the system are presented in Schwope et al. (2021).

## 5.5 Discussion

### 5.5.1 Summary

In this chapter, I presented a detailed report on a pipeline developed, and currently deployed, for the automated identification and classification of variables discovered during the eROSITA All-Sky Survey. A supervised approach to source classification was adopted, where a random forest classifier was trained on the features of a set of labelled sources, and then used this trained model to produce probabilistic classifications of new variable sources discovered during eRASS3. Assessing the classifier's performance using the micro average precision, I systematically explored the classifier's performance on various different feature sets, and found that the inclusion of multi-wavelength features, derived from the properties of the source counterparts identified in the *CatWISE* and *Gaia* source catalogues, yielded significant improvements in the classifier performance, relative to only the eROSITA-based features considered in this work. As a result of this, for sources without *CatWISE* and/ or *Gaia* counterparts, the classifications are less reliable, and such variable sources require further manual inspection. As the survey progresses and the number of labelled sources grows, the classifier can straightforwardly be retrained based on the updated training data set (i.e. the results described in this work reflect the trained classifier's performance at the time of writing), and it will also be possible to add in additional variable source

[16] `http://gsaweb.ast.cam.ac.uk/alerts/alert/Gaia21bxo/`

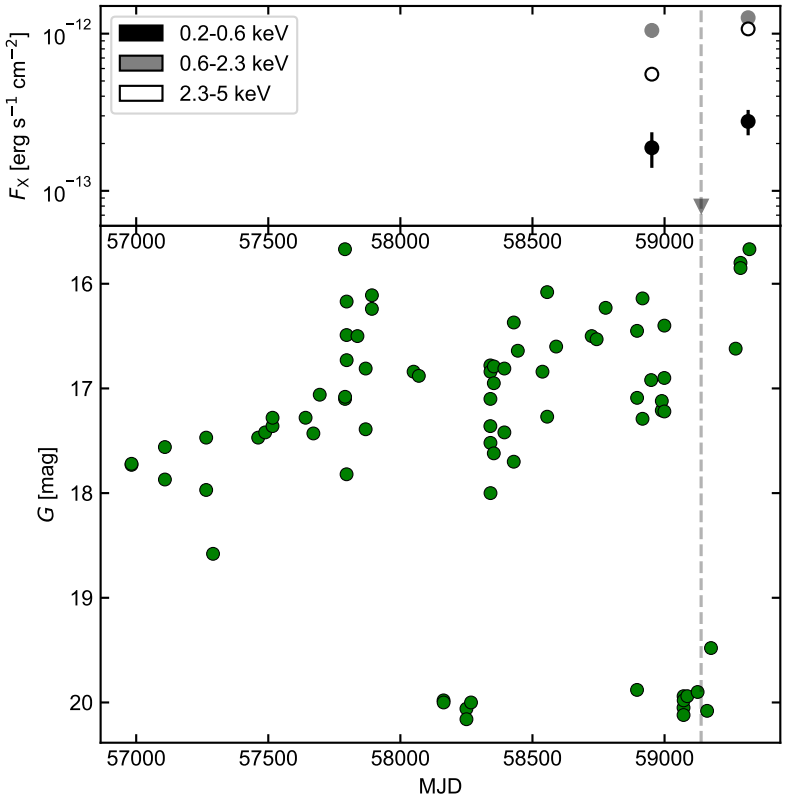


Figure 5.20: Joint eROSITA and *Gaia G*-band (apparent magnitude) light curve of the eclipsing two pole accretor eRASSt J192932.9-560346 (Schwope et al., 2021), which was initially identified and classified as a CV through the pipeline described in this work. The source was not detected in eRASS2, with the grey triangle marking the 0.6-2.3 keV band upper flux detection limit reported in Schwope et al. (2021).

classes (e.g. ultra-luminous X-ray sources, Earnshaw et al. 2019; or breaking up the XRBs into low and high mass XRBs). Finally, I presented example applications of the classifier during a segment of the third eRASS, highlighting how the pipeline has resulted in the detection of a number of new, interesting variable systems.

### 5.5.2 Future work

Taking place over a four year period, the eROSITA All-Sky Survey will offer an unprecedented insight into the dynamic X-ray sky. Whilst the pipeline presented in this work is exceptionally useful for quickly identifying large amplitude flux changes between the different eRASS (especially when deployed within a consortium with relatively limited man-power), there are a number of additional improvements to the pipeline that could be made in further work.

#### Variability identification

As discussed in section 5.4.3, the existing criteria for variability identification picks up sources which have changed fluxes between the different eRASS, and those which show variability within the light curve of a given eRASS may be missed. The design of the current criteria was guided on the basis of: i) the authors initially being most interested in selecting large amplitude flux changes within galactic nuclei over month-to-year-long timescales (e.g. due to changing-look AGN or TDEs), and ii) the light curves for each source within an eRASS were typically only available several weeks after their source catalogues were generated, and thus the pipeline was tailored towards promptly picking up flux changes based on the source catalogues. However, a relatively straightforward modification would be to expand the variability criteria used for variable source selection. For example, one may compute various statistics on the source light curves (e.g. the maximum amplitude deviation; the normalised excess variance, Nandra et al. 1997; the Bayesian excess variance, Buchner et al. 2021), and for sources which show significant variability, produce probabilistic classifications of these using the trained classifier.

#### Source classification

The current main bottleneck to the classifier's performance stems from assigning a *Gaia* and *CatWISE* counterpart for each source detected by eROSITA, which impacts the classifier's effectiveness at two key points. The first is during the training process, where the classifier will be better able to learn the typical characteristics of a source class, if it is shown features that are drawn from the underlying distribution for that class, instead of 'missing' features if no counterpart is identified, or 'wrong' features if the wrong counterpart is selected. Second, in applying the classifier to the unseen test dataset, the classifier will be less likely to misclassify an eROSITA source if it is given features from the correct counterpart. Whilst NWAY provides a much higher completeness and accuracy in counterpart identification relative to simple positional based cross-matching, counterpart identification in regions of higher source density, such as the Galactic plane, is still an ongoing challenge for eROSITA sources.

Given that a strong classification performance is obtained without using variability information, an additional future application of this classifier would be to use it for classifying the entire eROSITA point source sky, instead of just a subset of highly variable sources. For such a use case, it may be desirable to further split up the different source classes in the training dataset. For example, while currently actively accreting SMBHs are split into AGN versus QSO[17], an alternative could be to split these into jetted versus non-jetted AGN. A further division of the latter could then be made into obscured (type 1) and unobscured (type 2) AGN. However, whilst such systems do have different spectral energy distributions, it is not immediately clear how effective a classifier would be at distinguishing between these classes when only using the current eROSITA, *CatWISE* and *Gaia* source catalogues. Such an adapted classifier might therefore benefit from being trained on i) features derived from additional multi-band optical photometry obtained from the DESI Legacy Imaging Survey[18] (Dey et al., 2019), ii) features derived from radio source catalogues (although these would be very useful for distinguishing between jetted and non-jetted AGN, such catalogues typically have patchy sky coverage/ insufficient depth over the German eROSITA sky). An additional division of the current source classes could be performed for the young stellar objects, where one could split these off into class I, II or III YSO candidates (Lada, 1987), with class number increasing with stellar age (class I objects represent the youngest YSOs, which are still embedded in dense clouds of circum-stellar material).

Whilst the undersampling of the AGNs, QSOs and stars in this work did produce an improvement of the classifier's performance (on the minority classes), the final class balance after undersampling was chosen to very roughly reflect the expected distribution of eROSITA variables identified by the pipeline. As the classifier's performance worked sufficiently well for the purposes of automatically classifying interesting transients during the eRASS, there was no strong need to further fine tune the class balance in the training set in this work. However, if the classifier were to be adapted into a more general eROSITA source classifier, then further attention will probably need to be paid to the class balance in the training dataset, and how this compares to the 'true' class balance in the eROSITA sky. If not accounted for, then biases may creep into source classifications, which, for example, would complicate the analysis of source class populations identified through the adapted classifier.

### Transient identification

This work has predominantly focused on an automated pipeline for the identification and classification of variables detected during the eROSITA All-Sky Survey, and has largely avoided dealing with the transients picked up by the `searcher` pipeline. In this context, I refer to transients are as sources detected in the latest eRASS with no prior detection in any of the preceding

[17]This class splitting is sufficient for our current use case, as the AGN in the current training set are generally optically redder, and have a larger fraction of objects at lower `W1-W2` values, than the QSO class (Fig. 5.8). This leads to interesting variables from galaxies with `W1-W2`$\lesssim 0.5$ typically being classified as AGN instead of QSOs, which is a useful aspect of the classifier for TDA science (if these systems do host an AGN, then there is weak emission from the obscuring dusty torus relative to the host galaxy in the mid-infrared, and are potentially low-luminosity AGN which have recently brightened).

[18]This will use the filter set *g*, *r*, *i*, and *z*, cover the entire western Galactic hemisphere, and should become available within the next year.

scans). Transients would likely originate from a diverse range of astrophysical populations, and encompass more exotic phenomena such as X-ray bright tidal disruption events (TDEs; e.g. Komossa & Bade 1999; Esquej et al. 2008b; Burrows et al. 2011; Saxton et al. 2012; Holoien et al. 2016), fast X-ray transients (e.g. Jonker et al. 2013; Glennie et al. 2015; Bauer et al. 2017; Xue et al. 2019), shock breakout in supernovae (e.g. Soderberg et al. 2008) and gamma-ray burst afterglows (e.g. D'Avanzo et al. 2012). Developing an automatic pipeline for identifying and classifying such systems faces several additional challenges that would need to be overcome in future work.

Not every new eSASS-detected point source detected in the latest eRASS, which was not detected in previous scans, is associated with an astrophysical source. Spurious novel point source detections can arise from multiple origins: imperfect source detection within the vicinity of extended X-ray objects (such as a galaxy cluster being detected as a single extended object in eRASS1 and eRASS2, but as an extended object and point source in eRASS3[19]), (not yet corrected) hot pixel streaks triggering false source detections, or detection artifacts by extremely bright X-ray sources. This means that the candidate transient population uncovered by the `searcher` pipeline can be contaminated by false-positive events, and any classifier that were to run on such a population would produce a large number of false alerts and be ineffective. A possible solution would be to develop a new classification algorithm for distinguishing whether there is a new point source or not in a pair of X-ray image cutouts of the transient's position (one cutout from the current eRASS, and the other from the preceding eRASS). Such a tool would enable one to drop out all transients which have a high probability of being spurious (based on their image cutout), and therefore produce a stream of transients with a lower false-positive contamination. In addition, the above approach should be more sensitive to detecting transients within galaxy clusters relative to alternative methods for handling such events (e.g. removing all transients that are found within the extent of a galaxy cluster).

If a purer stream of transients can be obtained through the above modifications, then the next challenge is to produce a characterisation of the variability, which will be necessary for identifying the most interesting transients for follow-up. Since by definition it will not be possible to compute the variability flags described in Table 5.1 for these sources, one avenue would be to simply sort by the flare amplitude (relative to the flux upper detection limit at the transient's position derived from the previous eRASS coverage), or sort by the latest source flux (i.e. place higher interest in the events with the largest brightening). Alternatively, if the eSASS light curves become promptly available for these transients, then a sub-selection based on various variability criteria shown within the eRASS light curve could be performed (e.g. does the source brighten and fade to flux levels below the detection threshold, is it persistently bright, or is it detected only in a single 40s visit?).

Finally, one would want the pipeline to produce an approximate classification of the system, taking into account both the possible multi-wavelength counterparts, which will likely again be essential for the classification, and the variability properties. Whilst it might be possible to use a classifier (similar to what was described in this work) to filter out transients that have a high

[19]There is also the caveat that X-ray transients are also discovered in galaxy clusters (e.g. Maksym et al. 2013), so the detection of a point source within an extended galaxy cluster can also be due to a real X-ray brightening.

probability of being coronal flares, extra caution would need to be applied when using the multi-wavelength information for classification of non-coronal flaring transients here. For example, higher luminosity transients (e.g. off-axis gamma ray bursts, Xue et al. 2019; relativistic TDEs, Burrows et al. 2011; Levan et al. 2011) could be associated with high redshift, optically faint galaxies that lie below the limiting magnitude of *Gaia* or *CatWISE*. An interesting source could easily be misclassified and overlooked if a pipeline did not factor in both the variability and possible multi-wavelength properties.

# Chapter 6

# Conclusion

## 6.1 Summary

The work in this thesis commenced in October 2017, nearly two years prior to the launch of eROSITA. At this point in time, Khabibullin et al. (2014) had already predicted that eROSITA could detect 1000s of TDEs per year involving main-sequence (MS) stars and SMBHs during its four year All-Sky Survey. There had also been several WTDE candidates reported in the literature (see references within Chapter 3), with each of these unveiling an IMBH candidate with $M_{\rm bh} \lesssim 10^5 M_{\odot}$, a mass range which has previously been elusive for identifying strong BH candidates. With the advent of the launch of eROSITA, it was initially of great interest to produce a similar estimate for eROSITA's detection rate of WTDEs, with the primary aims of: i) preparing for detecting such a set of WTDE candidates, and ii) understanding whether WTDEs could be used for uncovering a population of IMBHs using eROSITA. In Chapter 3, I ran an extensive set of Monte-Carlo based simulations to explore eROSITA's detection sensitivity to WTDEs as a function of $M_{\rm bh}$, redshift and time offset between the event flaring and it first being observed during the eROSITA All-Sky Survey. I also presented a novel estimate of the rate density of WTDEs in the local Universe, and combined this with detection sensitivities to infer that eROSITA could detect $\sim$3 WTDEs during its four year survey. This work also demonstrated how the SIXTE simulator may be used for forecasting the transient detection abilities of future X-ray surveys.

The pre-launch expectations for the observational signatures of X-ray bright TDE candidates were based around the Rees (1988) model for their evolution (i.e. an ultra-soft, large amplitude, single peaked X-ray flare from a previously quiescent galaxy, that would subsequently undergo a power-law-like decay), since flares with such properties had almost ubiquitously been associated with TDE candidates in the literature. Furthermore, the ultra-soft X-ray spectra of these events had also effectively been used as a 'smoking-gun' for the flare being TDE-induced. In Chapter 4, I reported on a set of multi-wavelength observations of the unique, ongoing nuclear transient AT 2019avd, located in the nucleus of a previously quiescent galaxy. The key properties of AT 2019avd were: i) a net brightening in the 0.2-2 keV band by a factor of 600 relative to a previous $3\sigma$ flux upper detection limit, ii) an ultra-soft X-ray spectrum ($kT \sim 85$ eV in eRASS1), iii) a double-peaked optical light curve, with the rise and decay timescales of the first and second

peaks being TDE/SNe and AGN-like respectively, iv) transient He II, Bowen fluorescence, and coronal emission lines in the optical spectra. While the X-ray properties alone would have made AT 2019avd a strong TDE candidate, the broader multi-wavelength picture clearly shows a more complex scenario than that which has been seen in all previous TDE candidates. Thus the main outcome of this work is that the X-ray properties of a nuclear transient alone are not always able to distinguish between different physical origins of the large-amplitude variability seen in galactic nuclei, with more complex variability behaviour possibly present in the multi-wavelength datasets.

As discussed in Chapter 1, the dynamic nature of the X-ray sky is produced by a diverse population of different astrophysical systems. In Chapter 5, I presented a detailed report of a transient and variable source identification and classification pipeline, which is currently deployed during the eRASS survey. Adopting a supervised approach to classification, I trained a random forest-based classifier on a set of X-ray and multi-wavelength features (derived from the counterpart properties in the *Gaia* and *CatWISE* source catalogues) of known sources, and subsequently used this trained classifier to make probabilistic classifications of variables discovered during eRASS3. During the training process, I showed that the multi-wavelength features are essential for an accurate and precise source classification. This work has led to several extensive multi-wavelength follow-up campaigns of eROSITA-selected transients having already been undertaken, with further publications on these events anticipated for early 2022.

## 6.2 Future research

This thesis has laid foundations, and traced out initial steps, for systematically exploiting the transient sky of eROSITA. With a transient identification and classification pipeline now also in place for searching for X-ray transients, this work will be further built upon, and enable transient studies to be performed, by the eROSITA consortium in the coming years.

It is briefly worth mentioning that no strong WTDE candidates have yet been identified in the German half of the eROSITA sky during the first three eRASS scans. Whilst this is consistent with the predictions made in Chapter 3 of this thesis, it is important to stress that the search for WTDEs to date would have relied upon the transient being associated with a galaxy in optical imaging[1], so events which originate from isolated dwarf galaxies or globular clusters may have been missed. After working with the real eROSITA datasets over the first three eRASS (and experiencing the challenges of TDE identification raised by AT 2019avd), it is also now clear that for an eROSITA-selected transient to become a strong WTDE candidate[2], one will most likely need the detection of the predicted under-luminous, Ia-like supernova associated with the disruption (MacLeod et al., 2016). This means that there should also be: i) a detection of optical transient emission within the eROSITA localisation uncertainty (which would also offer much improved localisation accuracy over eROSITA and host identification), ii) a relatively well sampled optical light curve available for this transient (useful for constraining the evolution and energetics of the

[1]The eROSITA consortium is man-power limited, and there is unfortunately not enough time for inspecting and following-up all transients; focus and efforts have therefore been dedicated towards nuclear transient candidates.

[2]Assuming its X-ray properties are consistent with the predicted observational signatures discussed in Chapter 3.

event), and iii) an optical spectrum showing the predicted observational signatures for the TDE-induced thermonuclear burning (i.e. Ia-like, MacLeod et al. 2016). Due to these factors, the search for WTDEs with eROSITA may be more fruitful in the Russian half of the eROSITA sky, which has a much larger overlapping footprint with the Zwicky Transient Facility (arguably the 'best' optical transient survey for TDE searches currently operating) compared with the German sky.

The discovery of AT 2019avd (Chapter 4) has opened up a wealth of novel questions relating to nuclear transients. For example, was AT 2019avd induced by a TDE, or was it a more exotic nuclear transient that generated a TDE imposter (at least in the X-rays)? If the latter, then how many additional varieties of TDE imposters exist, and will it ever be possible to obtain a 'smoking gun' signature of a TDE for robust identification? Also, how common are AT 2019avd-like events? A highly promising avenue to begin answering such questions is to systematically follow-up eROSITA's nuclear transients with *Swift* XRT and UVOT, monitoring their multi-wavelength evolution in the months after their initial eROSITA detection. Such an observational campaign would enable the landscape of nuclear transients to be mapped out (at least partially), which will be extremely useful for grouping together the various transients based on their observational characteristics.

Complementary to this approach, the four-year eRASS will also provide an invaluable dataset to start answering such questions. During this period, eROSITA will monitor a sample of $\sim 10^6$ AGN every 6 months, which will allow for a measurement of their X-ray flux variability distributions over a range of timescales. With this, then it may be possible to produce a statistically defined criteria for major X-ray ignition events in galaxy centres (e.g. those systems which show a factor of 15 increase in their 0.2-5 keV flux between two successive eRASS scans). After identifying such a population, then one may also produce a systematic study of its X-ray properties, and produce a comparison of how the traditional expected observational signatures of TDEs compare to this eROSITA-selected population of X-ray ignition events detected during eRASS1-8.

Going beyond eRASS8, then the next generation of time-domain surveys are expected to commence operations, heralding a rich new era of transient discoveries. In the optical domain, the key discovery machine will be the Vera Rubin Observatory Legacy Survey of Space and Time (LSST; Ivezić et al. 2019), which will image the entire visible sky every ~3 nights and is expected to detect 1000s of TDEs per year (Bricman & Gomboc, 2020). The *Ultraviolet Transient Astronomy Satellite* (*ULTRASAT*; Sagiv et al. 2014) will be uniquely carrying out the first dedicated wide field, time-domain survey in the UV[3], and is anticipated to detect ~100s of TDEs per year (Sagiv et al., 2014). The X-ray sky will also be monitored by the *Einstein Probe* (*EP*; Yuan et al. 2018), which will be equipped with a wide-field X-ray telescope (3600 square degree FoV, 0.5-4 keV energy range) that aims to detect new, bright X-ray transients, which may then be automatically followed-up with its secondary X-ray telescope on-board. Whilst potentially detecting fewer TDEs than eROSITA, the *EP* may offer TDE light curves with higher cadence monitoring that will provide tighter constraints on the X-ray light curve evolution relative to eROSITA. Additional eyes on the high energy sky will be provided through the *enhanced X-ray*

[3] 220-280nm.

*Timing and Polarimetry mission* (*eXTP*; Zhang et al. 2016), and the *Space Variable Object Monitor* (*SVOM*; Wei et al. 2016). One of the important legacies of eRASS during this era will be its provision of deep X-ray imaging over the entire sky, which will be useful for constraining the nature and amplitude of any X-ray spectral changes of transients identified in future missions.

Lastly, the forthcoming decades in TDA will also stand out for the growing coordination between electromagnetic and gravitational wave (GW) astronomy. TDEs are predicted to produce bursts of gravitational waves mainly from: i) the changing quadrupole of the star-BH system during the disruption (Kobayashi et al., 2004; Rosswog et al., 2009; Haas et al., 2012; Anninos et al., 2018), and ii) the changing internal quadruple of the star around the moment of its maximal tidal compression at pericentre (Guillochon et al., 2009; Stone et al., 2013). For the former GW emission mechanism, Rossi et al. (2021) suggest that the characteristic strain and frequencies of GWs from MS TDEs would lie below the detection threshold of any GW observatory planned within the next 20 years. However, they may be detectable for WTDEs with future observatories, such as the *DECi-hertz Interferometer Gravitational Wave Observatory* (*DECIGO*; Sato et al. 2017), or the *Advanced Laser Interferometer Antenna* (*ALIA*; Bender et al. 2013), planned for launch in the 2030s. A similar issue arises for detection of GW emission from the changing internal quadrupole of the star during the disruption. Thus for the immediate future, it is likely that major advances in TDE science will be dominated by traditional electromagnetic observations. However, once the state of gravitational wave astronomy has sufficiently advanced to the point where a detection of this GW emission is feasible, then such a detection could eventually be used to i) distinguish between TDE and non-TDE induced origins for nuclear transients, and ii) pinpoint the moment of the star passing pericentre. This would then effectively provide a timestamp for the start of the TDE, which could then be used for mapping out the subsequent electromagnetic evolution of the flare.

# Appendix A

# Additional information for AT 2019avd

## A.1 Optical spectrum and lightcurve fitting

In Table A.1, we list the priors adopted in the fitting of the ZTF/ SEDM lightcurves, whilst in Table A.2, we list the priors used in our fitting of the NOT and WiFeS optical spectra.

## A.2 Long term lightcurve of AT 2019avd

In Fig. A.1, we plot the long term lightcurve of AT 2019avd, including the ASAS-SN data. ASAS-SN (Shappee et al., 2014) observed the location of AT 2019avd in $V$-band from Feb 2012 to Nov 2018 and in $g$-band from Oct 2017 to Sept 2020 (the time of this writing). The $V$- and $g$-band observations were reduced using a fully-automated pipeline detailed in Kochanek et al. (2017) based on the ISIS image subtraction package (Alard & Lupton, 1998; Alard, 2000). During each visit ASAS-SN observed three 90-second dithered images that are then subtracted from a reference image. For the $g$-band we modified the standard pipeline and rebuilt the reference image without any images with JD $\geq$ 2 458 518 to prevent any flux contamination from the outbursts.

All subtractions were inspected manually to remove data with clouds, cirrus, or other issues. We note, however, that the ASAS-SN light curve was negatively affected by two factors. First, there is a bright nearby star that is not resolved from the host galaxy in ASAS-SN data and added noise to the subtractions. Second, the location of AT 2019avd is right on the edge of two ASAS-SN fields. To help alleviate these issues and increase the ASAS-SN limiting magnitude we stacked the subtractions within a maximum of 10 days. We then used the IRAF package `apphot` to perform aperture photometry with a 2-pixel, or approximately $16\farcs0$, radius aperture on each subtracted image, generating a differential light curve. The photometry was calibrated using the AAVSO Photometric All-Sky Survey (Henden et al., 2015).

Table A.1: Priors adopted in the fitting of the ZTF lightcurves. The rise and decay timescales are in units of days, whilst $t_{\rm peak}$ is in MJD. $F_{\rm max}$ refers to the maximum observed flux within the given peak.

| | Priors |
|---|---|
| Peak 1 | $\log[\tau_{r,g}] \sim \mathcal{U}(0, \log[300])$, $\log[\sigma_{r,g}] \sim \mathcal{U}(0, \log[300])$<br>$\log[F_{\rm peak,r}] \sim \mathcal{U}(\log[0.9F_{\rm max,r}], \log[10F_{\rm max,r}])$<br>$\log[F_{\rm peak,g}] \sim \mathcal{U}(\log[0.9F_{\rm max,g}], \log[10F_{\rm max,g}])$<br>$t_{\rm peak} \sim \mathcal{U}(58450, 58650)$ |
| Peak 2 | $\log[\tau_{r,g}] \sim \mathcal{U}(0, \log[300])$<br>$\log[F_{\rm peak,r}] \sim \mathcal{U}(\log[0.9F_{\rm max,r}], \log[10F_{\rm max,r}])$<br>$\log[F_{\rm peak,g}] \sim \mathcal{U}(\log[0.9F_{\rm max,g}], \log[10F_{\rm max,g}])$<br>$t_{\rm peak} \sim \mathcal{U}(59000, 59300)$ |

Table A.2: Overview of the varying set of Gaussians used for modelling the emission lines in the NOT and WiFeS spectra.

| Region | Components |
|---|---|
| H$\gamma$ | Single Gaussian for each of H$\gamma$ and [O III] 4363 Å. |
| He II | Single Gaussian component for each of He II 4686 Å, and [N III] 4640 Å. |
| H$\beta$ | Broad and narrow Gaussian component. |
| H$\alpha$ | Broad and narrow Gaussian component for H$\alpha$, single Gaussian for each of [N II] 6549 and 6583 Å. |
| [S II] doublet | Single Gaussian for each of [S II] 6716 and 6731 Å. |
| [O III] 5007 Å, [Fe XIV] 5303 Å, [Fe X] 6375 Å | Single Gaussian for each. |

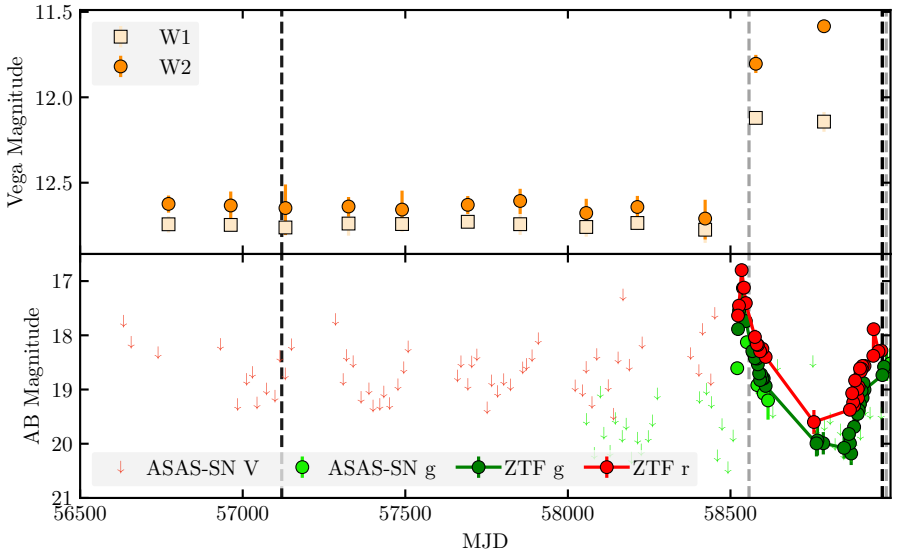


Figure A.1: Long term neoWISE, ASAS-SN and ZTF lightcurves of AT 2019avd. The early and late black dashed lines mark the 2015 *XMM-Newton* pointed and the 2020 eROSITA eRASS1 observations respectively. The early and late grey dashed lines mark the MJD that the NOT and first FLOYDS spectra were taken.

# Appendix B

# Additional information for the transient and variable source classification

## B.1 Feature sets

The following section contains the feature set names and the features they include for all considered feature subsets:

- **var_v001.001**: log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, log_ML_FLUX_1_over_w1_FLUX, log_ML_FLUX_1_over_w2_FLUX, log_ML_FLUX_2_over_w1_FLUX, log_ML_FLUX_2_over_w2_FLUX, log_ML_FLUX_3_over_w1_FLUX, log_ML_FLUX_3_over_w2_FLUX, w1mLQ, w2mLQ, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_1_over_phot_g_mean_flux, log_ML_FLUX_1_over_phot_bp_mean_flux, log_ML_FLUX_1_over_phot_rp_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_bp_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux, log_ML_FLUX_3_over_phot_bp_mean_flux, log_ML_FLUX_3_over_phot_rp_mean_flux, G_W1, G_W2.

- **var_v001.002**: log_X_RATIO_1, log_X_RATIO_3, ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, log_ML_FLUX_1_over_w1_FLUX, log_ML_FLUX_1_over_w2_FLUX, log_ML_FLUX_2_over_w1_FLUX, log_ML_FLUX_2_over_w2_FLUX, log_ML_FLUX_3_over_w1_FLUX, log_ML_FLUX_3_over_w2_FLUX, w1mLQ, w2mLQ, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_1_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_bp_mean_flux,

log_ML_FLUX_2_over_phot_rp_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux, G_W1, G_W2.

- **var_v001.003**: log_X_RATIO_1, log_X_RATIO_3, ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, log_ML_FLUX_1_over_w1_FLUX, log_ML_FLUX_1_over_w2_FLUX, log_ML_FLUX_2_over_w1_FLUX, log_ML_FLUX_2_over_w2_FLUX, log_ML_FLUX_3_over_w1_FLUX, log_ML_FLUX_3_over_w2_FLUX, w1mLQ, w2mLQ, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux, G_W1, G_W2.

- **var_v001.004**: log_X_RATIO_1, log_X_RATIO_3, ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, log_ML_FLUX_1_over_w2_FLUX, log_ML_FLUX_2_over_w1_FLUX, log_ML_FLUX_2_over_w2_FLUX, w1mLQ, w2mLQ, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux, G_W1, G_W2.

- **var_v001.005**: log_X_RATIO_1, log_X_RATIO_3, ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, log_ML_FLUX_2_over_w1_FLUX, log_ML_FLUX_2_over_w2_FLUX, w1mLQ, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux, G_W1, G_W2.

- **var_v001.006**: log_X_RATIO_1, log_X_RATIO_3, ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux, G_W1, G_W2.

- **var_v001.007**: log_X_RATIO_1, ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, G_W1, G_W2.

- **var_v001.008**: ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, G_W1, G_W2.

- **var_v001.009**: ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, G_W1, G_W2.
- **var_v001.010**: ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, bp_rp, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, G_W1, G_W2.
- **var_v001.011**: ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, bp_rp, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, G_W1, G_W2.
- **var_v001.012**: ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, bp_rp, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, G_W1.
- **var_v001.013**: ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, bp_rp, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, G_W2.
- **var_v001.014**: ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, bp_rp, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, G_W1.
- **var_v001.015**: ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, bp_rp, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, G_W1, log_X_RATIO_3.
- **var_v001.016**: ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, G_W1, log_X_RATIO_3.
- **var_v001.017**: ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, G_W1, log_X_RATIO_3.
- **var_v002.001**: log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, log_ML_FLUX_1_over_w1_FLUX, log_ML_FLUX_1_over_w2_FLUX, log_ML_FLUX_2_over_w1_FLUX, log_ML_FLUX_2_over_w2_FLUX, log_ML_FLUX_3_over_w1_FLUX, log_ML_FLUX_3_over_w2_FLUX, w1mLQ, w2mLQ.
- **var_v002.002**: log_X_RATIO_1, log_X_RATIO_3, ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_0_over_w2_FLUX, log_ML_FLUX_1_over_w1_FLUX, log_ML_FLUX_1_over_w2_FLUX, log_ML_FLUX_2_over_w1_FLUX, log_ML_FLUX_2_over_w2_FLUX, w1mLQ, w2mLQ.
- **var_v002.003**: log_X_RATIO_1, log_X_RATIO_3, ABS_BII, W1_W2, log_ML_FLUX_0_over_w1_FLUX, log_ML_FLUX_1_over_w1_FLUX, log_ML_FLUX_2_over_w1_FLUX, w1mLQ, w2mLQ.

- **var_v003.001**: log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_1_over_phot_g_mean_flux, log_ML_FLUX_1_over_phot_bp_mean_flux, log_ML_FLUX_1_over_phot_rp_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_bp_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux, log_ML_FLUX_3_over_phot_bp_mean_flux, log_ML_FLUX_3_over_phot_rp_mean_flux, ML_EXP_1, ML_EXP_2, ML_EXP_3.

- **var_v003.002**: log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_1_over_phot_g_mean_flux, log_ML_FLUX_1_over_phot_bp_mean_flux, log_ML_FLUX_1_over_phot_rp_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_bp_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux, log_ML_FLUX_3_over_phot_bp_mean_flux, log_ML_FLUX_3_over_phot_rp_mean_flux, ML_EXP_3.

- **var_v003.003**: log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_1_over_phot_g_mean_flux, log_ML_FLUX_1_over_phot_bp_mean_flux, log_ML_FLUX_1_over_phot_rp_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_bp_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux, log_ML_FLUX_3_over_phot_bp_mean_flux, log_ML_FLUX_3_over_phot_rp_mean_flux, ML_EXP_3.

- **var_v003.004**: log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_1_over_phot_g_mean_flux, log_ML_FLUX_1_over_phot_bp_mean_flux, log_ML_FLUX_1_over_phot_rp_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_bp_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux.

- **var_v003.004**: log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_1_over_phot_g_mean_flux, log_ML_FLUX_1_over_phot_bp_mean_flux, log_ML_FLUX_1_over_phot_rp_mean_flux,

log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_bp_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux.

- **var_v003.005**: log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_1_over_phot_g_mean_flux, log_ML_FLUX_1_over_phot_rp_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_bp_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux.
- **var_v003.006**: log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_1_over_phot_g_mean_flux, log_ML_FLUX_1_over_phot_rp_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_bp_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux.
- **var_v003.007**: log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_1_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_bp_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux.
- **var_v003.008**: log_X_RATIO_3, ABS_BII, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_1_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux, log_ML_FLUX_3_over_phot_g_mean_flux.
- **var_v003.009**: ABS_BII, phot_g_mean_mag, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_bp_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux.
- **var_v003.010**: ABS_BII, bp_rp, GAIA_ABS_MAG, PLX_SIG, PM_SIG, log_ML_FLUX_0_over_phot_g_mean_flux, log_ML_FLUX_0_over_phot_rp_mean_flux, log_ML_FLUX_2_over_phot_g_mean_flux, log_ML_FLUX_2_over_phot_rp_mean_flux.
- **var_v004.001**: ML_FLUX_1, ML_FLUX_2, ML_FLUX_3, ML_CTS_3, ML_EXP_3, log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII.
- **var_v004.002**: ML_CTS_3, ML_EXP_3, log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII.

- **var_v004.003**: ML_CTS_3, log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII.
- **var_v004.004**: ML_EXP_3, log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII.
- **var_v004.005**: log_X_RATIO_1, log_X_RATIO_2, log_X_RATIO_3, ABS_BII.

# Bibliography


Aasi J., et al., 2015, Classical and Quantum Gravity, 32, 115012

Acernese F., et al., 2015, Classical and Quantum Gravity, 32, 024001

Aizu K., 1973, Progress of Theoretical Physics, 49, 1184

Akerlof C., et al., 2003, Publications of the Astronomical Society of the Pacific, 115, 132

Alard C., 2000, Astronomy and Astrophysics Supplement Series, 144, 363

Alard C., Lupton R. H., 1998, The Astrophysical Journal, 503, 325

Alcock C., et al., 1997, Monthly Notices of the Royal Astronomical Society, 286, 483

Alexander K. D., Berger E., Guillochon J., Zauderer B. A., Williams P. K. G., 2016, The Astrophysical Journal, 819, L25

Alves Batista R., Silk J., 2017, Physical Review D, 96

Anninos P., Fragile P. C., Olivier S. S., Hoffman R., Mishra B., Camarda K., 2018, The Astrophysical Journal, 865, 3

Antonucci R. R. J., Cohen R. D., 1983, The Astrophysical Journal, 271, 564

Anzolin G., Tamburini F., de Martino D., Bianchini A., 2010, Astronomy and Astrophysics, 519, A69

Arcavi I., et al., 2014, The Astrophysical Journal, 793, 38

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

Assef R. J., et al., 2011, The Astrophysical Journal, 742, 93

Assef R. J., et al., 2013, The Astrophysical Journal, 772, 26

Assef R. J., Stern D., Noirot G., Jun H. D., Cutri R. M., Eisenhardt P. R. M., 2018, The Astrophysical Journal Supplement Series, 234, 23

Astier P., et al., 2006, Astronomy and Astrophysics, 447, 31

Bade N., Komossa S., Dahlem M., 1996, Astronomy and Astrophysics, 309

Bahcall N. A., Cen R., 1993, The Astrophysical Journal, 407, L49

Baker J. G., Menzel D. H., 1938, The Astrophysical Journal, 88, 52

Balbus S. A., Hawley J. F., 1998, Reviews of Modern Physics, 70, 1

Baldwin A., Phillips M. M., Terlevich R., 1981, Publications of the Astronomical Society of the Pacific, 93, 817

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, The Astrophysical Journal, 304, 15

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

Baron D., Poznanski D., 2017, Monthly Notices of the Royal Astronomical Society, 465, 4530

Baron D., Poznanski D., Watson D., Yao Y., Cox N. L. J., Prochaska J. X., 2015, Monthly Notices of the Royal Astronomical Society, 451, 332

Baron D., Stern J., Poznanski D., Netzer H., 2016, The Astrophysical Journal, 832, 8

Bauer F. E., et al., 2017, Monthly Notices of the Royal Astronomical Society, 467, 4841

Becker R. H., White R. L., Helfand D. J., 1995, The Astrophysical Journal, 450, 559

Begelman M. C., 2002, The Astrophysical Journal, 568, L97

Bellm E. C., 2014, arXiv preprint, arXiv:1410.8185

Bellm E. C., et al., 2019, Publications of the Astronomical Society of the Pacific, 131, 018002

Bender P. L., Begelman M. C., Gair J. R., 2013, Classical and Quantum Gravity, 30, 165017

Bentz M. C., et al., 2010, The Astrophysical Journal, 716, 993

Bernardini F., de Martino D., Mukai K., Russell D. M., Falanga M., Masetti N., Ferrigno C., Israel G., 2017, Monthly Notices of the Royal Astronomical Society, 470, 4815

Blagorodnova N., et al., 2017, The Astrophysical Journal, 844, 46

Blagorodnova N., et al., 2018, Publications of the Astronomical Society of the Pacific, 130, 035003

Blanchard P. K., et al., 2017, The Astrophysical Journal, 843, 106

Blanton M. R., Lupton R. H., Schlegel D. J., Strauss M. A., Brinkmann J., Fukugita M., Loveday J., 2005, The Astrophysical Journal, 631, 208

Bloom J. S., et al., 2011, Science, 333, 203

Bloom J. S., et al., 2012, Publications of the Astronomical Society of the Pacific, 124, 1175

Boella G., Butler R. C., Perola G. C., Piro L., Scarsi L., Bleeker J. A., 1997, Astronomy and Astrophysics Supplement Series, 122, 299

Boldt E., McDonald F. B., Riegler G., Serlemitsos P., 1966, Physical Review Letters, 17, 447

Boller T., Brandt W. N., Fink H., 1996, Astronomy and Astrophysics, 305, 53

Boller T., et al., 2021, Astronomy and Astrophysics, 647, A6

Bonnerot C., Rossi E. M., 2019, Monthly Notices of the Royal Astronomical Society, 484, 1301

Bonnerot C., Rossi E. M., Lodato G., 2016, Monthly Notices of the Royal Astronomical Society, 458, 3324

Boquien M., Burgarella D., Roehlly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, Astronomy and Astrophysics, 622, A103

Boroson T. A., Green R. F., 1992, The Astrophysical Journal Supplement Series, 80, 109

Bowen I. S., 1928, The Astrophysical Journal, 67, 1

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

Brassart M., Luminet J.-P., 2008, Astronomy and Astrophysics, 481, 259

Breiman L., 2001, Machine Learning, 45, 5

Bricman K., Gomboc A., 2020, The Astrophysical Journal, 890, 73

Brown T. M., et al., 2013, Publications of the Astronomical Society of the Pacific, 125, 1031

Brown M. L., et al., 2015, Proc. Sci., 2, 1365

Brown G. C., et al., 2017, Monthly Notices of the Royal Astronomical Society, 472, 4469

Buchner J., 2016, Statistics and Computing, 26, 383

Buchner J., 2019, Publications of the Astronomical Society of the Pacific, 131, 108005

Buchner J., et al., 2014, Astronomy and Astrophysics, 564, A125

Buchner J., Boller T., Bogensberger D., Malyali A., Nandra K., Wilms J., Dwelly T., Liu T., 2021, arXiv preprint, arXiv:2106.14529

Burgarella D., Buat V., Iglesias-Páramo J., 2005, Monthly Notices of the Royal Astronomical Society, 360, 1413

Burrows D. N., et al., 2011, Nature, 476, 421

Byram E. T., Chubb T. A., Friedman H., 1966, Science, 152, 66

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, The Astrophysical Journal, 533, 682

Cannizzo J. K., Gehrels N., 2009, The Astrophysical Journal, 700, 1047

Cannizzo J. K., Troja E., Lodato G., 2011, The Astrophysical Journal, 742, 32

Cappelluti N., et al., 2009, Astronomy and Astrophysics, 495, 9

Carpano S., Haberl F., Maitra C., Vasilopoulos G., 2018, Monthly Notices of the Royal Astronomical Society: Letters, 476, L45

Cash W., 1976, Astronomy and Astrophysics, 52, 307

Castro N., Protopapas P., Pichara K., 2017, The Astronomical Journal, 155, 16

Cenko B., et al., 2012, The Astrophysical Journal, 753, 77

Chambers K. C., et al., 2016, arXiv preprint, arXiv:1612.05560

Chan C.-H., Piran T., Krolik J. H., Saban D., 2019, The Astrophysical Journal, 881, 113

Charnock T., Moss A., 2017, The Astrophysical Journal, 837, L28

Childress M. J., Vogt F. P. A., Nielsen J., Sharp R. G., 2014, Astrophysics and Space Science, 349, 617

Chodil G., Mark H., Rodrigues R., Swift C. D., 1968, The Astrophysical Journal, 152, L45

Clerc N., et al., 2018, Astronomy and Astrophysics, 617, 92

Colbert E. J. M., Mushotzky R. F., 1999, The Astrophysical Journal, 519, 89

Corbel S., Coriat M., Brocksopp C., Tzioumis A. K., Fender R. P., Tomsick J. A., Buxton M. M., Bailyn C. D., 2013, Monthly Notices of the Royal Astronomical Society, 428, 2500

Coughlin E. R., Armitage P. J., Nixon C., Begelman M. C., 2017, Monthly Notices of the Royal Astronomical Society, 465, 3840

Coulter D. A., et al., 2017, Science, 358, 1556

Cropper M., 1990, Space Science Reviews, 54

D'Avanzo P., et al., 2012, Monthly Notices of the Royal Astronomical Society, 425, 506

Dai L., Escala A., Coppi P., 2013, The Astrophysical Journal, 775, L9

Dai L., McKinney J. C., Roth N., Ramirez-Ruiz E., Miller M. C., 2018, The Astrophysical Journal, 859, L20

De Colle F., Guillochon J., Naiman J., Ramirez-Ruiz E., 2012, The Astrophysical Journal, 760, 103

De Luca A., et al., 2020, Astronomy and Astrophysics, 634, L13

Decker French K., Zabludoff A. I., 2018, The Astrophysical Journal, 868, 99

Dennerl K., et al., 2020, in den Herder J.-W. A., Nakazawa K., Nikzad S., eds, Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray. No. December 2020. SPIE, p. 262

Dey A., et al., 2019, The Astronomical Journal, 157, 168

Djorgovski S., et al., 2008, Astronomische Nachrichten, 329, 263

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

Dong X.-B., Zhou H., Wang T., Wang J.-X., Li C., Zhou Y.-Y., 2005, The Astrophysical Journal, 620, 629

Dong X., Wang T., Wang J., Yuan W., Zhou H., Dai H., Zhang K., 2007a, Monthly Notices of the Royal Astronomical Society, 383, 581

Dong X., et al., 2007b, The Astrophysical Journal, 657, 700

Dopita M., Hart J., McGregor P., Oates P., Bloxham G., Jones D., 2007, Astrophysics and Space Science, 310, 255

Dopita M., et al., 2010, Astrophysics and Space Science, 327, 245

Drake A. J., et al., 2009, Astrophysical Journal, 696, 870

Duev D. A., et al., 2019, Monthly Notices of the Royal Astronomical Society, 489, 3582

Dwarkadas V. V., Gruszko J., 2012, Monthly Notices of the Royal Astronomical Society, 419, 1515

Earnshaw H. M., et al., 2016, Monthly Notices of the Royal Astronomical Society, 456, 3840

Earnshaw H. P., Roberts T. P., Middleton M. J., Walton D. J., Mateos S., 2019, Monthly Notices of the Royal Astronomical Society, 483, 5554

East W. E., 2014, The Astrophysical Journal, 795, 135

Edelson R., Nandra K., 1999, The Astrophysical Journal, 514, 682

Eisenhardt P. R. M., et al., 2020, The Astrophysical Journal Supplement Series, 247, 69

Esquej P., Saxton R. D., Freyberg M. J., Read A. M., Altieri B., Sanchez-Portal M., Hasinger G., 2007, Astronomy and Astrophysics, 462

Esquej P., et al., 2008a, Astronomy and Astrophysics, 489, 543

Esquej P., et al., 2008b, Astronomy and Astrophysics, 489, 543

Evans C. R., Kochanek C. S., 1989, The Astrophysical Journal, 346, L13

Evans C., Laguna P., Eracleous M., 2015, The Astrophysical Journal, 805, L19

Fabbiano G., Zezas A., Murray S. S., 2001, The Astrophysical Journal, 554, 1035

Farrell S. A., Murphy T., Lo K. K., 2015, The Astrophysical Journal, 813, 28

Feigelson E. D., Broos P., Gaffney III J. A., Garmire G., Hillenbrand L. A., Pravdo S. H., Townsley L., Tsuboi Y., 2002, The Astrophysical Journal, 574, 258

Feindt U., Nordin J., Rigault M., Brinnel V., Dhawan S., Goobar A., Kowalski M., 2019, Journal of Cosmology and Astroparticle Physics, 2019, 005

Fender R. P., Belloni T. M., Gallo E., 2004, Monthly Notices of the Royal Astronomical Society, 355, 1105

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

Feroz F., Hobson M. P., 2008, Monthly Notices of the Royal Astronomical Society, 384, 449

Filippenko A. V., 1989, The Astronomical Journal, 97, 726

Fragione G., Leigh N. W. C., Ginsburg I., Kocsis B., 2018, The Astrophysical Journal, 867, 119

Frail D. A., Kulkarni S. R., Ofek E. O., Bower G. C., Nakar E., 2012, The Astrophysical Journal, 747, 70

Frederick S., et al., 2019, The Astrophysical Journal, 883, 31

Frederick S., et al., 2020, arXiv preprint, arXiv:2010.08554

Freeman P., Doe S., Siemiginowska A., 2001, in Starck J.-L., Murtagh F. D., eds, Vol. 4477, Astronomical Data Analysis. SPIE, pp 76–87

Fremling C., et al., 2016, Astronomy and Astrophysics, 593, A68

Friedman H., Byram E. T., 1967, Science, 158, 257

Friedrich P., et al., 2014, in Takahashi T., den Herder J.-W. A., Bautz M., eds, Vol. 9144, Space Telescopes and Instrumentation 2014: Ultraviolet to Gamma Ray. p. 91444R

Furlong M., et al., 2015, Monthly Notices of the Royal Astronomical Society, 450, 4486

Garcia-Dias R., Prieto C. A., Almeida J. S., Ordovás-Pascual I., 2018, Astronomy & Astrophysics, 612, A98

Gaskell C. M., 2017, Monthly Notices of the Royal Astronomical Society, 467, stx094

Gaskell C. M., Rojas Lobos P. A., 2014, Monthly Notices of the Royal Astronomical Society: Letters, 438, L36

Gehrels N., et al., 2004, The Astrophysical Journal, 611, 1005

Gendre B., Galli A., Boer M., Galassi M., Palmer D., Fenimore E., 2008, in AIP Conference Proceedings. AIP, pp 72–75

Gendreau K. C., et al., 2016, in den Herder J.-W. A., Takahashi T., Bautz M., eds, Vol. 9905, Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray. p. 99051H

Gezari S., et al., 2008, The Astrophysical Journal, 676, 944

Gezari S., et al., 2009, The Astrophysical Journal, 698, 1367

Gezari S., et al., 2012, Nature, 485, 217

Gezari S., Cenko S. B., Arcavi I., 2017, The Astrophysical Journal, 851, L47

Gezari S., Frederick S., van Velzen S., Tartaglia L., Sollerman J., Goobar A., Perley D., Kulkarni S., 2020, The Astronomer's Telegram, 13717, 1

Giacconi R., 2003, Reviews of Modern Physics, 75, 995

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

Giacconi R., Kellogg E., Gorenstein P., Gursky H., Tananbaum H., 1971, The Astrophysical Journal, 165, L27

Giustini M., Miniutti G., Saxton R. D., 2020, Astronomy and Astrophysics, 636, L2

Glennie A., Jonker P. G., Fender R. P., Nagayama T., Pretorius M. L., 2015, Monthly Notices of the Royal Astronomical Society, 450, 3765

Gliozzi M., Williams J. K., 2020, Monthly Notices of the Royal Astronomical Society, 491, 532

Goldstein D. A., et al., 2015, The Astronomical Journal, 150, 82

Graham M. J., Djorgovski S. G., Drake A. J., Stern D., Mahabal A. A., Glikman E., Larson S., Christensen E., 2017, Monthly Notices of the Royal Astronomical Society, 470, 4112

Graham M. J., et al., 2019, Publications of the Astronomical Society of the Pacific, 131, 078001

Greene J. E., 2012, Nature Communications, 3, 1304

Greiner J., Di Stefano R., 2002, Astronomy and Astrophysics, 387, 944

Greiner J., Schwarz R., Zharikov S., Orio M., 2000, Astronomy and Astrophysics, 362

Gruber D. E., Blanco P. R., Heindl W. A., Pelling M. R., Rothschild R. E., Hink P. L., 1996, Astronomy and Astrophysics Supplement Series, 120, 641

Grupe D., Leighly K., 1999, Astronomy and Astrophysics, 350, L31

Güdel M., 2004, The Astronomy and Astrophysics Review, 12, 71

Guillochon J., Ramirez-Ruiz E., 2013, The Astrophysical Journal, 767, 25

Guillochon J., Ramirez-Ruiz E., 2015, The Astrophysical Journal, 809, 166

Guillochon J., Ramirez-Ruiz E., Rosswog S., Kasen D., 2009, The Astrophysical Journal, 705, 844

Guillochon J., Nicholl M., Villar V. A., Mockler B., Narayan G., Mandel K. S., Berger E., Williams P. K. G., 2018, The Astrophysical Journal Supplement Series, 236, 6

Gultekin K., Miller M. C., Hamilton D. P., 2004, The Astrophysical Journal, 616, 221

Gupta A., Galeazzi M., Williams B., 2011, The Astrophysical Journal, 731, 63

Gursky H., Giacconi R., Gorenstein P., Waters J. R., Oda M., Bradt H., Garmire G., Sreekantan B. V., 1966, The Astrophysical Journal, 146, 310

Gurzadian V. G., Ozernoi L. M., 1981, Astronomy and Astrophysics, 95, 39

HI4PI Collaboration et al., 2016, Astronomy and Astrophysics, 594, A116

Haardt F., Maraschi L., 1991, The Astrophysical Journal, 380, L51

Haardt F., Maraschi L., 1993, The Astrophysical Journal, 413, 507

Haas R., Shcherbakov R. V., Bode T., Laguna P., 2012, The Astrophysical Journal, 749, 117

Haiman Z., et al., 2005, arXiv preprint, arXiv:astro-ph/0507013

Harries J. R., McCracken K. G., Francey R. J., Fenton A. G., 1967, Nature, 215, 38

Harrison F. A., et al., 2013, The Astrophysical Journal, 770, 103

Hayasaki K., Stone N., Loeb A., 2013, Monthly Notices of the Royal Astronomical Society, 434, 909

Hayasaki K., Stone N., Loeb A., 2016, Monthly Notices of the Royal Astronomical Society, 461, 3760

Heise J., Zand J. i. Kippen R. M., Woods P. M., 2001, in , Vol. 2, Gamma-Ray Bursts in the Afterglow Era. Springer-Verlag, Berlin/Heidelberg, pp 16–21

Henden A. A., Levine S., Terrell D., Welch D. L., 2015, in American Astronomical Society Meeting Abstracts #225. p. 336.16

Henze M., et al., 2009, Astronomy and Astrophysics, 498, L13

Heymans C., et al., 2013, Monthly Notices of the Royal Astronomical Society, 432, 2433

Hills J. G., 1975, Nature, 254, 295

Hinkle J. T., et al., 2020, Monthly Notices of the Royal Astronomical Society, 500, 1673

Holoien T. W.-S., et al., 2014, Monthly Notices of the Royal Astronomical Society, 445, 3263

Holoien T. W.-S., et al., 2016, Monthly Notices of the Royal Astronomical Society, 463, 3813

Holoien T. W.-S., et al., 2019a, The Astrophysical Journal, 880, 120

Holoien T. W.-S., et al., 2019b, The Astrophysical Journal, 883, 111

Hung T., et al., 2017, The Astrophysical Journal Supplement Series, 238, 15

Irwin J. A., et al., 2016, Nature, 538, 356

Ishida E. E., et al., 2019, arXiv preprint, arXiv:1909.13260

Israel G. L., et al., 2017, Science, 355, 817

Ivezić , et al., 2019, The Astrophysical Journal, 873, 111

Jagoda N., Austin G., Mickiewicz S., Goddard R., 1972, IEEE Transactions on Nuclear Science, 19, 579

Jansen F., et al., 2001, Astronomy and Astrophysics, 365, L1

Jonker P. G., et al., 2013, The Astrophysical Journal, 779, 14

Jonker P. G., Stone N. C., Generozov A., Velzen S. v., Metzger B., 2020, The Astrophysical Journal, 889, 166

Kahabka P., van den Heuvel E. P. J., 1997, Annual Review of Astronomy and Astrophysics, 35, 69

Kajava J. J. E., Giustini M., Saxton R. D., Miniutti G., 2020, Astronomy and Astrophysics, 639, A100

Kara E., Dai L., Reynolds C. S., Kallman T., 2018, Monthly Notices of the Royal Astronomical Society, 474, 3593

Kasen D., 2006, The Astrophysical Journal, 649, 939

Kaspi S., Maoz D., Netzer H., Peterson B. M., Vestergaard M., Jannuzi B. T., 2005, The Astrophysical Journal, 629, 61

Kataoka J., et al., 2008, The Astrophysical Journal, 672, 787

Kauffmann G., et al., 2003, Monthly Notices of the Royal Astronomical Society, 346, 1055

Kawana K., Tanikawa A., Yoshida N., 2018, Monthly Notices of the Royal Astronomical Society, 477, 3449

Keller S. C., et al., 2007, Publications of the Astronomical Society of Australia, 24, 1

Kepler S. O., Kleinman S. J., Nitta A., Koester D., Castanheira B. G., Giovannini O., Costa A. F. M., Althaus L., 2007, Monthly Notices of the Royal Astronomical Society, 375, 1315

Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, Monthly Notices of the Royal Astronomical Society, 372, 961

Khabibullin I., Sazonov S., 2014, Monthly Notices of the Royal Astronomical Society, 444, 1041

Khabibullin I., Sazonov S., Sunyaev R., 2014, Monthly Notices of the Royal Astronomical Society, 437, 327

Kim D.-C., Yoon I., Evans A. S., 2018, The Astrophysical Journal, 861, 51

King A., 2020, Monthly Notices of the Royal Astronomical Society: Letters, 493, L120

King A. R., Davies M. B., Ward M. J., Fabbiano G., Elvis M., 2001, The Astrophysical Journal, 552, L109

Kobayashi S., Laguna P., Phinney E. S., Meszaros P., 2004, The Astrophysical Journal, 615, 855

Kochanek C. S., et al., 2017, Publications of the Astronomical Society of the Pacific, 129, 104502

Kollatschny W., 2003, Astronomy and Astrophysics, 407, 461

Kollatschny W., Ulbrich K., Zetzl M., Kaspi S., Haas M., 2014, Astronomy and Astrophysics, 566, A106

Kollmeier J. A., et al., 2017, arXiv preprint, arXiv:1711.03234

Komossa S., 2015, Journal of High Energy Astrophysics, 7, 148

Komossa S., Bade N., 1999, Astronomy and Astrophysics, 343, 775

Komossa S., Greiner J., 1999, Astronomy and Astrophysics, 349, L45

Korista K. T., Ferland G. J., 1989, The Astrophysical Journal, 343, 678

Korista K. T., et al., 1995, The Astrophysical Journal Supplement Series, 97, 285

Kormendy J., Ho L. C., 2013, Annual Review of Astronomy and Astrophysics, 51, 511

Kotze M. M., Charles P. A., 2012, Monthly Notices of the Royal Astronomical Society, 420, 1575

Kouveliotou C., Meegan C. A., Fishman G. J., Bhat N. P., Briggs M. S., Koshut T. M., Paciesas W. S., Pendleton G. N., 1993, The Astrophysical Journal, 413, L101

Kouveliotou C., et al., 1998, Nature, 393, 235

Kovačević J., Popović L., Dimitrijević M. S., 2010, The Astrophysical Journal Supplement Series, 189, 15

Krolik J. H., Piran T., 2011, The Astrophysical Journal, 743, 134

Krolik J. H., Piran T., 2012, The Astrophysical Journal, 749, 92

Krone-Martins A., Moitinho A., 2014, Astronomy and Astrophysics, 561, A57

Kulkarni S. R., 2020, arXiv preprint, arXiv:2004.03511

Lacy J. H., Townes C. H., Hollenbach D. J., 1982, The Astrophysical Journal, 262, 120

Lada C. J., 1987, Symposium - International Astronomical Union, 115, 1

Lada C. J., 2006, The Astrophysical Journal, 640, L63

Lasota J.-P., 2001, New Astronomy Reviews, 45, 449

Law-Smith J., MacLeod M., Guillochon J., Macias P., Ramirez-Ruiz E., 2017a, The Astrophysical Journal, 841, 132

Law-Smith J., Ramirez-Ruiz E., Ellison S. L., Foley R. J., 2017b, The Astrophysical Journal, 850, 22

Lawrence A., Gezari S., Elvis M., Ward M., Smartt S., Smith K., Wright D., 2012, EPJ Web of Conferences, 39, 03002

Lehmer B. D., et al., 2012, The Astrophysical Journal, 752, 46

Leloudas G., et al., 2019, The Astrophysical Journal, 887, 218

Levan A. J., et al., 2011, Science, 333, 199

Lin D., et al., 2017a, Nature Astronomy, 1, 1

Lin D., Godet O., Ho L. C., Barret D., Webb N. A., Irwin J. A., 2017b, Monthly Notices of the Royal Astronomical Society, 468, 783

Lin D., et al., 2018, Nature Astronomy, 2, 656

Liu Q. Z., van Paradijs J., van den Heuvel E. P. J., 2001, Astronomy and Astrophysics, 368, 1021

Liu Q. Z., van Paradijs J., van den Heuvel E. P. J., 2006, Astronomy and Astrophysics, 455, 1165

Liu Q. Z., van Paradijs J., van den Heuvel E. P. J., 2007, Astronomy and Astrophysics, 469, 807

Liu F. K., Li S., Chen X., 2009, The Astrophysical Journal, 706, L133

Liu Z., Li D., Liu H.-Y., Lu Y., Yuan W., Dou L., Shen R.-F., 2020, The Astrophysical Journal, 894, 93

Lo K. K., Farrell S., Murphy T., Gaensler B. M., 2014, The Astrophysical Journal, 786, 20

Lochner M., McEwen J. D., Peiris H. V., Lahav O., Winter M. K., 2016, The Astrophysical Journal Supplement Series, 225, 31

Lodato G., King A. R., Pringle J. E., 2009, Monthly Notices of the Royal Astronomical Society, 392, 332

Loeb A., Ulmer A., 1997, The Astrophysical Journal, 489, 573

Long K. S., Helfand D. J., Grabelsky D. A., 1981, The Astrophysical Journal, 248, 925

Loveday J., 1997, The Astrophysical Journal, 489, 29

Luminet J., Pichon B., 1989a, Astronomy and Astrophysics, 209, 85

Luminet J., Pichon B., 1989b, Astronomy and Astrophysics, 209, 103

Lynden-Bell D., 1969, Nature, 223, 690

MacLeod M., Goldstein J., Ramirez-Ruiz E., Guillochon J., Samsing J., 2014, The Astrophysical Journal, 794, 9

MacLeod C. L., et al., 2016, Monthly Notices of the Royal Astronomical Society, 457, 389

Magorrian J., Tremaine S., 1999, Monthly Notices of the Royal Astronomical Society, 309, 447

Mainetti D., Campana S., Colpi M., 2016, Astronomy and Astrophysics, 592, 1

Mainzer A., et al., 2011, The Astrophysical Journal, 731, 53

Mainzer A., et al., 2014, The Astrophysical Journal, 792, 30

Maksym W. P., Ulmer M. P., Eracleous M. C., Guennou L., Ho L. C., 2013, Monthly Notices of the Royal Astronomical Society, 435, 1904

Malyali A., Rau A., Nandra K., 2019, Monthly Notices of the Royal Astronomical Society, 489, 5413

Malyali A., et al., 2020, The Astronomer's Telegram, 13712, 1

Malyali A., et al., 2021, Astronomy and Astrophysics, 647, A9

Mandel I., Levin Y., 2015, The Astrophysical Journal, 805, L4

Marsaglia G., 1972, The Annals of Mathematical Statistics, 43, 645

Martin D. C., et al., 2005, The Astrophysical Journal, 619, L1

Marton G., et al., 2019, Monthly Notices of the Royal Astronomical Society, 487, 2522

Masci F. J., et al., 2019, Publications of the Astronomical Society of the Pacific, 131, 018003

Matsuoka M., et al., 2009, Publications of the Astronomical Society of Japan, 61, 999

Matt G., Guainazzi M., Maiolino R., 2003, Monthly Notices of the Royal Astronomical Society, 342, 422

McConnachie A. W., 2012, Astronomical Journal, 144, 4

McConnell N. J., Ma C.-P., 2013, The Astrophysical Journal, 764, 184

McHardy I. M., Koerding E., Knigge C., Uttley P., Fender R. P., 2006, Nature, 444, 730

Meidinger N., et al., 2020, in den Herder J.-W. A., Nakazawa K., Nikzad S., eds, Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray. No. December 2020. SPIE, p. 260

Merloni A., et al., 2012, arXiv preprint, arXiv:1209.3114

Merloni A., et al., 2015, Monthly Notices of the Royal Astronomical Society, 452, 69

Metzger B. D., Quataert E., Thompson T. A., 2008, Monthly Notices of the Royal Astronomical Society, 385, 1455

Meusinger H., et al., 2010, Astronomy and Astrophysics, 512, A1

Meyer F., Meyer-Hofmeister E., 1981, Astronomy and Astrophysics, 104, L10

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

Middleton M., Uttley P., Done C., 2011, Monthly Notices of the Royal Astronomical Society, 417, 250

Miller B. P., Gallo E., Greene J. E., Kelly B. C., Treu T., Woo J.-H., Baldassare V., 2015, The Astrophysical Journal, 799, 98

Miniutti G., et al., 2019, Nature, 573, 381

Mockler B., Guillochon J., Ramirez-Ruiz E., 2019, The Astrophysical Journal, 872, 151

Möller A., de Boissière T., 2020, Monthly Notices of the Royal Astronomical Society, 491, 4277

Montmerle T., Grosso N., Tsuboi Y., Koyama K., 2000, The Astrophysical Journal, 532, 1097

Mukai K., 2017, Publications of the Astronomical Society of the Pacific, 129, 062001

Muthukrishna D., Narayan G., Mandel K. S., Biswas R., Hložek R., 2019, Publications of the Astronomical Society of the Pacific, 131, 118002

Nagao T., Taniguchi Y., Murayama T., 2000, The Astronomical Journal, 119, 2605

Nandra K., Pounds K. A., 1994, Monthly Notices of the Royal Astronomical Society, 268, 405

Nandra K., George I. M., Mushotzky R. F., Turner T. J., Yaqoob T., 1997, The Astrophysical Journal, 476, 70

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

Nenkova M., Sirocky M. M., Nikutta R., Ivezić v., Elitzur M., 2008, The Astrophysical Journal, 685, 160

Netzer H., Elitzur M., Ferland G. J., 1985, The Astrophysical Journal, 299, 752

Neustadt J. M. M., et al., 2020, Monthly Notices of the Royal Astronomical Society, 494, 2538

Noda H., Done C., 2018, Monthly Notices of the Royal Astronomical Society, 480, 3898

Nordin J., Brinnel V., Giomi M., Santen J. V., Gal-yam A., Yaron O., Schulze S., 2019, Transient Name Server Discovery Report, 2019-236, 1

Nousek J. A., et al., 2006, The Astrophysical Journal, 642, 389

Nyholm A., et al., 2020, Astronomy and Astrophysics, 637, A73

Nysewander M., Fruchter A. S., Pe'er A., 2009, The Astrophysical Journal, 701, 824

Ofek E. O., et al., 2013, The Astrophysical Journal, 763, 42

Ozawa H., Grosso N., Montmerle T., 2005, Astronomy and Astrophysics, 438, 661

Padovani P., et al., 2017, The Astronomy and Astrophysics Review, 25, 2

Pasham D. R., Strohmayer T. E., Mushotzky R. F., 2014, Nature, 513, 74

Pasham D. R., et al., 2015, The Astrophysical Journal, 805, 68

Payne A. V., et al., 2021, The Astrophysical Journal, 910, 125

Pelat D., Alloin D., Bica E., 1987, Astronomy and Astrophysics, 182, 9

Penston M. V., Pérez E., 1984, Monthly Notices of the Royal Astronomical Society, 211, 33P

Perlmutter S., et al., 1999, The Astrophysical Journal, 517, 565

Phinney E., 1989, Symposium - International Astronomical Union, 136, 543

Piran T., Svirski G., Krolik J., Cheng R. M., Shiokawa H., 2015, The Astrophysical Journal, 806, 164

Planck Collaboration XIII 2016, Astronomy and Astrophysics, 594, A13

Pojmański G., 2001, in Paczynski B., Chen W.-P., Lemme C., eds, Astronomical Society of the Pacific Conference Series Vol. 246, IAU Colloq. 183: Small Telescope Astronomy on Global Scales. p. 53

Predehl P., 2017, Astronomische Nachrichten, 338, 159

Predehl P., et al., 2021, Astronomy and Astrophysics, 647, A1

Protopapas P., Giammarco J. M., Faccioli L., Struble M. F., Dave R., Alcock C., 2006, Monthly Notices of the Royal Astronomical Society, 369, 677

Pye J. P., Rosen S., Fyfe D., Schröder A. C., 2015, Astronomy and Astrophysics, 581, A28

Rakshit S., Stalin C. S., Chand H., Zhang X.-G., 2017, The Astrophysical Journal Supplement Series, 229, 39

Ramirez-Ruiz E., Rosswog S., 2009, The Astrophysical Journal, 697, L77

Ramsay G., Wheatley P. J., Norton A. J., Hakala P., Baskill D., 2008, Monthly Notices of the Royal Astronomical Society, 387, 1157

Rau A., et al., 2009, Publications of the Astronomical Society of the Pacific, 121, 1334

Rees M. J., 1988, Nature, 333, 523

Reiprich T. H., Bohringer H., 2002, The Astrophysical Journal, 567, 716

Reis R. C., Miller J. M., Reynolds M. T., Gültekin K., Maitra D., King A. L., Strohmayer T. E., 2012, Science, 337, 949

Reis R. C., Wheatley P. J., Gänsicke B. T., Osborne J. P., 2013, Monthly Notices of the Royal Astronomical Society, 430, 1994

Reis I., Poznanski D., Baron D., Zasowski G., Shahaf S., 2018, Monthly Notices of the Royal Astronomical Society, 476, 2117

Revsbech E. A., Trotta R., van Dyk D. A., 2018, Monthly Notices of the Royal Astronomical Society, 473, 3969

Reylé C., 2018, Astronomy and Astrophysics, 619, L8

Ricci C., et al., 2020, The Astrophysical Journal, 898, L1

Ricker G. R., et al., 2003, in Ricker G. R., Vanderspek R. K., eds, American Institute of Physics Conference Series Vol. 662, Gamma-Ray Burst and Afterglow Astronomy 2001: A Workshop Celebrating the First Year of the HETE Mission. pp 3–16

Ricker G. R., et al., 2014, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Ritter H., Kolb U., 2003, Astronomy and Astrophysics, 404, 301

Rodriguez C. L., Morscher M., Pattabiraman B., Chatterjee S., Haster C.-J., Rasio F. A., 2015, Physical Review Letters, 115, 051101

Romano P., et al., 2015, Astronomy and Astrophysics, 576, L4

Roming P. W. A., et al., 2005, Space Science Reviews, 120, 95

Ross N. P., et al., 2018, Monthly Notices of the Royal Astronomical Society, 480, 4468

Rossi E. M., Stone N. C., Law-Smith J. A. P., Macleod M., Lodato G., Dai J. L., Mandel I., 2021, Space Science Reviews, 217, 40

Rosswog S., Ramirez-Ruiz E., Hix W. R., 2008, The Astrophysical Journal, 679, 1385

Rosswog S., Ramirez-Ruiz E., Hix W. R., 2009, The Astrophysical Journal, 695, 404

Roth N., Kasen D., 2018, The Astrophysical Journal, 855, 54

Roth N., Kasen D., Guillochon J., Ramirez-Ruiz E., 2016, The Astrophysical Journal, 827, 3

Rozyczka M., Bodenheimer P., Lin D. N. C., 1995, Monthly Notices of the Royal Astronomical Society, 276, 597

Rutledge R. E., 1998, Publications of the Astronomical Society of the Pacific, 110, 754

Sagiv I., et al., 2014, The Astronomical Journal, 147, 79

Salvato M., et al., 2018, Monthly Notices of the Royal Astronomical Society, 473, 4937

Sandage A., et al., 1966, The Astrophysical Journal, 146, 316

Sato S., et al., 2017, Journal of Physics: Conference Series, 840, 012010

Savitzky A., Golay M. J. E., 1964, Analytical Chemistry, 36, 1627

Saxton R. D., Read A. M., Esquej P., Komossa S., Dougherty S., Rodriguez-Pascual P., Barrado D., 2012, Astronomy and Astrophysics, 541, 1

Saxton R. D., Read A. M., Komossa S., Lira P., Alexander K. D., Wieringa M. H., 2017, Astronomy & Astrophysics, 598, A29

Saxton R. D., et al., 2019, Astronomy and Astrophysics, 630, A98

Saxton R., Komossa S., Auchettl K., Jonker P. G., 2020, Space Science Reviews, 216, 85

Schechter P., 1976, The Astrophysical Journal, 203, 297

Schmid C., 2012, PhD thesis, Friedrich-Alexander-Universität Erlangen Nürnberg

Schmidt B. P., et al., 1998, The Astrophysical Journal, 507, 46

Schwope A. D., Scipione V., Traulsen I., Schwarz R., Granzer T., Pires A. M., Thorstensen J. R., 2014, Astronomy and Astrophysics, 561, 1

Schwope A., Buckley D. A. H., Malyali A., Potter S., König O., Arcodia R., Gromadzki M., Rau A., 2021, arXiv preprint, arXiv:2106.14540

Sell P. H., Maccarone T. J., Kotak R., Knigge C., Sand D. J., 2015, Monthly Notices of the Royal Astronomical Society, 450, 4198

Shakura N. I., Sunyaev R. A., 1973, Astronomy and Astrophysics, 500, 33

Shappee B. J., et al., 2014, The Astrophysical Journal, 788, 48

Shcherbakov R. V., Pe 'er A., Reynolds C. S., Haas R., Bode T., Laguna P., 2013, The Astrophysical Journal, 769, 85

Shields G. A., 1978, Nature, 272, 706

Shiokawa H., Krolik J. H., Cheng R. M., Piran T., Noble S. C., 2015a, The Astrophysical Journal, 804, 85

Shiokawa H., Krolik J. H., Cheng R. M., Piran T., Noble S. C., 2015b, The Astrophysical Journal, 804, 85

Shklovsky I. S., 1967, The Astrophysical Journal, 148, L1

Sidoli L., Postnov K. A., Belfiore A., Marelli M., Salvetti D., Salvaterra R., De Luca A., Esposito P., 2019, Monthly Notices of the Royal Astronomical Society, 487, 420

Simm T., Salvato M., Saglia R., Ponti G., Lanzuisi G., Trakhtenbrot B., Nandra K., Bender R., 2016, Astronomy and Astrophysics, 585, A129

Simmonds C., Buchner J., Salvato M., Hsu L.-T., Bauer F. E., 2018, Astronomy and Astrophysics, 618, A66

Skrutskie M. F., et al., 2006, The Astronomical Journal, 131, 1163

Smart R. L., et al., 2021, Astronomy and Astrophysics, 649, A6

Smartt S. J., et al., 2017, Nature, 551, 75

Smith J. F., Courtier G. M., 1976, Proceedings of the Royal Society of London. A., 350, 421

Smith K. W., et al., 2019, Research Notes of the AAS, 3, 26

Śniegowska M., Czerny B., 2019, arXiv preprint, arXiv:1904.06767

Soderberg A., 2010, arXiv preprint, arXiv:0902.3674

Soderberg A. M., et al., 2006, Nature, 442, 1014

Soderberg A. M., et al., 2008, Nature, 453, 469

Sooknunan K., et al., 2021, Monthly Notices of the Royal Astronomical Society, 502, 206

Springel V., et al., 2005, Nature, 435, 629

Starrfield S., Truran J. W., Sparks W. M., Krautter J., MacDonald J., 1990, International Astronomical Union Colloquium, 122, 306

Stern D., et al., 2012, The Astrophysical Journal, 753, 30

Stone N. C., Metzger B. D., 2016, Monthly Notices of the Royal Astronomical Society, 455, 859

Stone N., Sari R., Loeb A., 2013, Monthly Notices of the Royal Astronomical Society, 435, 1809

Storchi-Bergmann T., Schimoia J. S., Peterson B. M., Elvis M., Denney K. D., Eracleous M., Nemmen R. S., 2017, The Astrophysical Journal, 835, 236

Strüder L., et al., 2001, Astronomy and Astrophysics, 365, L18

Sunyaev R., et al., 2021, arXiv preprint, arXiv:2104.13267

Tanaka Y., Inoue H., 1994, Publications of the Astronomical Society of Japan, 46, L37

Tananbaum H., Gursky H., Kellogg E. M., Levinson R., Schreier E., Giacconi R., 1972a, The Astrophysical Journal, 174, L143

Tananbaum H., Gursky H., Kellogg E., Giacconi R., Jones C., 1972b, The Astrophysical Journal, 177, L5

Taylor B. G., Andresen R. D., Peacock A., Zobl R., 1981, Space Science Reviews, 30, 479

Tenzer C., Warth G., Kendziorra E., Santangelo A., 2010, in Holland A. D., Dorn D. A., eds, Vol. 7742, Spie. p. 77420Y

Thorp S., Chadwick E., Sesana A., 2019, Monthly Notices of the Royal Astronomical Society, 488, 4042

Tody D., 1986, in Crawford D. L., ed., Vol. 0627, Instrumentation in Astronomy VI. p. 733

Tohline J. E., Osterbrock D. E., 1976, The Astrophysical Journal, 210, L117

Tonry J. L., et al., 2018, Publications of the Astronomical Society of the Pacific, 130, 064505

Torrealba G., et al., 2019, Monthly Notices of the Royal Astronomical Society, 488, 2743

Tout C. A., Pols O. R., Eggleton P. P., Han Z., 1996, Monthly Notices of the Royal Astronomical Society, 281, 257

Trakhtenbrot B., et al., 2019a, Nature Astronomy, 3, 242

Trakhtenbrot B., et al., 2019b, The Astrophysical Journal, 883, 94

Trakhtenbrot B., Arcavi I., Ricci C., Burke J., 2020, Transient Name Server Classification Report, 2020-1391, 1

Traulsen I., Reinsch K., Schwarz R., Dreizler S., Beuermann K., Schwope A. D., Burwitz V., 2010, Astronomy and Astrophysics, 516, A76

Trümper J., 1982, Advances in Space Research, 2, 241

Tsunemi H., Kitamoto S., Manabe M., Miyamoto S., Yamashita K., Nakagawa M., 1989, Publications of the Astronomical Society of Japan, 41, 391

Ulmer A., 1999, The Astrophysical Journal, 514, 180

Ulmer A., Paczynski B., Goodman J., 1998, Astronomy and Astrophysics, 333, 379

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

Valenti S., et al., 2014, Monthly Notices of the Royal Astronomical Society: Letters, 438, L101

Véron-Cetty M.-P., Véron P., 2010, Astronomy and Astrophysics, 518, A10

Vestergaard M., Peterson B. M., 2006, The Astrophysical Journal, 641, 689

Villar V. A., Cranmer M., Contardo G., Ho S., Lin J. Y.-Y., 2020, arXiv preprint, arXiv:2010.11194, pp 1–6

Walton D. J., Miller J. M., Harrison F. A., Fabian A. C., Roberts T. P., Middleton M. J., Reis R. C., 2013, The Astrophysical Journal, 773, L9

Wang T.-G., Zhou H.-Y., Komossa S., Wang H.-Y., Yuan W., Yang C., 2012, The Astrophysical Journal, 749, 115

Wei J., et al., 2016, arXiv preprint, arXiv:1610.06892

Weisskopf M. C., Tananbaum H. D., Van Speybroeck L. P., O'Dell S. L., 2000. pp 2–16

Wevers T., et al., 2019, Monthly Notices of the Royal Astronomical Society, 488, 4816

Wheatley P. J., Mauche C. W., Mattei J. A., 2003, Monthly Notices of the Royal Astronomical Society, 345, 49

Winkler C., et al., 2003, Astronomy and Astrophysics, 411, L1

Wolter H., 1952, Annalen der Physik, 445, 94

Woosley S. E., 1993, The Astrophysical Journal, 405, 273

Wright E. L., et al., 2010, The Astronomical Journal, 140, 1868

Xue Y. Q., et al., 2019, Nature, 568, 198

Young P. J., Shields G. A., Wheeler J. C., 1977, The Astrophysical Journal, 212, 367

Yu Y.-W., Chen A., Li X.-D., 2019, The Astrophysical Journal, 877, L21

Yuan W., et al., 2018, in den Herder J.-W. A., Nakazawa K., Nikzad S., eds, Vol. 1069925, Space Telescopes and Instrumentation 2018: Ultraviolet to Gamma Ray. SPIE, p. 76

Zalamea I., Menou K., Beloborodov A. M., 2010, Monthly Notices of the Royal Astronomical Society: Letters, 409, L25

Zauderer B. A., et al., 2011, Nature, 476, 425

Zauderer B. A., Berger E., Margutti R., Pooley G. G., Sari R., Soderberg A. M., Brunthaler A., Bietenholz M. F., 2013, The Astrophysical Journal, 767, 152

Zhang S. N., et al., 2016, Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray, 9905, 99051Q

Zhang G., Gelfand J. D., Russell D. M., Lewis F., Masetti N., Bernardini F., Andruchow I., Zibecchi L., 2017, Monthly Notices of the Royal Astronomical Society, 469, 4236

de Jong R. S., et al., 2019, The Messenger, 175, 3

van Velzen S., 2018, The Astrophysical Journal, 852, 72

van Velzen S., et al., 2011, The Astrophysical Journal, 741, 73

van Velzen S., Mendez A. J., Krolik J. H., Gorjian V., 2016, The Astrophysical Journal, 829, 19

van Velzen S., et al., 2019, The Astrophysical Journal, 872, 198

van Velzen S., Holoien T. W., Onori F., Hung T., Arcavi I., 2020, Space Science Reviews, 216, 124

van Velzen S., et al., 2021, The Astrophysical Journal, 908, 4

van den Heuvel E. P. J., Bhattacharya D., Nomoto K., Rappaport S. A., 1992, Astronomy and Astrophysics, 262, 97

# Acknowledgements

I am very grateful and fortunate to have been supported by many different people over the course of this PhD.

Firstly, I would like to thank both of my supervisors, Arne and Paul, for their immeasurable support, care[1], and teaching over the last four years.

The High Energy Group at MPE has always been a pleasure to work in, and is full of supportive, friendly people. In particular, I would like to thank Mara (for her extensive support throughout the PhD, and for taking a chance on a random email from an undergraduate student in England back in 2016 and organising an eight-week research project- without Mara this thesis would not have been written), Andrea (for his invaluable scientific feedback through this thesis), and to Johannes and J Michael (for their statistical expertise and teachings/ preachings). In addition, I would like to thank Birgit, who has always been able to help with any general administrative problems, and Harald, who has been incredibly patient with any IT-related problems I have had- no matter how trivial!

Living in Munich has been an unforgettable experience, and was made even better by the new people I've met here. Special thanks go to Felix, Jacob, Ricky and Sebastiano, for being not like normal physicists, and also to my friends back in London for coming to visit pre-corona times, and meeting up regularly (virtually) during the pandemic.

Lastly, I am deeply indebted to my Mum, Dad and Brother, who have provided immense support and encouragement over the last 26 years, and to my incredible girlfriend, Rebecca, who has helped and supported me enormously since we first met.

[1]For the interested reader, this not only encompasses academic-related matters, but also includes Friday Zoom Quarantinis during lockdown!