Probing the Physics and Geometry of Active Galactic Nuclei

JAKE ARTHUR JACK MITCHELL

Probing the Physics and Geometry of Active Galactic Nuclei

Jake Arthur Jack Mitchell

Abstract: Super Massive Black Holes (SMBH) inhabit the centre of every major galaxy, some of which power Active Galactic Nuclei (AGN) via mass accretion. In this thesis I explore two of the main approaches taken to study AGN, firstly the use of a wide sample, and secondly detailed observations of a single object. Multi-wavelength emission from AGN can be broken down into physical components which dominate in different wavebands, namely the dusty torus (IR), the accretion disc (optical/UV), the soft X-ray excess (UV/soft X-ray) and hot corona (hard X-rays). I first outline our new sample of around 700 SDSS Optical/UV/X-ray selected AGN, the SOUX sample, where the SDSS emission lines (H α, H β and Mg II) give single epoch black hole mass estimates while the UV/X-rays determine the continuum. I bin and stack the sample in mass and luminosity to look for trends as a function of these two important parameters. I perform detailed continuum fitting on the SOUX sample including the disc, soft X-ray excess and hot corona, and demonstrate clear issues in all current accretion flow models. Simple discs cannot match the UV extent of the most massive AGN irrespective of black hole spin as the high spin fits used in previous studies did not include the strong gravitational redshift which results from general relativistic ray tracing from the inner disc to the observer. I conclude that either there is a systematic over-prediction in black hole masses, or, more likely, the accretion flow in AGN does not take the form of a standard disc but may be fully Comptonised or take the form of something altogether different. I then switch gear to use the variable accretion flow emission to constrain the physical size scale of the dusty torus and Paschen Broad Line Region (BLR) from near-IR reverberation mapping. For Mrk 509, I develop a technique of photometric re-scaling and am able to place loose constraints on the size scales of these regions despite the challenging dataset. This technique is repeated with the much more robust data-set on Mrk 817, observed as part of the wider STORM2 collaboration. Both of these show that the Paschen BLR is co-spatial with the Balmer BLR, which is important as the Paschen BLR is now resolvable with GRAVITY data in a few objects, so can be used to test single epoch mass estimates. I then summarise the presented work and discuss potential avenues of investigation to take this further.

Probing the Physics and Geometry of Active Galactic Nuclei

Jake Arthur Jack Mitchell

A thesis presented for the degree of
Doctor of Philosophy

Centre For Extragalactic Astronomy
Department of Physics
Durham University
United Kingdom

December 2022

Dedicated to

Mum and Dad

Also to

Grandad Sam

Contents

List of Figures

List of Tables

Declaration

The work in this thesis is based on research carried out in the Centre For Extragalactic Astronomy, Department of Physics, Durham University, United Kingdom. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

The content presented in each named chapter is taken from the specified paper which has been submitted to MNRAS:

- Chapter 3: *"The SOUX AGN Sample: SDSS-XMM-Newton Optical, Ultraviolet and X-ray selected active galactic nuclei spanning a wide range of parameter space – Sample definition"*
- Chapter 4: *"The SOUX AGN sample: Optical/UV/X-ray SEDs and the nature of the disc"*
- Chapter 5: *"The first spectroscopic IR reverberation programme on Mrk 509"*

I am second author on the paper *"The SOUX AGN Sample: SDSS-XMM-Newton Optical, Ultraviolet and X-ray selected active galactic nuclei spanning a wide range of parameter space – Sample definition"*, which was primarily authored by Dr. Daniel Kynoch and is presented in Chapter 3. I was primarily responsible for the radio matching and analysis and performed some initial spectral fitting for black hole mass determination along with visual inspection and quality control of the optical spectra and *XMM-Newton* data. Dr. Daniel Kynoch was primarily responsible for the sample definition, black hole mass estimations and the α_{ox} analysis which is presented.

The work presented in Chapter 4, *"The SOUX AGN sample: Optical/UV/X-ray SEDs and the nature of the disc"*, was completed in close collaboration with Professor Chris Done.

In Chapter 5 the spectral decomposition and temperature estimation were carried out and detailed by Dr. Hermine Landt, who also summarised the LCOGT and GROND observations. The optical and near-IR photometry were provided by J. V. Hernández Santisteban. In Chapter 6 the optical and near-IR photometry were provided by John Montano as part of the STORM2 collaboration, permission to present these data here was given by Dr. Gerard Kriss. I carried out the vast majority of spectral observations presented in Chapters 5 and 6 through the remote observing mode of the NASA IRTF facility. I was the PI for two semesters of Mrk 817 observations and also PI of a pilot study on near-IR spectro-polarimetry of AGN using the 8 m Subaru telescope. In addition to this I attended an in person observational summer school in Bulgaria based at the 2 m Rohzen facility.

Acknowledgements

Throughout my life I have been fortunate enough to be surrounded by a great many people whose support and guidance has been invaluable and without whom this work would not have been possible. To all of these people, I owe an immense gratitude.

Firstly, I would like to thank Martin Ward for his expert advice, relaxed supervision and interesting conversation over that last four years. I also owe a massive thanks to Chris Done for her invaluable support and guidance and for schooling me in the mystic ways of SED fitting with XSPEC. Thank you to Hermine Landt for introducing me to the IRTF and showing me the observational ropes during our virtual trips to Hawaii. I am very grateful to Daniel Kynoch for his patient advice and unending willingness to lend me his time and expertise. I also extend a massive thanks to Scott Hagan for lending me his amazing computing skills and for providing extremely valuable advice and support in the latter stages of my PhD. I would also like to thank Dave Alexander, Paula Chadwick and Tim Roberts for their support throughout. I extend my gratitude to Alastair Edge, Mark Swinbank and John Lucey as I would not have even applied for this PhD without their astronomical enthusiasm and caring guidance during my undergraduate degree.

I have had the privilege of meeting many great friends throughout my time as a PhD student and without them I certainly would not have been able to complete this thesis. For all of the long afternoon walks and early morning breakfasts that have kept me sane I thank Amy Etherington, this would have been a lot more difficult without you. To Vicky, Jack, Aiden, Ellen, Joaqiun, Max, Ra'ad, Ryota, Martina, Emmy, Duncan, Victor, Alice, Cam, Ed and James thanks for brightening the office each day.

Outside of the department, I am blessed with an incredible support network and have countless people to thank for their love and support. I am very grateful to Will and Paul, for being the most understanding and supportive housemates possible and for making great memories over the last three years. Jamie, you have helped me through this every step of the way, without your consistent voice of calm, reason and unconditional support I wouldn't have come

close to completing this. Freddie, Sam, Elliott and Sion thank you for all of the good times throughout the Durham years and beyond and of course for your continued support. Will, Linton, Fraser and Harry, your voices have now formed part of my subconscious. I can't thank you enough for the marathon conversations and for filling my days with unending laughter. A special thank you to Reeves, for all of the laughs and of course the unwavering support.

To my family, I can't thank you enough for all of the love and support that you show me every day, I love you all. Karen and Simon, you have been like a second set of parents to me, thank you for the countless times that you have helped me in life. Emma, Laura, Ashley and Beth I am so grateful to have you, thank you for everything. To Charlie, Edie and Lane, thank you for brightening my life every single day. To my Nan and Grandad Sam, and Nanna and Grandad Halifax, you have always been a source of love and guidance, without you none of this would have been possible. To Ben thank you for always being there, I wouldn't be the person I am today without you. To my parents, from day one you have given me unconditional love and support and have helped me every step of the way. Thank you for your steadfast belief in me without it I could never have come this far. You are always there when I need you and for that I am grateful beyond words.

Last but by no means least, Natalie, you are my best friend and a constant source of light and joy in my life. Without you this journey would have been immeasurably more difficult, your love and support mean the world to me, thank you and I love you.

Chapter 1

Introduction

1.1 Super Massive Black Holes

Super Massive Black Holes (SMBH) exist at the centre of every major galaxy. SMBHs are the most massive compact objects in the known universe weighing hundreds of thousands to tens of billions of solar masses. Whilst they themselves do not radiate, the effect that they have on their local environment and the wider universe is far from invisible.

In our own galaxy the radio source Sagittarius A* marks the position of our SMBH. Until very recently, the most compelling evidence that this was a SMBH came from observations of central bulge stars rapidly orbiting an unseen mass (Ghez et al., 2008). However, due to the advancement of observational techniques we now have direct radio images of two local SMBHs, first M87 and secondly Sagittarius A* (Event Horizon Telescope Collaboration et al. 2019, 2022: see Figure 1.1).

Radio astronomy has not only provided us with our first direct glimpse of SMBHs in the local universe, but was key in the discovery and understanding of much more distant and luminous systems. Radio surveys in the 1950's such as the third Cambridge radio survey (Edge et al., 1959), were responsible for the discovery of quasars. In 1964, observations of the quasi-stellar radio source (quasar) 3C 273 inspired the idea that mass accretion onto a central SMBH could be an extremely efficient mechanism for generating the power exhibited by this source (Salpeter, 1964; Zel'dovich, 1965). Quasars are now known to be in galaxies where the central SMBH is accreting matter, in other words, galaxies with active nuclei.

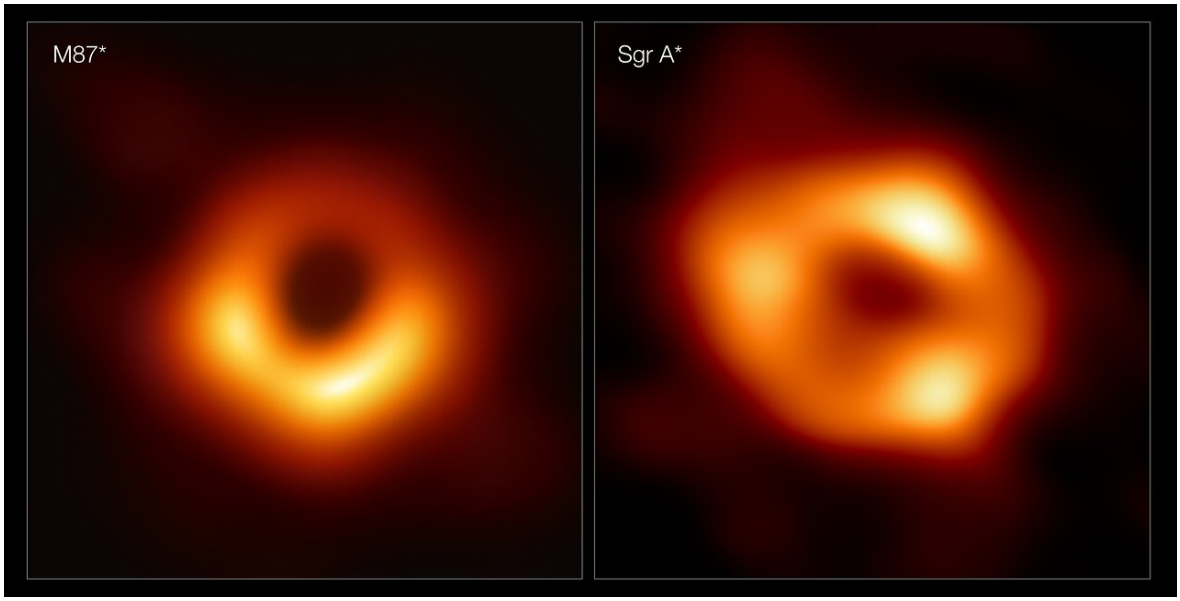

Figure 1.1: The shadows of the SMBHs at the centre of the galaxy M87 (left), and the Milky Way (right), against radio emission from a ring of hot in-falling matter. Imaged by the Event Horizon Telescope, these are the first direct images of Black Holes ever taken (Event Horizon Telescope Collaboration et al., 2019, 2022).

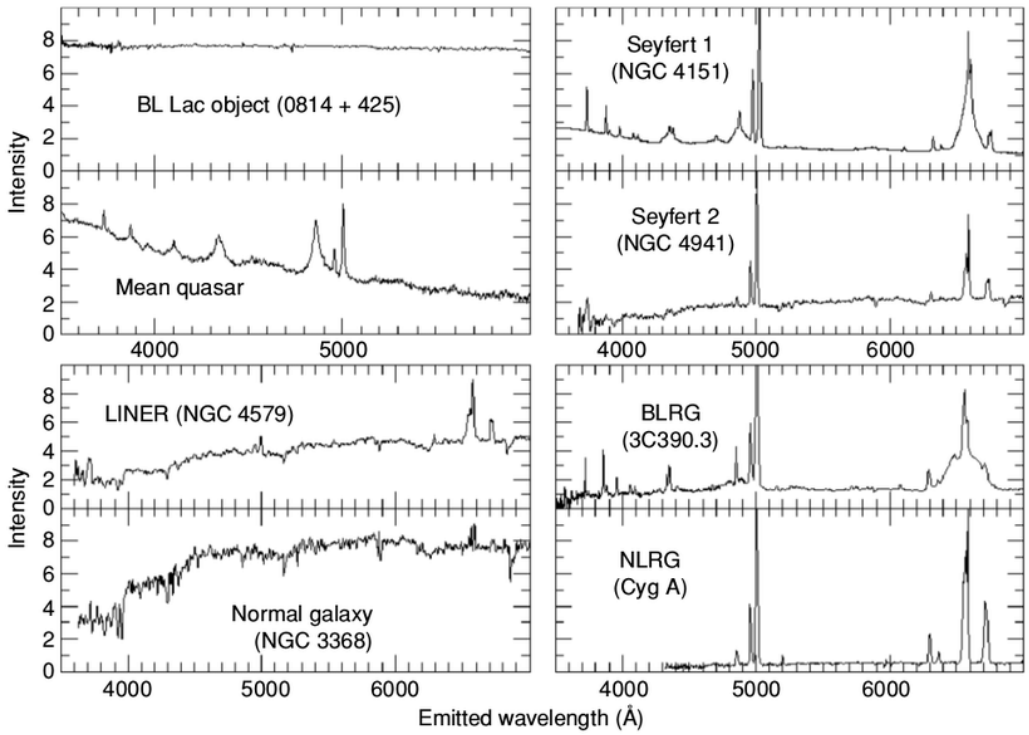

Figure 1.2: A variety of typical AGN optical spectra of differing classification from Andika (2016).

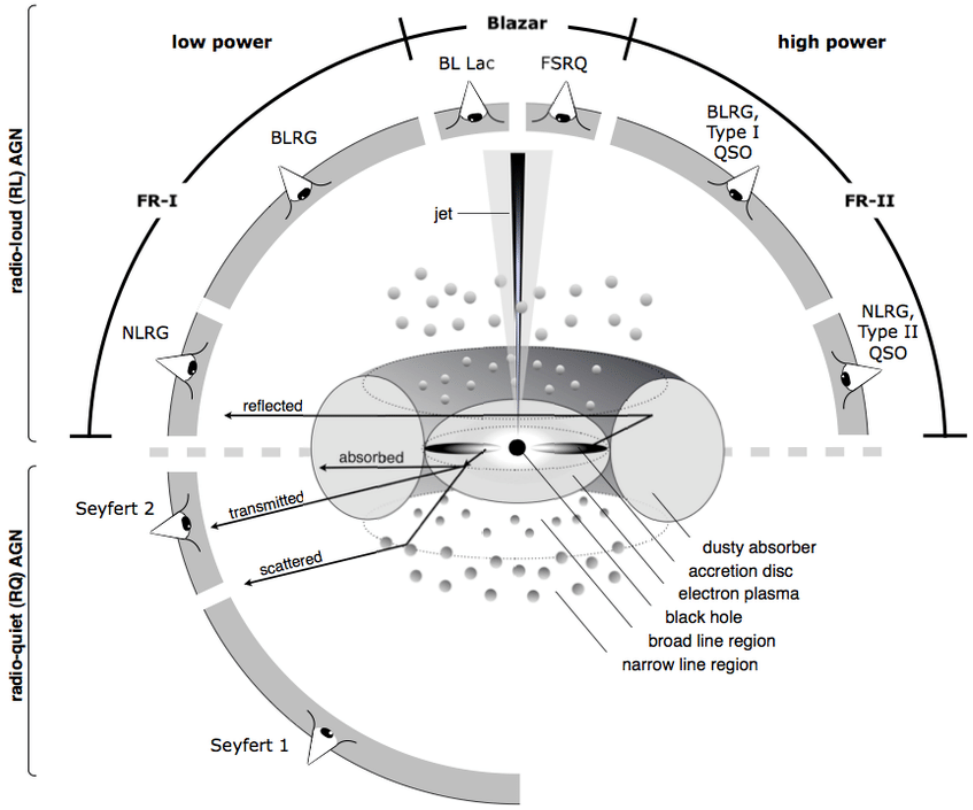

Figure 1.3: A schematic of the unified model of AGN from Beckmann & Shrader (2012).

1.2 Active Galactic Nuclei

In some galaxies, the extreme gravitational field of the central SMBH causes in the in-fall of matter across large distances. The immense gravitational potential energy liberated as this matter falls towards the event horizon is so intense that it can be visible across cosmic distances. These sources are known as Active Galactic Nuclei (AGN).

Most AGN are relatively distant meaning that their compact luminous central regions cannot be spatially resolved. Despite the inability to directly image their specific geometries we have gained incredible insights into the structures comprising AGN through many indirect and subtle analysis techniques. Throughout the decades of study on these sources and the resulting litany of literature, the most abundantly clear fact about AGN is their impressive diversity. This variety has resulted in a confusing zoo of AGN nomenclature. The most immediately obvious point of variance is the vast range of luminosity over which AGN can be observed. Coupled with this is the diversity in their optical spectra, examples of which are exhibited in Figure 1.2. Whilst there are many points of divergence between these examples of prototypical AGN of differing classification, the most notable is the heterogeneity in emission

line properties.

There are two distinct types of lines on display, narrow and broad emission lines originating from a Narrow Line Region (NLR) and a Broad Line Region (BLR) respectively. Seyfert 1 sources exhibit both narrow and broad lines, whereas in Seyfert 2 sources there are only strong narrow emission lines. The AGN unification scheme of Antonucci (1993) proposed a physical mechanism for this dichotomy, where the presence of an equatorial obscurer leads to differences in observed emission characteristics with changing viewing angle. Figure 1.3 shows a schematic of the unified model.

However, whilst orientation with respect to the observer is surely a key factor in the classification of AGN, we also observe systematic differences in the Spectral Energy Distribution (SED) shape with both mass and luminosity. This strongly implies that there is an intrinsic change in the SED shape as a function of these parameters. Therefore the origins of AGN diversity are more complex than is shown by Figure 1.3. The presence or lack of powerful radio jets is one intrinsic dichotomy that is highlighted in Figure 1.3. The result is that each half of the diagram represents an intrinsically different AGN paradigm.

Whilst I have briefly touched upon radio and optically based classifications, AGN emit across a wide range of the electromagnetic spectrum, with each of the physical components shown in Figure 1.3 dominating over a different bandpass. Therefore multi-wavelength observations are essential to build a truly holistic picture of the physics and geometry of AGN. With the advent of large telescopes and new observational techniques we are now able to probe the geometry of AGN across almost the entire electromagnetic spectrum and even spatially resolve the inner components of some sources (Gravity Collaboration et al., 2018, 2020). As these datasets widen, it will be possible to test the canonical paradigm of AGN physics more robustly than ever before.

1.3 Multi-Wavelength Observations of AGN

AGN radiate across a wide range of the electromagnetic spectrum with different physical components dominating the emission at different energies. It is therefore possible to model the physical components responsible for this emission and to construct a composite Spectral Energy Distribution (SED). Multi-wavelength SEDs are one of the most powerful diagnostic tools available to us when studying AGN. From an SED one can infer the mass accretion rate, bolometric luminosity and to an extent the physical geometry of a source (Jin et al.,

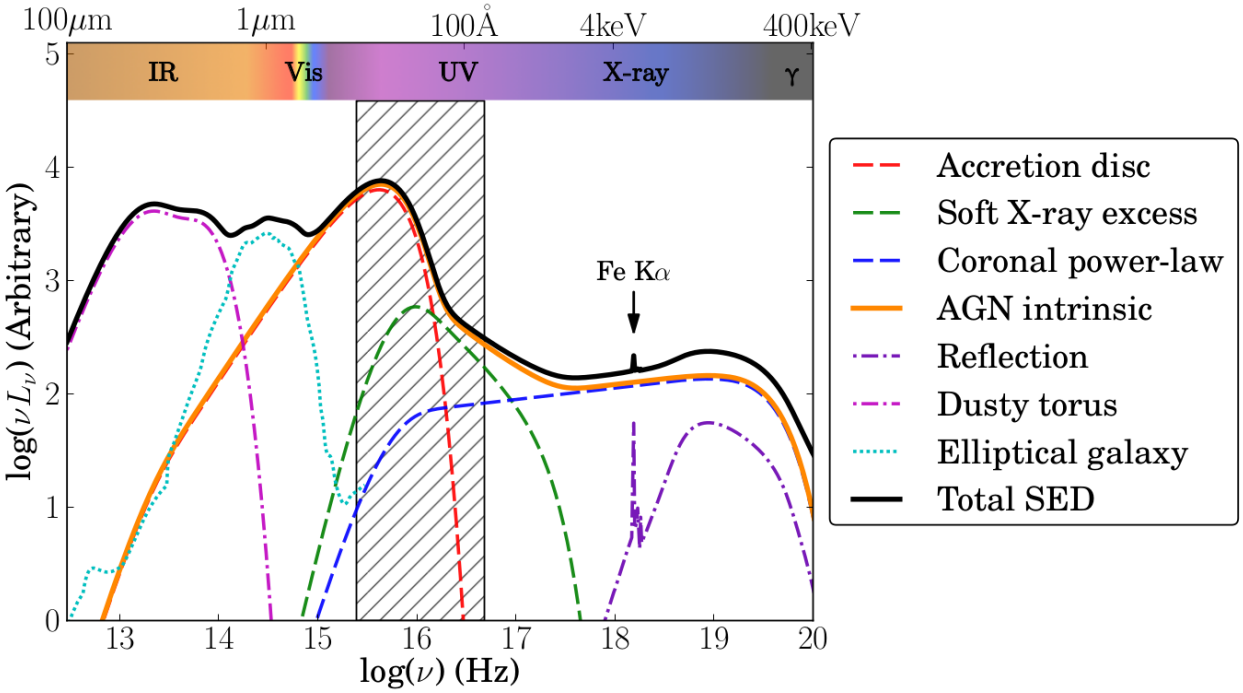

Figure 1.4: A simple schematic of a wide band-pass AGN SED. The total composite SED (black) is shown alongside the approximate shape and extent of the individual components thought to comprise the geometry of AGN. The host galaxy template is that of an elliptical galaxy, and the hatched region represents the spectral region which suffers from a high degree of absorption by the ISM. This schematic is taken from Collinson et al. (2017).

2012a,b; Kubota & Done, 2018). Figure 1.4 shows an example of a simple AGN schematic, broken down into different physical components, from (Collinson et al., 2017). The AGN SED generally peak in the UV but this is almost completely hidden by the Inter Stellar Medium in our own galaxy even in sources where our line of sight does not intersect the nuclear torus in the AGN host galaxy. This unfortunate fact makes us effectively blind to much of the accretion power in extragalactic sources.

The basic intrinsic AGN components included in Figure 1.4, are an inner X-ray emitting corona, a Comptonised Soft X-ray emitting region and a standard accretion disc. These are the key emission driving physical components that are included in many modern models (Kubota & Done, 2018, 2019). The dusty torus surrounds the accretion disc in the equatorial plane and radiates in the infrared, accompanied of course by the emission from stars in the host galaxy in the optical to Infrared. Predictably, the overall picture is much more complicated. Two key components missing from this schematic are the broad and narrow line regions, but in some sources we also observe phenomena such as disc winds and even colossal kpc-scale radio jets.

Multi-wavelength campaigns have catalogued a wide range of SED shapes in AGN using a

range of facilities from ground based Infrared to X-ray space telescopes. These SED shapes have been found to vary systematically with mass and mass accretion rate, indicating major intrinsic physical differences across the population (Jin et al., 2012a; Kubota & Done, 2018).

1.4 Accretion Power

If an AGN is thought of as an engine, then its inner regions are the combustion chamber, the source of almost all of its immense power. The driving mechanism here is accretion.

1.4.1 The Eddington Luminosity

First proposed in the context of mass accretion onto a star, the Eddington luminosity is the luminosity at which the radiation pressure and gravitational force on an in-falling particle reaches parity. Beyond this limit, radiation pressure would begin to blow material away. This principle is, widely applied to other massive accreting bodies such as neutron stars and black holes.

The Eddington Luminosity, for pure Hydrogen gas falling onto a BH of mass $M_{\rm BH}$, is defined as:

$$L_{\rm Edd} = \frac{4\pi G M_{\rm BH} m_{\rm p} c}{\sigma_{\rm T}} \simeq 1.26 \times 10^{38} \left(\frac{M_{\rm BH}}{M_\odot}\right) {\rm erg\ s^{-1}}, \tag{1.4.1}$$

where η is the efficiency, $\sigma_{\rm T}$ is the Thomson scattering cross section and $m_{\rm p}$ is the proton mass. The mass accretion rate of a source is often defined in terms of $L_{\rm Edd}$:

$$\dot{m} = \frac{\dot{M}}{\dot{M}_{\rm Edd}} = \frac{L}{L_{\rm Edd}}, \tag{1.4.2}$$

where $\dot{M}_{\rm Edd}$ is the accretion rate required to reach $L_{\rm Edd} = \eta \dot{M}_{\rm Edd} c^2$.

While the accretion flow is not spherical, the Eddington limit is an indicative limit of the point at which we would expect the accretion flow to change its structure.

1.4.2 Black Hole Binarys (BHB)

BHBs are systems in which a star orbits a stellar Mass black hole ($\sim 10 M_\odot$) such that the L_1 lagrange point lies close to or at the stars surface. The gas is stripped from the stars

surface and falls towards the BH forming an accretion disc (AD) orbiting the black hole. If the energy thermalises then it gives a temperature:

$$\sigma T(R)^4 = L(R) = \frac{3GM\dot{M}\, f(R)}{8\pi R^3}, \tag{1.4.3}$$

where R is the radius from the BH, and f(R) includes the stress free inner boundary condition.

In accretion disc models the flow extends down to the Inner Most Stable Circular Orbit, $R_{\rm ISCO}$, which is the general relativistic limit for a particle to maintain a stable orbit around a black hole. Within this limit a particle will quickly spiral into the BH so this effectively gives the inner edge of the accretion disc. $R_{\rm ISCO}$ is often expressed in terms of gravitational radii, $R_{\rm g}$, which is defined as follows:

$$R_{\rm g} = \frac{1}{2}R_{\rm S} = \frac{GM_{\rm BH}}{c^2}, \tag{1.4.4}$$

where $R_{\rm S}$ is the Schwarzchild radius, the radius of the event horizon of a BH. $R_{\rm ISCO}$ is not only dependent on the mass but is also reliant on one of the other fundamental properties held by a black hole, its spin (a_*).

For a zero spin, or Schwarzchild BH, $R_{\rm ISCO} = 6R_{\rm g}$, however this changes with BH spin. If a particle is orbiting in the same direction as the spin in a maximally spinning BH ($a_* = 0.998$), then $R_{\rm ISCO} \simeq 1.24R_{\rm g}$ whereas for a maximally counter rotating BH where $a_* = -1$, $R_{\rm ISCO} \simeq 9R_{\rm g}$. The upper limit limit on co-rotating spin is set at $a_* < 0.998$ due to counter-torque from the angular momentum of radiation from the disc being accreted by the BH.

There is a well documented transition from a high luminosity soft X-ray state towards a low luminosity hard X-ray state at around $\dot{m} = 0.02$. This transition occurs when the accretion flow switches from a geometrically thin, cool, optically thick disc similar to the standard disc models (Shakura & Sunyaev, 1973) to an optically thin, hot, geometrically thick flow such as the Advection Dominated accretion flow models (ADAFs) (Narayan & Yi, 1995). This stark spectral transition is exhibited in the spectra of Cyngus X-1 from Gierliński et al. (1999) shown in Figure 1.5.

1.4.3 AGN - Scaling up

BHBs give a zeroth order look at the accretion flow and how it might scale up to the SMBHs of AGN.

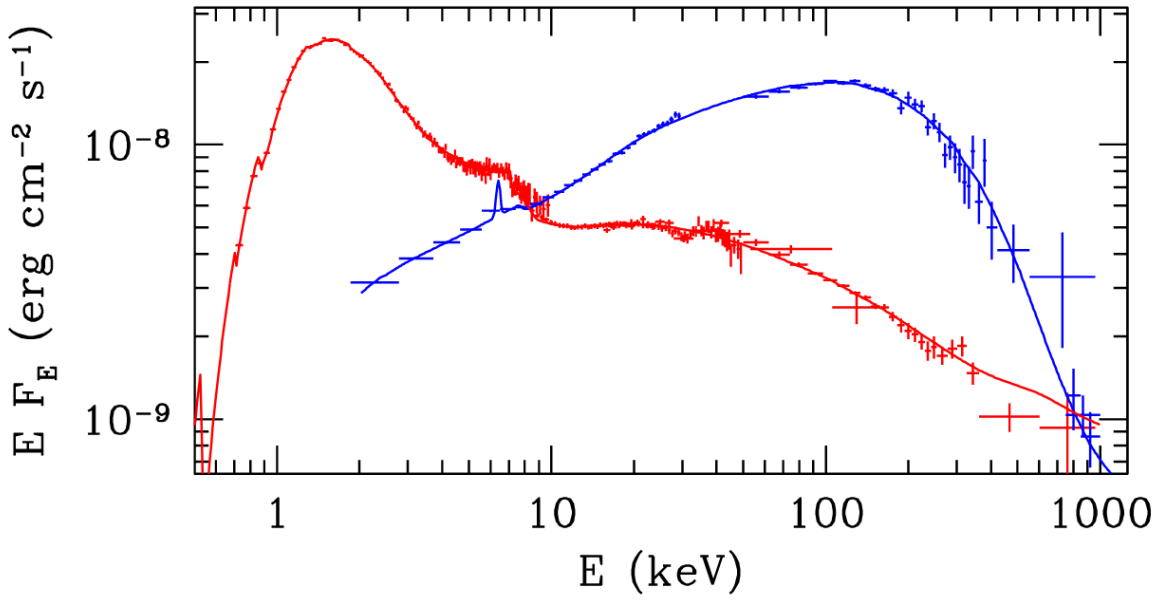

Figure 1.5: The SED of Cygnus X-1 transitioning from a soft (red) to a hard (blue) state.

The hot hard flow exhibited by low $L/L_{\rm Edd}$ BHBs should not be affected by the change in BH mass when scaling up to AGN. However the disc dominated emission would be expected to scale with the increase in mass as:

$$T_{\rm peak} \propto \left(\frac{\dot{m}}{M}\right)^{1/4}, \tag{1.4.5}$$

at the innermost stable circular orbit, giving a predicted peak luminosity in the UV for AGN rather than in the soft X-rays as seen in BHB. Unfortunately this results in the peak of emission being hidden by absorption in the ISM.

AGN are also known to undergo a state change around the same threshold of $\dot{m} = 0.02$ that marks the BHB spectral transition. AGN that cross this threshold, 'changing look' AGN display a dramatic change in the shape of their SED (see Figure 1.6). The optical/UV continuum almost completely collapses resulting in most of the power coming from the hard X-rays, giving an extremely hard SED shape. The physical interpretation of this transition is that the accretion flow switches from an optical/UV dominated thermally emitting AD to an ADAF in the form of a hot X-ray emitting corona.

The change in state is not only evident in the continuum emission, but also has a drastic effect on the BLR. Figure 1.7 shows the dramatic change in spectral shape and brightness in the quasar J1021+1645 as it crosses the 'changing look' threshold over the course of years. In the high-soft state (red line) J1021+1645 exhibits strong broad emission lines, however in the low-

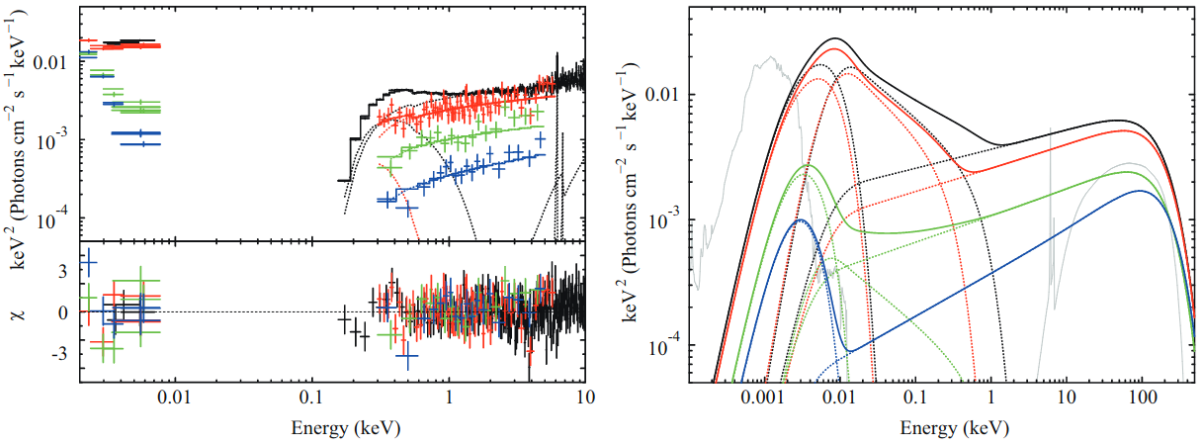

Figure 1.6: **Left:** Mrk 1018 optical/UV/X-ray spectrum in three different epochs spanning eight years. The data is fit with a model including the accretion disc, soft X-ray and hot X-ray corona with a torus and host galaxy component, with reddening and absorption. **Right:** The best fit models after reddening and absorption are removed. Grey shows model spectra of an S0-type host galaxy emission, and a torus reflection component (Noda & Done, 2018).

hard state (black line) the broad lines have almost totally disappeared leaving mainly strong narrow emission lines. These two spectra appear to embody the classic features associated with Type I (red line) and a Type 2 (black line) sources. It is clear that this change of state is not well fit by obscuration (blue and grey lines). Therefore this transition must reflect the intrinsic change of state between an AD dominated and ADAF regime. This type of process is not by any means exceptional for AGN, some large surveys suggest that essentially all AGN exhibit extreme changes of state given enough time (Morganson et al., 2014).

However AGN and BHBs diverge in their SED shape above the transition. AGN show clear signs of a more complex geometry than a simple AD. Unlike the BHB, AGN almost always display a significant amount of non-disc emission, in the form of a soft excess and hard X-ray tail. This has lead to the creation of AGN models consisting of three main components, an inner hot X-ray corona, then a warm Comptonised region of the disc, and an outer standard disc component. Understanding the specifics of the accretion flow structure by testing these models shall be a point of focus for this thesis.

1.5 AGN Variability

AGN are ubiquitously variable (Peterson et al., 2004). Not only do they all vary, but they all vary across their entire range of emission. Their amplitude is not constant and they appear to be aperiodic. Historically, variability was first documented in BL Lac and Optically Violent Variable (OVVs) objects which display the some of the most extreme examples of variability.

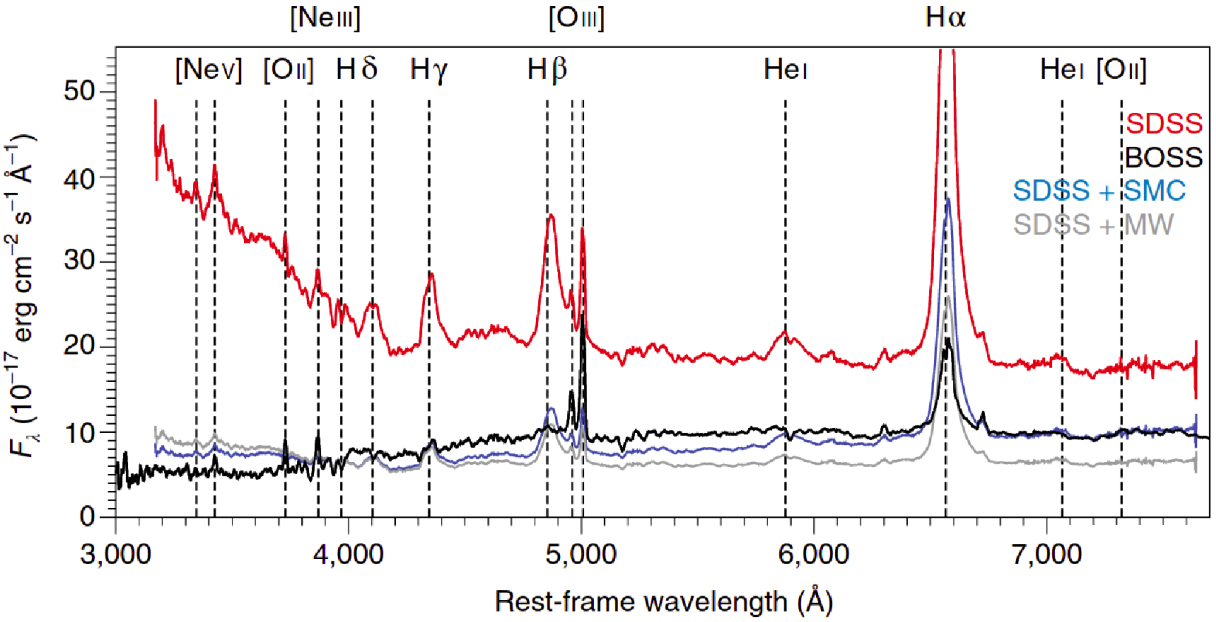

Figure 1.7: The dramatic change in spectral shape and brightness in the quasar J1021+1645. The red and black lines are the data. The blue and grey lines are failed attempts to model the collapse with variable extinction. These Figures are adapted from (Lawrence, 2018).

However these are not representative of the AGN population as a whole. We now know that this extreme variability comes from relativistic beaming of jet emission that is pointed directly towards the observer. In this thesis I will not discuss jet dominated sources in detail but rather focus on the more isotropically emitting central regions of AGN (see Figure 1.8).

Despite the prevalence of variability as an AGN phenomenon, its origin remains somewhat of a puzzle. For some time now, AGN have been in crisis (Lawrence, 2018), and resting at the heart of the issue is variability. The timescales on which AGN vary are orders of magnitude shorter than those expected. The mechanism which causes variability in BHB discs also cannot explain the variability seen in AGN emission. The concept of X-ray reprocessing, suggested by Clavel et al. (1992), solved some of the issues facing the AD model. In this model a central, rapidly varying X-ray source illuminates the disc, causing a variable heating effect. Any radius in the disc then has two sources of variability, the fast varying reprocessed X-ray emission, and the slowly varying thermal emission from viscous heating. However, there are limits to reprocessing as the UV variability must be intrinsic in the changing look AGN.

It is not only the continuum emission that varies in AGN, but also the broad emission lines. Figure 1.8 shows a basic schematic of the structural components responsible for the variable emission at different wavebands in the regime of X-ray reprocessing in NGC 5548,

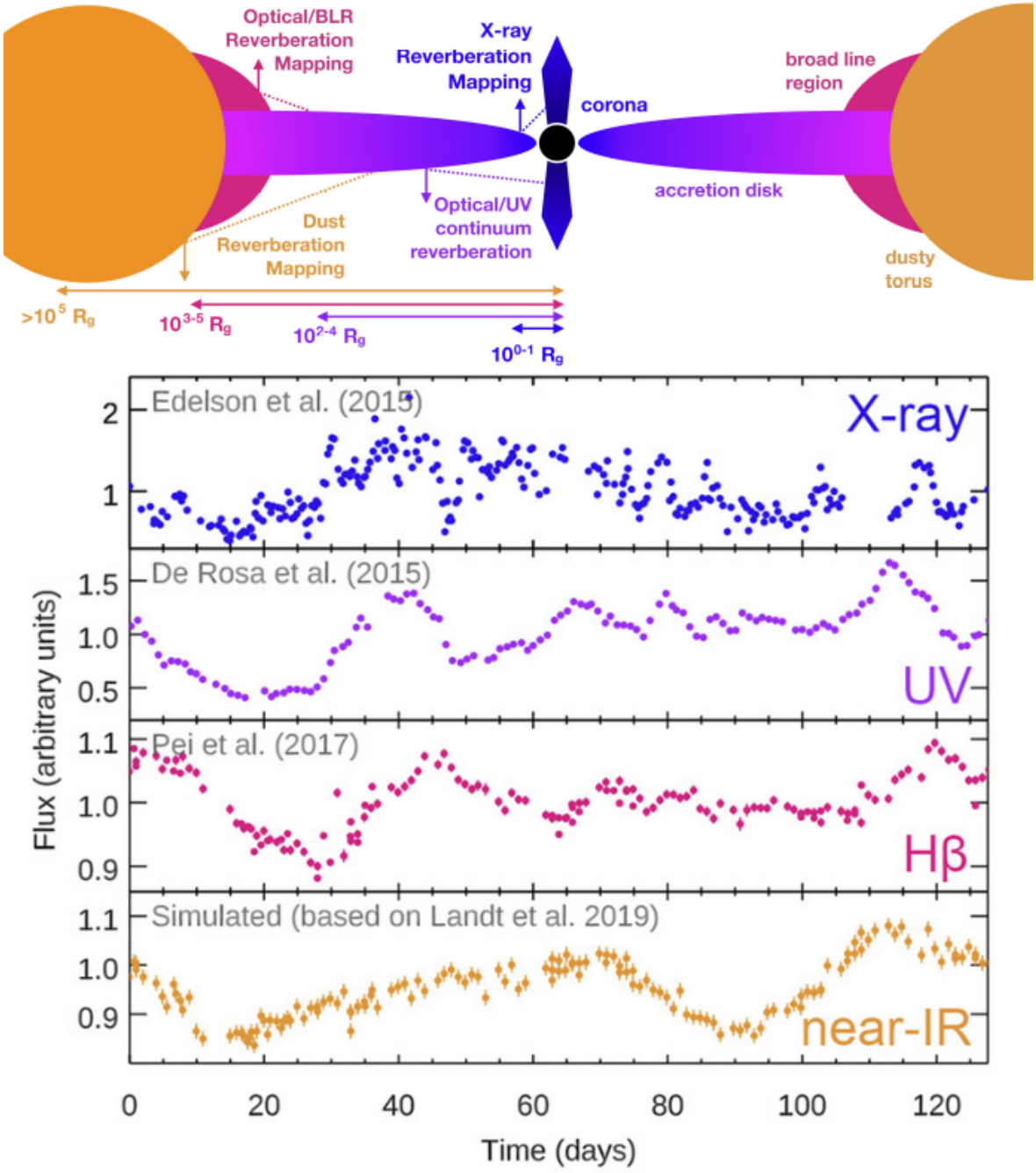

Figure 1.8: Lightcurves of NGC 5548, demonstrating variability in the X-ray corona, accretion disc, BLR and dusty torus, adapted from Cackett et al. (2021). The schematic highlights the physical component thought to be responsible for the emission in each lightcurve along with size scales as measured through Reverberation Mapping (RM).

the lightcurves for which are displayed beneath. The lightcurves clearly show similarities strongly implying some reprocessing mechanism.

Whilst the mystery over the nature of specific emission mechanisms persists, the process of variability itself remains an important tool. Variability is crucial in Reverberation Mapping techniques used to weigh black holes and measure the spatial extent of physical structures such as the dusty torus. This makes the understanding of the processes driving variability of utmost importance to the field of AGN study.

1.6 Reverberation Mapping

Reverberation mapping (RM) is a technique that relies on the measurement of a time delay between the variable signal in two separate wavebands. Using light travel time arguments a distance between the respective emission regions can be inferred.

1.6.1 The Broad Line Region

Understanding the nature of the BLR is of key importance to understanding AGN in general. The kinematics and geometry of this region are intrinsically linked to the accretion and outflow processes that power the emission and need to be better understood, and emission from the BLR is used to estimate black hole mass. As discussed in Section 1.4, black hole mass is extremely important in understanding the accretion process.

RM was primarily suggested to recover a delay between the optical-UV continuum emission originating from the AD and the variable signal of broad emission lines (Blandford & McKee, 1982). Therefore distances to the BLR can be inferred by assuming that this is the light travel time. These distances can be used in conjunction with measurements of the Full Width at Half Maximum (FWHM) of the broad emission lines and an assumption of a virialized geometry to provide an estimate of black hole mass. This can be calculated using the virial theorem:

$$M_{\rm BH} = f\frac{R_{\rm BLR}V^2}{G} = f\frac{c\tau\Delta V^2}{G}, \tag{1.6.1}$$

where $R_{\rm BLR}$ is the radius of the BLR, ΔV is the velocity of the gas, G is the gravitational constant, f is the virial factor which accounts for the geometry and inclination to the line of sight, c is the speed of light and τ is the measured time delay between the continuum source

and the emission line response. There are a few assumptions upon which the technique of RM is built (Peterson et al., 2004).

Firstly RM assumes that the most important timescale to consider is the light travel time, $\tau_{\rm LT} = R/c$, between the 'driving' and the 'responding' emission regions. This is a safe assumption as $\tau_{\rm LT}$ is usually of order tens to hundreds of days, whereas the timescales for emission line gas of typical BLR densities, $n_e \approx 10^{10}{\rm cm}^{-3}$, to re-establish photoionization equilibrium , $\tau_{\rm rec} = (n_e \alpha_{\rm B})^{-1}$, is only $\tau_{\rm rec} \approx 0.1$ hrs. This is negligible with respect to $\tau_{\rm LT}$. In addition, the dynamical timescale of BLR gas, $\tau_{\rm dyn} \approx R/\Delta V$, in a typical luminous quasar is $\sim (3-5)$ years, meaning that RM campaigns significantly shorter than this should avoid any smearing effects from changing BLR geometry (Peterson et al., 2004).

Secondly, the size scale of the continuum emitting region, the AD, is assumed to be small enough with respect to the size scale of the BLR that it can be considered to be a single central source.

Finally, it is assumed that there is a simple, relationship between the observed continuum and ionising continuum. Specifically, as the line variations are driven by the ionising continuum, it is assumed that the observed continuum must vary in phase with the ionising continuum.

AGN broad emission lines can be divided into two distinct populations, the Low Ionisation Lines (LIL) such as the Balmer series and High Ionisation Lines (HIL) such as He II and C IV. RM results have confirmed that the size scales of the emission regions for these sources differ, with the HIL region exhibiting lower delays and existing closer to the ionising source (Peterson & Wandel, 1999; Zajaček et al., 2021).

One theory as to the origin of the irradiated material constituting the BLR is a line driven wind (Murray et al., 1995). The launching of such a wind which would provide a source of HIL, and the formation of clouds within it has been substantiated by numerical studies and advanced computer simulations (Proga et al., 2000; Waters et al., 2022). The material composing the LIL region could emerge from a dust driven wind in the form of a Failed Radiatively Accelerated Dusty Outflow (FRADO) (Czerny & Hryniewicz, 2011). The formation of this wind is determined by the sublimation temperature of the dust particles which is of the order of thousands of Kelvin. Dust sublimation sets the inner edge of the torus but dust can also exist within this radius if it is embedded and shielded by the accretion disc itself. The structure can change from that of a failed wind at low mass accretion rates, to an escaping wind at high mass accretion rates (Czerny & Hryniewicz, 2011).

RM has widely been considered the 'gold standard' of black hole mass estimation since its

inception as a technique, however recently a few famous and luminous sources have had their BLR and torus spatially resolved (Gravity Collaboration et al., 2018, 2020). This approach is still extremely limited due to the necessary intensity of observation and the limitations on luminosity and cosmic distance. For the time being, RM remains the most reliable widely applicable mass estimation technique available to us. Not only does RM provide mass estimates, it also provides us with a radius-luminosity relationship than can be used to calculate single epoch virial mass estimates from single spectra, allowing bulk mass estimation where RM is not possible.

Single Epoch Virial Mass Estimates

Single epoch virial mass estimation is the cheapest mass estimation technique available, and allows for bulk mass estimation from spectral survey data.

As in RM, the width of the broad emission lines can be used to estimate the velocity of the material in the BLR. Here the FWHM of the broad emission lines is used as a proxy for V. $R_{\rm BLR}$ can be approximated using a radius-luminosity relationship calibrated through the technique of Reverberation Mapping, where $R_{\rm BLR} \propto L^{\alpha}$ and α is a constant. Therefore $M_{\rm BH}$ is calculated as follows:

$$M_{\rm BH} = K(L_{\lambda})^{\alpha}({\rm FWHM}^2), \tag{1.6.2}$$

where L_{λ} is a monochromatic luminosity, FWHM is measured from an emission line, and K is a constant. These mass estimates are referred to as single epoch virial mass estimates and are commonly considered to be accurate to a factor ~ 3. Much of the work presented in this thesis is contingent on single epoch virial mass estimation. We mainly adopt the relations of Mejía-Restrepo et al. (2016).

One alternative method of black hole mass estimation involves scaling the relationship between the X-ray variability timescale and black hole mass seen in Galactic black holes up to AGN (McHardy et al., 2006). As we gain a greater understanding of the origins of variability and the nature of the accretion flow in AGN, techniques such as this, which are not reliant on RM may be useful in testing the scaling relations derived from it.

1.6.2 Continuum

Reverberation mapping is not only useful in determining the size scales of the BLR for the purposes of $M_{\rm BH}$ estimation. The principles of this technique are equally applicable to aspects of the continuum emission in AGN. Through RM, constraints have been placed on the size scales of the AD and the inner edge of the dusty torus (McHardy et al., 2016; Hernández Santisteban et al., 2020; Koshida et al., 2014).

Optical/UV Continuum - The Accretion Disc

A growing number of intensive optical/UV/X-ray continuum RM campaigns of individual sources are being carried out to explore the central assumption that much of the UV/optical variability is from reprocessing. The basic concept is that the fast variable X-ray flux illuminates the disc resulting in fast variable optical/UV flux that is lagged by the light travel time. In the regime of a centrally illuminated thin accretion disc model, the time lag would be expected to vary as:

$$\tau = \frac{R}{c} \propto (M_{BH}^2 \dot{m})^{1/3} \lambda^{\beta}, \tag{1.6.3}$$

where, τ is the measured lag, λ is the emission wavelength and $\beta = 4/3$ for a standard AD (Cackett et al., 2007). This model takes into account both viscous heating and irradiation, assuming that the disc is irradiated from above at a height much larger than its own thickness but much smaller than the distance to each irradiated annulus. Not only would τ be expected to be larger in the optical than in the UV, but the amplitude variability in the optical should be less and the response more smoothed with respect to higher energy emission. This is due to the emission originating from a larger region at longer λ.

The expected relations seem well matched by data between the UV-optical, however the inclusion of X-rays complicates matters significantly. In a select few sources, the X-ray variability has even been observed to follow UV rather than lead it. Perhaps the most consistent result however is that the X-rays deviate from the expected $\tau \propto \lambda^{4/3}$ for a standard disc, implying the existence of an extended geometry or even a clumpy disc structure (McHardy et al., 2014; McHardy et al., 2018; Edelson et al., 2015, 2019). RM studies on continuum reverberation are proving extremely valuable in advancing our understanding of the structure of accretion discs, the origin of variability and the complex composition of AGN continuum emission and the role played by the BLR in it.

Infrared Continuum - The Dusty Torus

Not all AGN display broad emission lines in their spectra, some sources only exhibit strong narrow emission lines, these sources are known as Type 2 AGN. The lack of any obvious emission from a BLR in these sources combined with strong infrared (IR) continuum emission at $\lambda > 1\mu m$ is thought to be due to the presence of an obscurer. In fact some Type 2 AGN exhibit broad lines in their polarised spectra, lending further weight to the idea that obscuration is the reason for a lack of broad lines in these specific sources (Antonucci & Miller, 1985; Antonucci, 1993; Cohen et al., 1999; Lumsden et al., 2001; Tran, 2003). Historically this obscurer was modelled as a dusty toriodal structure known as the 'Dusty Torus', though the precise geometry, despite the nomenclature, remains unknown.

Emission from the torus is thought to be thermally reprocessed optical-UV emission from the AD. Therefore, much like the BLR, RM can be used to determine the size scales on which the torus exists, a technique implemented in this work. Currently there are ∼20 sources that have been reverberation mapped in the near-IR a technique most sensitive to the re-processed emission from the inner edge of the structure (Clavel et al., 1989; Sitko et al., 1993; Nelson, 1996; Oknyanskij & Horne, 2001; Glass, 2004; Minezaki et al., 2004; Suganuma et al., 2006; Koshida et al., 2014; Vazquez et al., 2015). Therefore any size scales estimated through this approach are inner radii and not indicative of the distance to the bulk of IR emission. This could explain why near-IR reverberation results are systematically smaller than the radii obtained from interferometric measurements. One expression of the potential geometries of these different regions is shown in Figure 1.9. In this configuration, the torus is not a simple toroidal structure, but has multiple components including an outflow resulting in a polar conical geometry (Hönig & Kishimoto, 2017; Hönig, 2019).

The actual geometry of the torus may become clearer with the advent of near and mid-IR interferometry which has successfully been used to spatially resolve the torus in a few objects, setting its spatial extent to ∼(0.1-6) pc (Tristram et al., 2009; Pott et al., 2010; Kishimoto et al., 2011; Burtscher et al., 2013; Kishimoto et al., 2013). These findings are roughly consistent with a power-law size-luminosity relation $R_{\rm dust} = L^{0.5}$ (Kishimoto et al., 2009, 2011). This is a technique that can only be applied sparingly due to the intensive observational requirements, and the limitations on the luminosity and distance of a potential source.

Not only is progress being made in determining the geometry of the torus, but also its astro-chemistry. Building on, and confirming the findings of Landt et al. (2019), this work presents

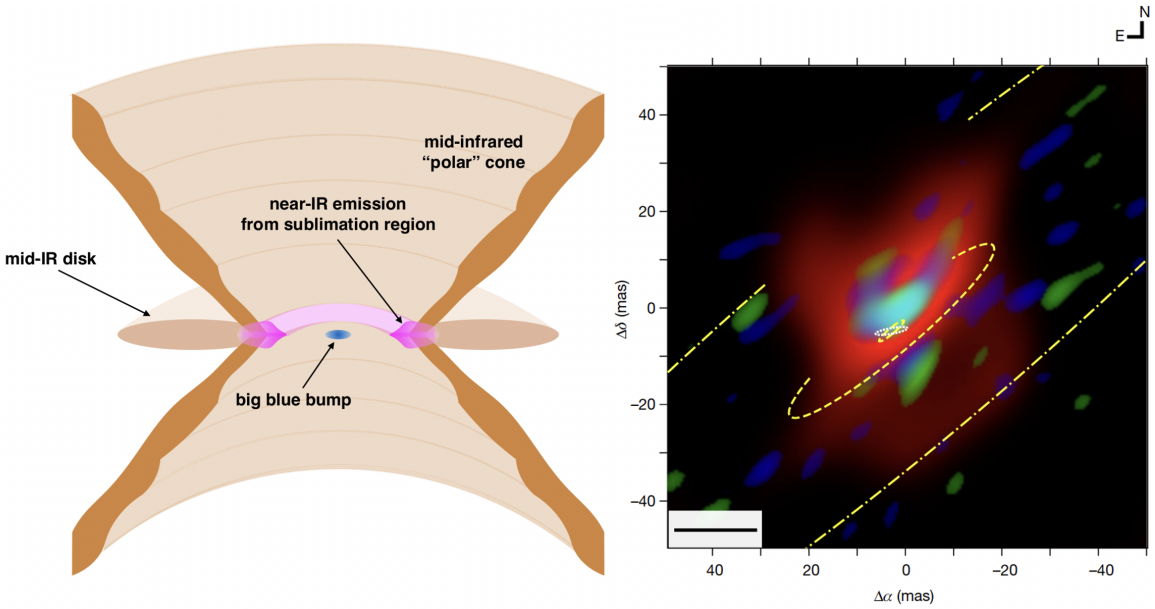

Figure 1.9: **Left panel:** A schematic showing a possible geometry of the pc sized IR emission regions of an AGN taken from Hönig (2019). The components are as follows; a falsely enlarged AD (blue), representing the 'driving' optical-UV emission, the sublimation region (pink) which is the inner edge of the dusty structure dominating the (3-5)μm emission, a hollow dusty cone towards the polar regions (dark brown) and a geometrically thin disc in the equatorial plane (light brown) both of which emit in the infrared. **Right panel:** IRBis reconstructed images of NCG 1068 from (Rosas et al., 2022). Interferometric colour composite image of the 3.4μm (blue), 4.6μm (green), 12μm (red) emission. The innermost small elipses possible positions of the masked sublimation limit around the SMBH. The outermost dot-dashed lines ellipse shows the maximum measured disc extent (10pc).

results indicating that the torus is predominantly composed of large carbonaceous dust grains. This conclusion is achieved through means of spectroscopic near-IR RM, which is much the same in its principles as photometric near-IR RM, but with the added benefit that temperature can be measured simultaneously with time delay. To measure the temperature a spectral decomposition can be performed, allowing the hot dust component to be fit with blackbody profiles representing differing astrochemical compositions. From this the luminosity weighted radius is calculated as:

$$\frac{L_{\rm UV}}{4\pi R^2_{\rm d,lum}} = 4\sigma T^4 \langle Q^{\rm em} \rangle \tag{1.6.4}$$

where σ is the Stefan-Boltzmann constant and $\langle Q^{\rm em} \rangle$ is the Planck-averaged value of $Q_\lambda(a)$. These radii can then be compared to the response weighted radii to determine the astrochemistry (Landt et al., 2019). These results are presented in Chapter 5.

1.7 AGN and the wider universe

The work presented in this thesis is a focused attempt to better understand certain aspects of the physics and geometry of AGN. Given the size and complexity of the wider known universe, AGN may seem like a small component, however they are far from insignificant with respect to the bigger picture. Unravelling the mysteries of these galactic powerhouses may shed some light on long standing cosmological mysteries including galaxy and large scale structure formation.

From where, and through what processes did SMBHs originate? This is perhaps the most perplexing mystery surrounding AGN and a question with profound impacts on our perception of galaxy formation itself. The most distant quasars, which have SMBHs of millions to billions of solar masses have been detected at $z \sim 7$ despite the fact that universe was only $\sim$1 Gyr old at this epoch (Mortlock et al., 2011; Wang et al., 2021). This is not enough time for a stellar collapse BH seed of $\sim 100 M_{\odot}$ to grow, even at $\dot{m}$=1. This implies that SMBHs must form from much larger BH seeds, implying the existence of intermediate mass BHs. But there is no evidence for this population. Whatever the mechanism of formation, the intrinsic properties of BHs retain information about their past. $M_{\rm BH}$ is a record of the amount of mass accreted by a BH over its lifetime and spin can provide an insight into the mode of accretion and the merger history of a system.

The impact of AGN activity on the host galaxy is not well understood. Despite this, there are several relationships that indicate a tightly connected evolutionary process between an AGN and its host. Perhaps the most notable is the tightly correlated and well documented $M_{\rm BH} - \sigma$ relationship (Gültekin et al., 2009) (see Figure 1.10).

One way in which SMBHs are thought to impact the evolution of massive galaxies, is through the quenching of star formation via AGN feeback McNamara & Nulsen (2007); Fabian (2012); Gaspari et al. (2020). AGN feedback refers to the ejection of matter and energy from an AGN core into its host galaxy and the IGM (Eckert et al., 2021). Figure 1.11 shows two examples of such feedback processes. Firstly via jets, e.g., the cluster MS 0735+7421, and secondly via winds, e.g., the P-Cygni profile caused by an Ultra Fast Outflow (UFO) in the quasar PDS 456. Hydrodynamical simulations of galaxy formation including and excluding an AGN feedback mechanism have been carried out. It has been concluded from these studies and others that AGN feedback is required to produce the halo baryon fraction and the shape of the galaxy luminosity function (McCarthy et al., 2010; Oppenheimer et al., 2021). It is also important to understand the fuelling processes which power SMBH accretion to build a full

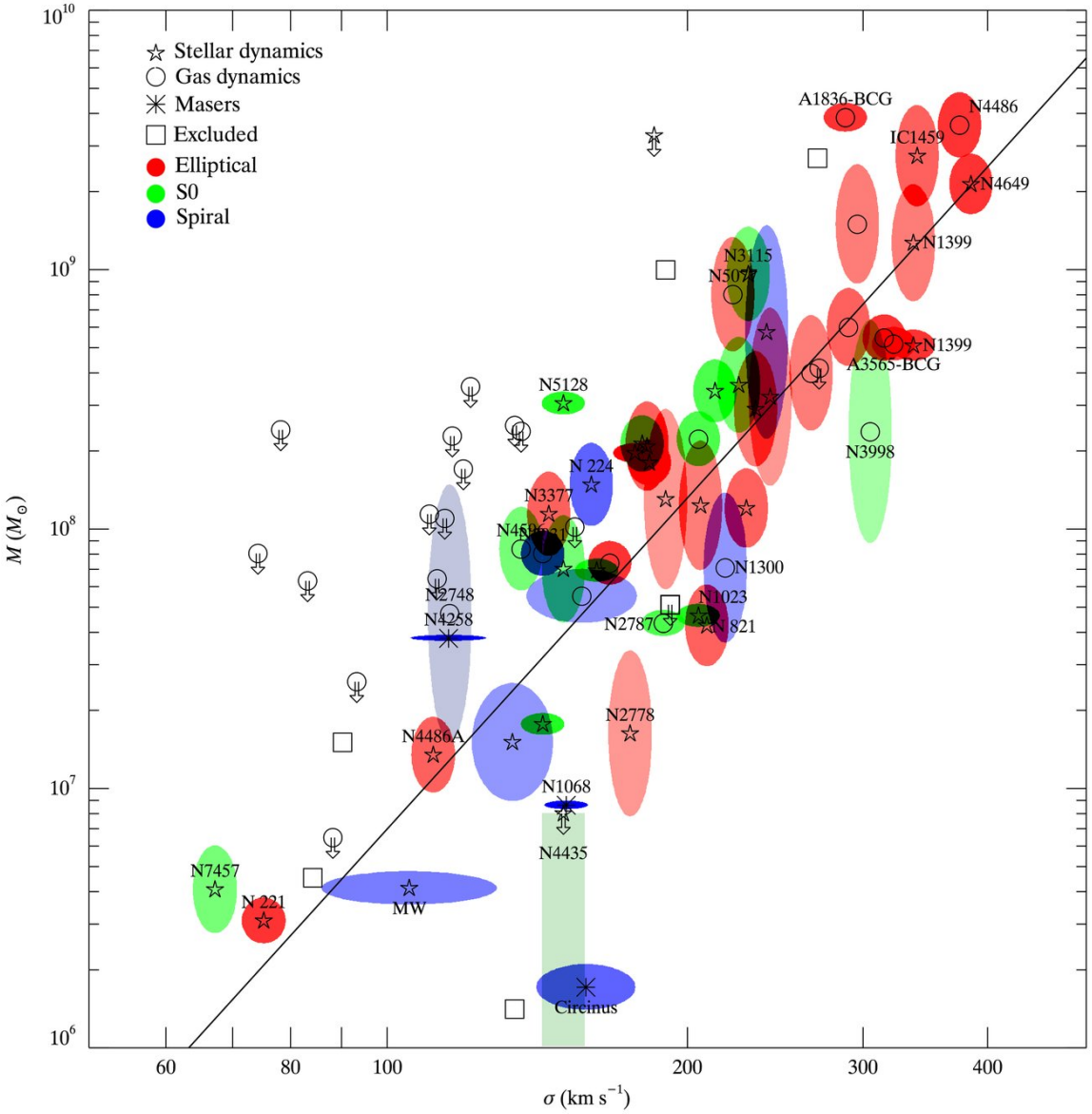

Figure 1.10: $M-\sigma$ relation for galaxies with dynamical measurements. the symbols indicate the method of $M_{\rm BH}$ estimation. The color of the error ellipse indicates the Hubble type of the host galaxy.The black line represents the best fit to the full sample Gültekin et al. (2009).

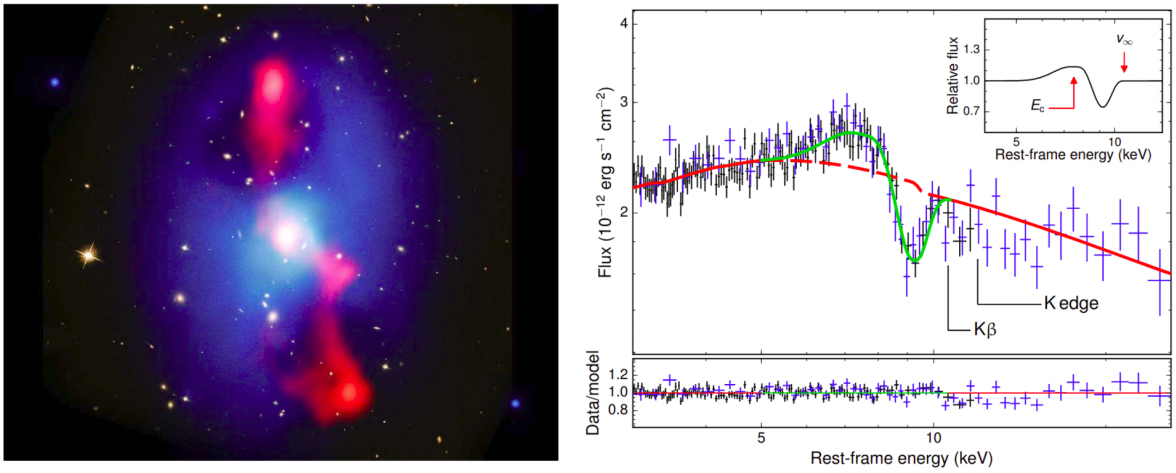

Figure 1.11: **Right:** Chandra X-ray image (blue) and Very Large Array (VLA) 300 MHz radio image (red) with the Hubble space telescope optical image of galaxy cluster MS 0735+7421. The image size is ∼800 kpc by 800 kpc (Gitti et al., 2012). **Left:** P-Cygni profile from a spherically symmetric outflow showing evidence for an Ultra Fast Outflow (UFO) in the quasar PDS 456 (Nardini et al., 2015).

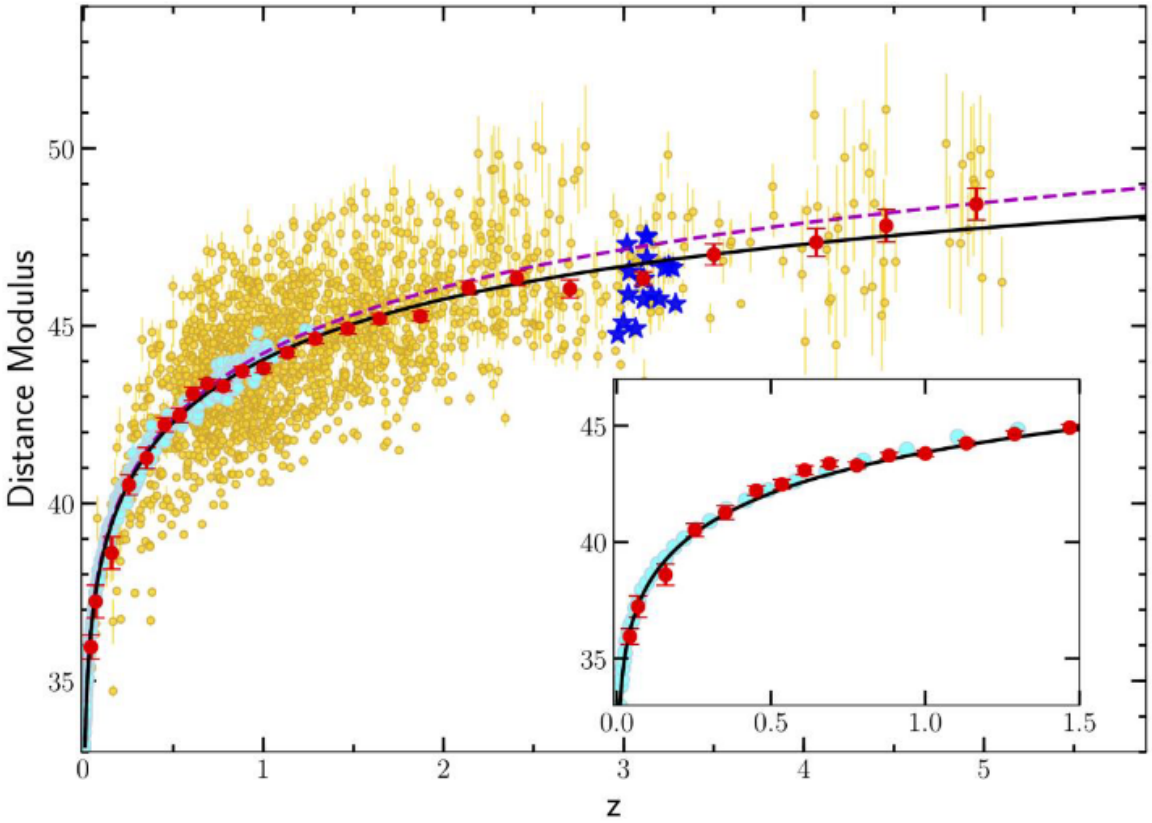

Figure 1.12: A Hubble diagram of both the supernovae from the JLA survey2 (cyan points) and quasars (yellow points). The mean of the quasar population is shown in red for narrow redshift bins as a visual aid. The dashed magenta line is the flat ΛCDM model with $\Omega_M = 0.31 \pm 0.05$ fitted to data of z<1.4 and extrapolated. The best fit to the third order expansion of log(1+z) is shown in black. The new *XMM-Newton* sample of z>3 quasars is shown in blue. The inset shows a zoom of the quasar and supernovae averages binned by redshift Risaliti & Lusso (2018).

picture of the feeback process, though this requires a better understanding of the molecular gas which exists close to the AGN. Rose et al. (2019) have recently detected inflows of such gas in some of the brightest cluster galaxies, providing evidence for one potential fuelling mechanism of AGN accretion and subsequent feedback. From studies such as these it seems clear that the energetic processes linked to AGN accretion are a key component in galaxy formation. This means that our understanding of AGN and the accretion physics driving feedback processes is incredibly important to building a full understanding of galaxy formation. This is especially true since the largest scale cosmological simulations, such as EAGLE (Schaye et al., 2015), are not yet of a sufficient resolution to include proper hydrodynamical models of each AGN and therefore have these feedback processes in-built at the sub-grid level.

Quasars may also be used as probes of cosmic distances and therefore be used to test the cosmological ΛCDM model over a range of redshift otherwise inaccessible (Risaliti & Lusso, 2015, 2018). Using the non-linear relationship between the UV and X-ray luminosities of quasars, Risaliti & Lusso (2015) are able to build a quasar Hubble diagram that matches well to the supernovae analog in the usual z=0.1-1.4 redshift range but extends out to z∼6. Using this, they test the assumptions of ΛCDM over an unprecedented timescale, and are able to find deviations from the model at high redhishift, perhaps indicating an evolution in the state of dark energy, increasing with time (Risaliti & Lusso, 2018) (see Figure 1.12). This work is still based on relatively small sample sizes, however with the ever increasing number of observed high redshift sources, quasars seem set to provide the means to refining accepted cosmological paradigms.

1.8 This work

This is a thesis of two halves, both of which are tied together by a common goal; to further the collective understanding of the physics and geometry of AGN.

Broadly speaking, there are two main approaches taken when studying AGN, the detailed studies of individual objects and wider studies of samples. Both sides of this coin are essential in achieving a full understanding of these objects. The work presented here is a mixture observational and theoretical approaches and aims to highlight areas in which the accepted knowledge in each of these paradigms do not align, in hopes of pushing the field towards a more accurate and complete understanding of AGN.

Here I give a brief outline of the work presented in this thesis:

- **Chapter 2:** A short summary of the astronomical facilities and surveys most relevant to the work presented in the subsequent chapters of this thesis.

- **Chapter 3:** We define a sample of SDSS Optical/UV/X-ray selected AGN, the SOUX sample. We analyse the samples mass distribution, α_{ox} characteristics, and radio properties of the sample.

- **Chapter 4:** Using the SOUX sample, as defined in Chapter 3 we perform detailed SED fitting on a 2-D grid of black hole mass and luminosity testing the assumptions made in the new AGN SED model QSOSED (Kubota & Done, 2018). We expand our findings to the wider SDSS population and discuss major divergences between observational data and current AGN models.

- **Chapter 5:** We switch gear to individual objects. We present the results from a near-IR spectroscopic RM campaign on Mrk 509. We develop method of photometric correction in a source for which the narrow lines show signs of extension beyond the slit. We perform a spectral decomposition and measure simultaneous luminosity weighted and response weighted radii, and place constraints on the response in four spectral bandpasses and Pa ϵ and Pa β.

- **Chapter 6:** We extend the work on Mrk 509 to Mrk 817 and apply the same scaling approach to search for a reverberated signal in four spectral bandpasses as well as for Pa β. The data presented here is part of the large multi-wavelength STORM2 campaign, allowing the comparison of our Pa β result with contemporaneous H β reverberation results.

- **Chapter 7:** I briefly summarise the conclusions of and insights gained from each of the projects undertaken as part of this thesis and give a short description of possible directions of future works.

Chapter 2

Facilities and Surveys

2.1 Telescopes

2.1.1 Infrared Telescope Facility (IRTF)

The NASA Infrared Telescope Facility (IRTF) is a 3.2 m telescope optimized for infrared (IR) observations located on Maunakea on the Big Island of Hawai'i. The IRTF is a versatile facility with a range of instruments allowing for photometry and medium-high mid and near IR spectroscopy.

I have carried out two separate observing campaigns with the IRTF, firstly observing Mrk 509 and then Mrk 817 as part of the wider STORM2 campaign. In total there were 57 nights of successful observations, the data from which is presented in Chapters 5 and 6. All of these observations were carried out with the SpeX instrument, a medium-resolution (0.7-3.5) μm spectrograph.

2.1.2 Subaru

Subaru is an 8.2 m telescope located on Manaukea on the Big Island of Hawai'i, owned and operated by the National Astronomical Observatory of Japan (NAOJ). There are a number of instruments that can be fitted to Subaru that allow observations from the optical to the IR, in both photometric and spectroscopic modes.

I have been awarded time for two separate observation runs in order to observe five AGN using the IRCS instrument in near-IR spectro-polarimetry mode.

Figure 2.1: Image of the IRTF

Figure 2.2: Image of Subaru

2.2 Surveys

Much of the analysis presented in this work has been reliant on the usage of published surveys covering different wavebands. Here I present a short description of each of the surveys relevant to this thesis.

2.2.1 Radio

We utilise two separate radio surveys, both sets of observations taken with the National Radio Astronomy Observatory (NRAO) Very Large Array (VLA). The VLA is one of the worlds most sophisticated radio observatories. Located on the Plains of San Agustin, New Mexico it consists of 27, 25 m diameter, antenna in a Y-shaped configuration. These signals from each anntena can be combined in order to provide the resolution of an antenna 36 km wide and the equivalent sensitivity of a dish 130 m in diameter.

NVSS

The NRAO VLA Sky Survey (NVSS) (Condon et al., 1998) mapped the entire 1.4 GHz radio sky above $\delta = -40^{\circ}$ between 1993 and 1997 with a beam width of 45". The catalogue contains almost 2 million discrete sources with a 5σ detection limit of ~ 2.25 mJy.

FIRST

The Very Large Array (VLA), Faint Images of the Radio Sky at Twenty-Centimetres (FIRST) survey (Becker et al., 1995) is a 1.4 GHz covering 10000 square degrees of sky in the North and South galactic caps.

2.2.2 Optical/UV

SDSS

The Sloan Digital Sky Survey (SDSS) is an optical photometric and spectroscopic survey mapping one quarter of the northern sky (York et al., 2000). The survey uses a 2.5 m telescope at Apache Point Observatory in New Mexico, United States. The survey took photometry in five broad optical bands to a depth of $g' \sim 23$ mag. There was a subsequent spectroscopic survey imaging $\sim 10^6$ brightest galaxies and $\sim 10^5$ brightest quasars in its first

iteration. The most recent version, the sixteenth release of the SDSS quasar catalogue (Lyke et al., 2020), contains 750 414 quasars. We primarily make use of the fourteenth release of the SDSS quasar catalogue, which contains 526 356 quasars (Pâris et al., 2018). We use this catalogue due to the in depth analysis carried out in Rakshit et al. (2020) the results from which are utilised in Chapter 3.

XMM-SUSS

XMM–Newton Serendipitous Ultraviolet Source Survey (XMM-SUSS) is a catalogue of sources detected serendipitously by the Optical Monitor (OM) on board XMM-Newton (Page et al., 2012). The OM a small 30 cm optical/UV telescope imagine in the U, B, $UVW2$, $UVM2$ and $UVW1$ bands (Mason et al., 2001). This work uses the XMM-SUSS4.1 catalogue which spans the obseration period from 2000 and July 2017 and contains 8 176 156 entries corresponding to 5 503 765 sources, 1 035 453 of which have multiple entries.

2.2.3 X-ray

XMM-Newton

The X-ray Multi-Mirror Mission (XMM-Newton: Jansen et al. 2001) is a European Space Agency satellite launched aboard an Ariane 504 on 10^{th} December 1999 . Almost all X-rays are blocked by the Earth's atmosphere meaning that ground based X-ray observatories are not viable and that space telescopes are essential for viewing this high energy emission. The orbit of XMM-Newton is highly eccentric meaning that the satellite is located beyond the Van Allen radiation belts for much of its ∼48 hour orbital allowing for un-interrupted long exposures.

XMM-Newton is host to three European Photon Imaging Camera (EPIC: Strüder et al. 2001) detectors and two Reflection Grating Spectrometer (RGS: Den Herder et al. 2001) modules allowing for simultaneous imaging and spectroscopy. These detectors are used in conjunction with three high-throughput X-ray telescopes. XMM-Newton also carries a small optical telescope, the Optical Monitor (OM: Mason et al. 2001).

We use the ninth data release of the fourth XMM-Newton catalogue, released in 2019 December (4XMM-DR9) (Webb et al., 2020). The observations were taken between 2000 February 3 and 2019 February 26 and the catalogue contains 810795 X-ray detections pertaining to 550124 unique sources.

Chapter 3

The SOUX AGN sample: Sample Definition

3.1 Introduction

Active galactic nuclei (AGN) are powered by mass accretion onto the supermassive black holes at the centres of galaxies. Most of the gravitational potential energy of the infalling material is liberated as radiation across the full electromagnetic spectrum. For typical, luminous AGN, the majority of their power emerges in the ultraviolet and optical bands as thermal radiation from a viscous accretion disc (Shakura & Sunyaev 1973; Novikov & Thorne 1973). The non-thermal, hard X-ray emission emerges via the Compton upscattering of soft seed photons from the disc in a hot ($T_{\rm e} \sim 100$ keV) plasma above the disc (the 'corona' e.g. Haardt & Maraschi 1991). A strong relationship between the ultraviolet and X-ray emission of AGN was found in early studies (e.g. Tananbaum et al. 1979) and its nature has been explored further in more recent investigations (e.g. Lusso & Risaliti 2017, Salvestrini et al. 2019). The non-linear and remarkably tight relation between the ultraviolet and X-ray luminosities enables luminous AGN to be used as 'standardisable candles' to test cosmological models (e.g. Sacchi et al. 2022). Clearly the disc and the corona are intimately related, but a detailed description of how has yet to be formulated.

For the majority of AGN, the peak of their energy output occurs in the far-UV/soft X-ray band which is unobservable because of Galactic extinction. This complicates the determination of the energetics and the disc-corona relationship. One means of circumventing this problem is to make studies of high-z AGN for which the spectral peak is redshifted to observable optical-UV wavelengths, as was done by Collinson et al. (2015). Making detailed

studies of the optical and X-ray spectra of 11 sources, Collinson et al. (2015) were able to accurately determine the bolometric luminosities and accretion rates of the AGN, demonstrating estimates of the bolometric luminosity made by scaling from a single wavelength or narrow-band luminosity (e.g. at 2500 Å or 2–10 keV) are generally inaccurate (see also Netzer 2019). However, obtaining high-quality multiwavelength spectra for high-z AGN necessarily restricts the sample to a small number of bright quasars, thereby limiting the parameter space which can be explored. Another approach is to sample the far-UV and soft X-ray emission on either side of the peak and use a physical model (e.g. Done et al. 2012) to 'bridge the gap' and recover the intrinsic spectral energy distribution (SED). This method was adopted by Jin et al. (2012a), who explored the optical and X-ray spectroscopic properties of 51 AGN in a series of works (Jin et al. 2012a,b,c). They showed that the broadband SEDs of AGN could generally be fit with three components: a standard accretion disc, a hot corona, and an intermediate 'warm corona' responsible for the observed excess of soft X-ray emission (e.g. Magdziarz et al. 1998; Gierliński & Done 2004). The findings of Jin et al. (2012a,b,c) have enabled the development and refinement of new, physical models of AGN SEDs (Kubota & Done 2018, 2019). These models have been constructed with reference to a small number of representative AGN SEDs. In the decade since the work of Jin et al. (2012a,b,c) the continuing SDSS (Blanton et al. 2017; Pâris et al. 2018) and *XMM-Newton* (Webb et al. 2020) surveys have greatly increased the number of AGN with quality optical/UV and X-ray spectra. We now have the opportunity to test the predictions made by current physical models against the observed properties of a large and diverse sample of AGN.

Achieving a more rigorous understanding of the accretion flow and disc-corona relationship will also enable us to better address another outstanding problem in astrophysics: the origin of relativistic jets. How AGN launch and power relativistic jets of outflowing matter is an active area of research. Progress towards a fuller understanding of the nature of jets and their relation to the accretion flow may be made via comparisons to accreting stellar-mass black holes (X-ray binaries: XRBs, e.g. Fender et al. 2004) in which the disc-corona system likely regulates the launching of jets. Many studies have explicitly investigated the analogies between AGN and XRBs in terms of the accretion flow properties (e.g. Noda & Done 2018; Ruan et al. 2019; Arcodia et al. 2020) and the disc-jet coupling (e.g. Merloni et al. 2003; Körding et al. 2006; Svoboda et al. 2017; Fernández-Ontiveros & Muñoz-Darias 2021; Moravec et al. 2022). Generally these studies are supportive of a unified scheme of mass accretion and ejection across the mass scale, although many open questions have yet to be resolved. Studies of high accretion rate AGN in the local Universe (narrow-line Seyfert 1s:

NLS1s) have found a relative deficit of radio-loud sources: only ≈ 5 per cent of NLS1s are radio-loud (e.g. Komossa et al. 2006; Rakshit et al. 2017) compared with ≈ 15–20 per cent of quasars (e.g. Kellermann et al. 1989), again suggesting a relationship between the accretion flow (the disc-corona configuration) and the presence of a radio jet.

Previous studies have mostly either investigated very large and diverse AGN samples without a detailed exploration of their spectroscopic properties (e.g. Svoboda et al. 2017) or have presented a thorough analysis of much smaller, focussed samples of AGN with high-quality spectroscopic data (e.g. Jin et al. 2012a,b,c and Collinson et al. 2015, 2017).

3.2 This Chapter

In this Chapter we expand on the work of Jin et al. (2012a,b,c) and Collinson et al. (2015, 2017) and make use of recent optical, UV and X-ray catalogs to compile a large sample of AGN (696 unique sources) with both optical and X-ray spectra in addition to broad multi-wavelength coverage. This new sample has the advantage of being large enough to investigate population statistics whilst still having good quality multi-wavelength data. It has been selected to probe a diverse range of AGN properties; as we will demonstrate, the sample spans several orders of magnitude in black hole mass and luminosity (accretion rate). The sample is a well-suited selection with which to perform detailed investigations of the evolution of the disc-corona system with black hole mass and accretion rate (Chapter 4). In this Chapter we make key measurements of AGN properties and take a preliminary look at the evolution of the spectral energy distribution with both mass and luminosity. In Section 3.3 we describe the assembly of the sample; in Section 3.4 we obtain multi-wavelength measurements for the selected AGN; in Section 3.5 we determine black hole masses from optical/UV data; we investigate the UV/X-ray energy index $\alpha_{\rm ox}$ and make comparisons with a physical model in Section 3.6; finally in Sections 3.7 and 3.8 we discuss our results and summarise our conclusions. To convert fluxes to luminosities we have assumed a flat ΛCDM cosmology with $H_0 = 70\ {\rm km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\rm m} = 0.3$ and $\Omega_\Lambda = 0.7$.

3.3 Sample selection

3.3.1 The parent catalogues

Optical spectra: SDSS-DR14Q

The Quasar Catalog of the Fourteenth SDSS Data Release (SDSS-DR14Q, Pâris et al. 2018) contains 526356 quasars. Each quasar has been the subject of least one optical spectroscopic observation recorded on or before 2016 December 5. The catalogue also contains SDSS photometric data from imaging observations. Multiwavelength data from *ROSAT*, *XMM-Newton*, *GALEX*, 2MASS, *WISE*, UKIDSS and FIRST are included, where available. The *XMM-Newton* data are taken from the 3XMM-DR7 catalogue (an earlier version of the serendipitous survey catalogue: see below). 14736 of the SDSS-DR14Q quasars (≈ 3 per cent) have an X-ray source detection within a 5 arcsec matching radius. Although the catalogue contains some X-ray data obtained from an *XMM-Newton* catalogue (3XMM-DR7), it does not contain any optical/UV photometry recorded with the *XMM-Newton* Optical Monitor (OM).

X-ray data: 4XMM-DR9

4XMM-DR9 is the first iteration of the fourth source catalogue compiled from the *XMM-Newton* Serendipitous Survey; this catalogue, released in 2019 December, is the ninth data release from the survey overall (Webb et al. 2020). The catalogue contains 810795 X-ray detections pertaining to 550124 unique sources from observations made between 2000 February 3 and 2019 February 26.

Optical and ultraviolet photometry: XMM-SUSS4.1

We take *XMM-Newton* optical/UV photometric data from the fourth *XMM-Newton* Serendipitous Ultraviolet Source Survey (SUSS: Page et al. 2012) catalogue, XMM-SUSS4.1. This catalogue, released in 2018 January, compiles data from observations made up to 2017 July. It contains 8176156 entries from 5503765 sources.

Selection biases

As we primarily utilise SDSS for the sample selection, our sample is subject to the biases inherent in the SDSS selection pipeline. One such bias is towards the selection of blue

quasars rather than reddened sources which are often missed by SDSS due to the in-built colour selection criteria. This however should not present much of an issue to the work presented in this Chapter and in Chapter 4 as we intend to focus on sources which do not display significant levels of intrinsic reddening. This is particularly true of the work presented in Chapter 4 where we aim to model the intrinsic SED's of our sample which we assume to be devoid of intrinsic reddening.

In addition as we require a minimum X-ray count in XMM we shall miss many sources which may be bright in SDSS but are X-ray weak. This may mean that we are less likely to observe some super Eddington sources for example which are bright in the optical/UV but relatively weak in the X-rays. We do expand our fitting to the wider SDSS parameter space in Chapter 4 and provide evidence to suggest that the X-ray selection does not bias our conclusions regarding the shape of the optical/UV continuum with regards to the wider SDSS population.

We do however note that the completeness of our sample is fundamentally limited by the biases inherent in the SDSS quasar selection pipeline and the XMM observational procedure along with the slection criteria that we ourselves impose. Despite this, in both this Chapter and Chapter 4 we are able present results that have relevance to the wider AGN population, particularly in addressing some of the tensions which exist between canonical observational and theoretical paradigms of AGN study.

3.3.2 Initial source selection criteria

Quasars

We first establish which optical quasars have an X-ray source detection in the latest *XMM-Newton* data release. To do this, we cross-match the optical catalogue with the slimline version of the X-ray catalogue. The slimline 4XMM-DR9 catalogue contains just one row per unique X-ray source (rather than one row per detection as in the full catalogue). There are 550050 X-ray sources in the catalogue, excluding 74 which are flagged as 'CONFUSED' (i.e. with a nonzero probability of being associated with more than one distinct source). Following Pâris et al. 2018, we choose a match radius of 5 arcsec and retain all of the best (nearest) matching X-ray sources. Fig. 3.1 shows the distribution of angular separations between the optical quasars and X-ray sources, along with the distribution of match density as a function of separation. We model the false association tail of the cumulative match curve and use

the intercept as an indicator of the number of 'real' associations that we would expect. The values obtained consistent with the number of matches we obtain using a 5" matching radius. Our cross-match returns 17336 X-ray detected optical quasars (approximately 3.3 per cent of those in SDSS-DR14Q). We record the unique X-ray source ID ('SRCID') of each quasar. We do not remove broad absorption line (BAL) quasars from the optical catalogue initially, but will weed these out of our sample by visually inspecting the optical spectra.

Before performing a cross-match to the full X-ray catalogue, we clean it by excluding the sources which:

- have a poor source summary flag (SUM_FLAG $\geqslant 3$);
- are flagged as 'CONFUSED';
- were observed with a high background.

The clean catalogue then contains 88 per cent of the total number of records, with 713866 X-ray detections from 484726 unique sources.

Using the X-ray source IDs, we find all clean observations of the optical quasars. We match all records in the cleaned, full 4XMM-DR9 catalogue to our pre-selected optical quasars. We find that 17299 of the pre-selected 17336 quasars have a clean X-ray detection. Of these, 5364 (31 per cent) were observed by *XMM-Newton* more than once; Fig. 3.2 shows the number of clean X-ray observations per optical quasar.

We then match with sources in the OM catalog. This matching requires that the OM source is within 1 arcsecond of the optical quasar coordinates and that the observation ID of the OM record matches that of the X-ray observation (i.e. our OM and EPIC data were recorded simultaneously). 4499 quasars meet these criteria of which 868 have more than one simultaneous OM-EPIC observation.

At $z \gtrsim 0.4$, the OM photometry may be compromised by the strong Ly α λ1216 UV emission line and Ly α forest absorption blueward of 912 Å (rest frame). To determine which OM filters would be free of Ly α emission and absorption, we calculated the quasar redshifts for which the red wing of Ly α would fall just outside of the effective bandpass of each filter. Then, for each redshift bin we require an OM detection in one of the following:

- $z < 0.40$: any filter;
- $0.40 \leqslant z < 0.55$: V, B, U, UVW1 or UVM2 filters;

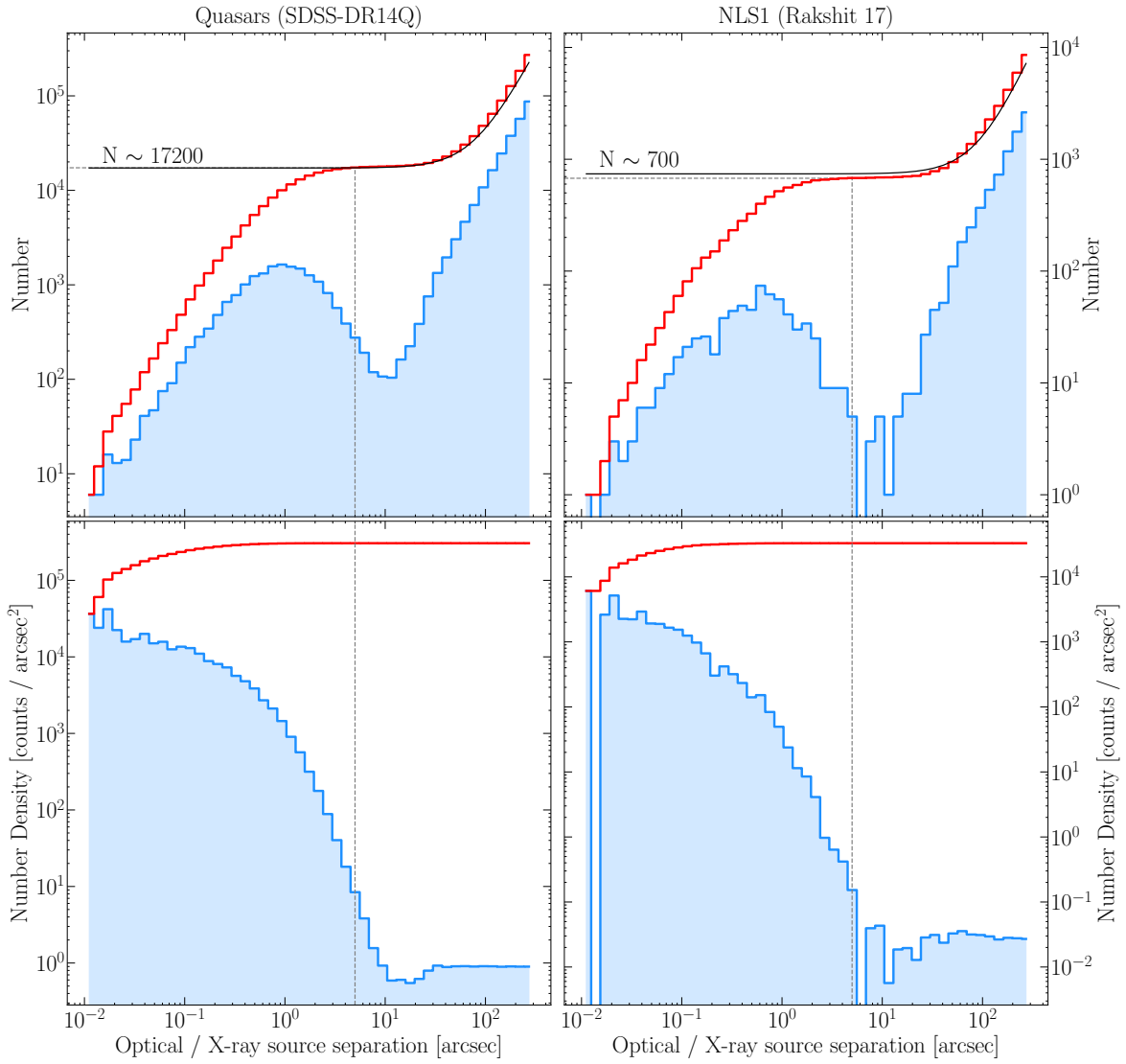

Figure 3.1: *Left-hand Panels*: Matching results between the SDSS-DR14 quasar catalogue and the 4XMM-DR9 catalogue. *Right-hand Panels*: Matching results between the Rakshit et al. 2017 NLS1 catalogue and the 4XMM-DR9 catalogue. *Top Panels*: Histogram of the angular separations between the two optical source catalogues and the 4XMM-DR9 X-ray sources as a function of matching radii. *Bottom Panels*: Histogram of the density of source separations between the two optical source catalogues and the 4XMM-DR9 X-ray sources as a function of matching radii. In all panels, the cumulative total is shown in red and the counts / density for each given matching radius is shown in blue. We fit the cumulative associations curve above a radius of 45" in the top two panels with a quadratic and use the intercept to give an indication of the number of 'real' associations that we would expect. These values are consistent with the 17336 quasars and 676 NLS1 sources matched using our selected matching radius of 5", marked with the dashed line on all panels.

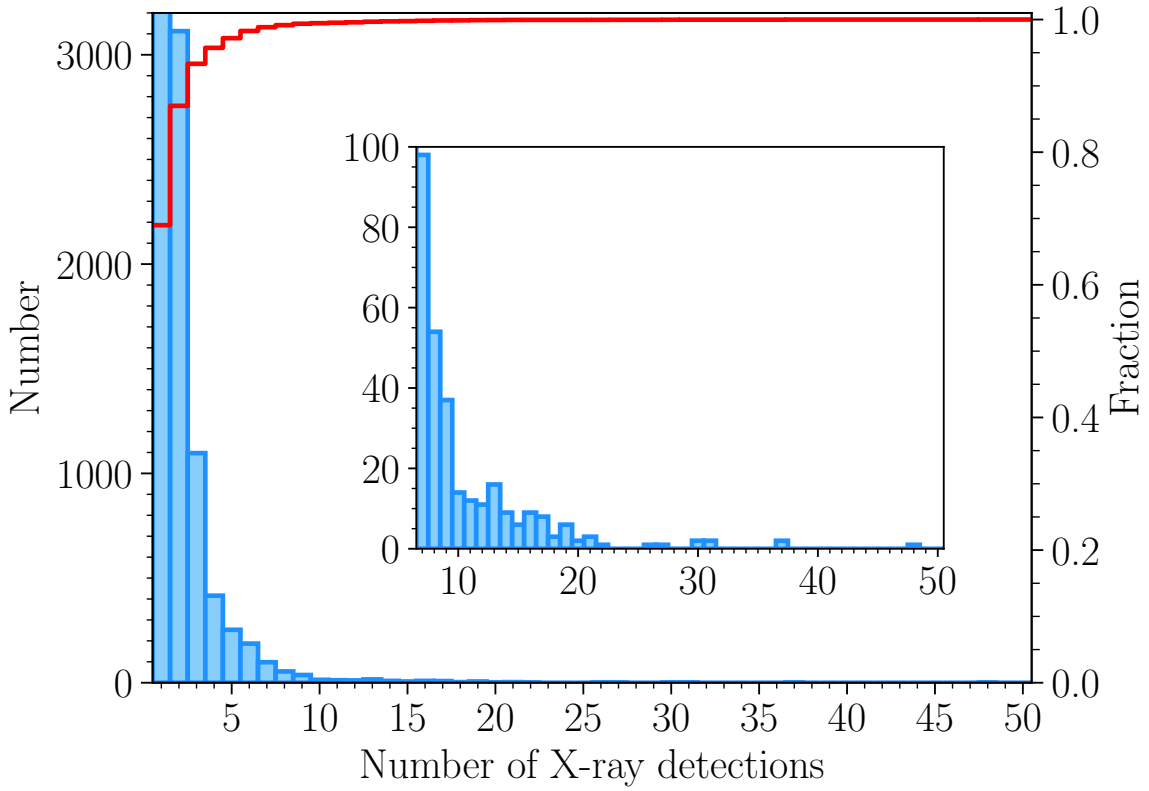

Figure 3.2: The number of clean X-ray detections for 17299 X-ray detected optical quasars. 11935 quasars (69 per cent) have only 1 clean X-ray detection. Only 109 have 10 or more detections. The red line shows the cumulative fraction, with the fraction indicated on the right-hand ordinate axis.

- $0.55 \leqslant z < 0.90$: V, B, U or UVW1 filters;
- $0.90 \leqslant z < 1.35$: V, B, or U filters;
- $1.35 \leqslant z < 1.95$: V or B filters;
- $1.95 \leqslant z < 2.50$: V filter only.

1768 quasars have at least one useful OM photometry point.

We wish to obtain reasonable constraints on the X-ray spectral parameters of these quasars, for which we require relatively high-quality spectra. We therefore make a quality cut on the number of X-ray counts recorded by an EPIC detector in the 0.2–12 keV band. The number of sources as a function of X-ray counts is shown in Fig. 3.3. It can be seen in Fig. 3.3 that the counts distribution peaks $\gtrsim 200$ counts and that only 17 per cent of detections are made with 1000 counts or more. We chose to make a cut at 250 counts, which includes 46 per cent of the clean detections. For quasars with more than one X-ray observation, we retain the one with the greatest number of counts in any detector (so that we now have only one simultaneous EPIC and OM observation of each quasar). After applying this cut, we are left with 782 quasars.

To make an estimate of the black hole mass, we require spectral coverage of at least one of the emission lines H α, H β or Mg II. H α or H β will be visible in the spectra of low-redshift sources and the Mg II emission line will be visible in the wavelength range of the BOSS spectrograph up to $z \lesssim 2.5$. We therefore consider only the SDSS quasars with $z \leqslant 2.5$. Following this redshift cut, our sample contains 768 quasars.

Narrow-line Seyfert 1s

Narrow-line Seyfert 1s (NLS1s) are a subset of unobscured AGN with relatively narrow broad lines. They are generally defined as optical type 1 AGN with FWHM(H β)$\lesssim 2000$ km s^{-1}, weak [O III]λ5007 relative to H β and strong Fe II emission (e.g. Osterbrock & Pogge 1985, Goodrich 1989). NLS1s are typically found in spiral galaxies in the local Universe. Because Seyfert AGN are less luminous than quasars, many NLS1s are absent from the SDSS quasar catalogues. For example, of the 12 NLS1s in the sample of Jin et al. (2012a), only 2 are present in our selection of quasars. Nevertheless, NLS1s represent a unique and important region of parameter space to explore, and enable contrasts to be made with typical broad-line AGN and quasars (e.g. Jin et al. 2012a,b,c).

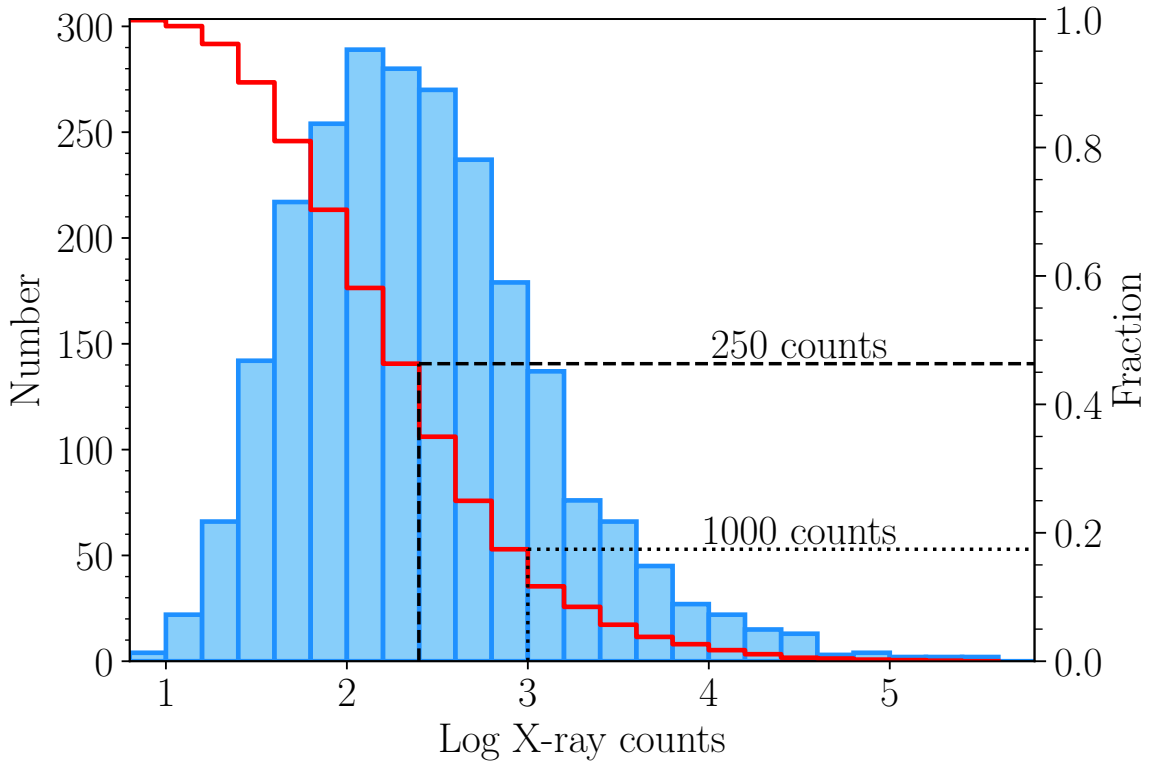

Figure 3.3: The X-ray counts distribution for 2374 'clean' X-ray observations with simultaneous *XMM*-OM photometry redward of Ly α. For each observation, we take X-ray counts to be the greatest number of counts recorded by any of the three EPIC detectors (pn, MOS1 or MOS2). The left-hand ordinate axis gives the total number of observations and the right gives the fraction with a minimum number of counts (shown by the red line).

Rakshit et al. (2017) (hereafter R17) assembled a sample of 11101 NLS1s from the SDSS-DR12 spectroscopic database, by analysing the spectra of 'QSO'-type sources and retaining those with FWHM(H β) $\leqslant$ 2200 km s^{-1} and flux ratio [O III]λ5007/H β $<$ 3. We identify NLS1s with *XMM-Newton* OM and EPIC data by matching the R17 catalog to the X-ray and OM serendipitous catalogs, following the method described in Section 3.3.2. Our criteria return 147 NLS1s: 63 new sources and 84 that were already included from the quasar catalogue.

3.3.3 Optical spectral fitting

Rakshit et al. (2020) (hereafter R20) performed fits to the optical spectra of all SDSS DR14Q quasars using the the publicly-available PYTHON code PyQSOFit (Guo et al. 2018, Guo et al. 2019, Shen et al. 2019). For the quasars in our sample, we can take many emission line and continuum measurements from R20. However, the 63 NLS1s not present in the quasar catalog have only optical spectral measurements from R17 who used a private, custom spectral fitting routine to obtain their measurements of NLS1s. These two routines differ in several respects. The broad emission lines were often modelled with a single Lorentzian profile in R17; Lorentzian profiles are not currently an option in PyQSOFit and instead one or more Gaussians may be used to fit a line. R20 chose to use 3 Gaussians to model broad H β and H α, for example. The host galaxy subtraction, which is often required for the low-luminosity NLS1s, is also approached in different ways. R17 removed the host galaxy contribution by performing simple stellar population modelling whereas principal component analysis (PCA: e.g. Yip et al. 2004a,b) is employed in PyQSOFit. Naturally these different methods will yield somewhat different results for the continuum and emission line parameters. Derived quantities such as the virial black hole mass and Eddington ratio will then also differ.

For consistency of the optical spectral measurements of our sample, we adapted PyQSOFit to replicate the spectral fitting procedure of R20 and refit the spectra of the sources drawn from R17. We briefly summarise the adopted PyQSOFit spectral fitting routine here. In this setup, the Schlegel et al. (1998) dust map is queried to determine the line-of-sight extinction and the spectrum is dereddened using the extinction curve of Fitzpatrick (1999). The spectrum is then transformed to the rest-frame. The PCA method is used to decompose stellar and AGN emission with 5 and 20 eigenspectra used to construct the host galaxy and quasar contributions, respectively. The host galaxy subtracted continuum is then modelled with a power-law plus Fe II templates and a functional Balmer continuum. Emission line complexes are then modelled separately in several broad windows:

- 6400–6800 Å: H α, the [N II] doublet ($\lambda = 6549, 6585$ Å) and the [S II] doublet ($\lambda = 6718, 6732$ Å);
- 4640–5100 Å: H β, the [O III] doublet ($\lambda = 4959, 5007$ Å) and He II;
- 2700–2900 Å: Mg II.

Following the fitting procedure of R20, the broad components of H α and H β and He II are each modelled with 3 Gaussian components; broad Mg II is modelled with 2 Gaussians. The broad Gaussians have a minimum FWHM of 900 km s^{-1} and maximum velocity offset $\pm$3000 km s^{-1} from their centre. Narrow Gaussians are used to model forbidden lines and the narrow components of the permitted lines; these are allowed a width in the range $100 \leqslant$ FWHM $\leqslant$ 900 km s^{-1}. [O III]$\lambda\lambda$4959, 5007 additionally has Gaussian wings that may be broader than the narrow cores (in the range $400 \leqslant$ FWHM $\leqslant$ 2000 km s^{-1}). The velocity offsets of the narrow lines are not allowed to exceed $\pm$1000 km s^{-1}. The flux ratios of the [O III] and [N II] doublets are fixed to their theoretical value 1:3. In each emission line complex the widths and offsets of the narrow lines are tied together. Uncertainties on the emission line and continuum paramaters are estimated using the Monte Carlo method, resampling each spectrum 50 times. Following the procedures of R20 and Calderone et al. (2017), we added a quality flag for each measured emission line and continuum point; the flags are listed in Appendix A.3.

PyQSOFit uses optical galaxy eignenspectra to perform the host galaxy subtraction; these eigenspectra do not extend into the UV and therefore the fitted spectrum is truncated to match the wavelength range of the eigenspectra ($\approx$ 3500–9000 Å). Consequently, there are no measurements of the UV continuum and emission lines for sources in which host galaxy subtraction was performed in the optical. To obtain the UV spectral measurements for these AGN we checked for spectra that contained an Mg II line that was not initially measured and then fit the short-wavelength part of these spectra ($\lambda_{\rm rest}$ = 1000–4000 Å) without performing host-galaxy subtraction. The measurements for Mg II and the UV continuum were added to our catalogue, and we have added a flag ('MGII_SEP') to indicate the AGN for which the Mg II region was fitted separately from the optical region.

Identification of broad- and narrow-line type 1 AGN

To ease the comparison between AGN with relatively narrow and broad permitted emission lines, we assigned each AGN a linewidth flag ('B' for broad-line type 1 AGN; 'N' for narrow-

line type 1 AGN). The determination is based on the fitted width of H β where it is available, Mg II for the higher-redshift sources without H β, and H α only if neither H β or Mg II were measured. We set the division between the subsets at the usual 2000 km s^{-1}. The differences in fitting routine and the division between narrow- and broad-line AGN mean that our classifications differ from those of R17 (those authors measured lines by fitting a Lorentzian profile and chose a limit of FWHM[H β]$\leqslant$ 2200 km s^{-1}). Consequently, not all AGN drawn from R17 are flagged 'N' in our sample. Typically, the broad line FWHMs measured by PyQSOFit are slightly broader than those measured by R17. However, the differences in linewidth measured by the two routines are relatively minor: the median difference in FWHM(H β) is $\approx +100$ km s^{-1} and for approximately two-thirds of the AGN from the NLS1 catalogue in our sample, FWHM(H β) as measured by PyQSOFit is within $\pm$350 km s^{-1} of the width measured by R17 (350 km s^{-1} is the median uncertainty on FWHM[H β] from PyQSOFit). We note that the narrow-line type 1 AGN in our sample are not necessarily NLS1s by the standard definition described in Section 3.3.2. We have not always used broad H β as the characteristic line width and we have made no assessment of the strengths of [O III] λ5007 or the optical [Fe II] emission.

3.3.4 Quality checks and selection of the final SOUX sample

We performed a visual inspection of all 831 optical spectra and removed any sources which had low signal to noise or any absorption features particularly affecting the line regions from which black hole mass measurements were derived. In addition any objects with a diffuse emission flag in their X-ray spectra were removed from the sample. Following this quality control, our final sample consists of 696 AGN[1]. Of these, 636 are drawn from SDSS-DR14Q, 60 unique sources are added from R17 and 72 are common to both catalogues.

3.4 Multiwavelength measurements

We will present a detailed study of the broadband spectral energy distributions (SEDs) of our AGN in Chapter 4. However, the SED shape can be characterised by two common parameters: the radio loudness and UV-X-ray energy index. The radio-loudness parameter indicates the power of AGN outflows relative to the accretion flow, with radio-loud AGN exhibiting

[1]In this work, we consider only one optical spectrum per source: that listed as the 'science primary' spectrum at the time of the catalogue compilation. However, 116 AGN have multiple spectra recorded on different nights; we describe these in Appendix A.2

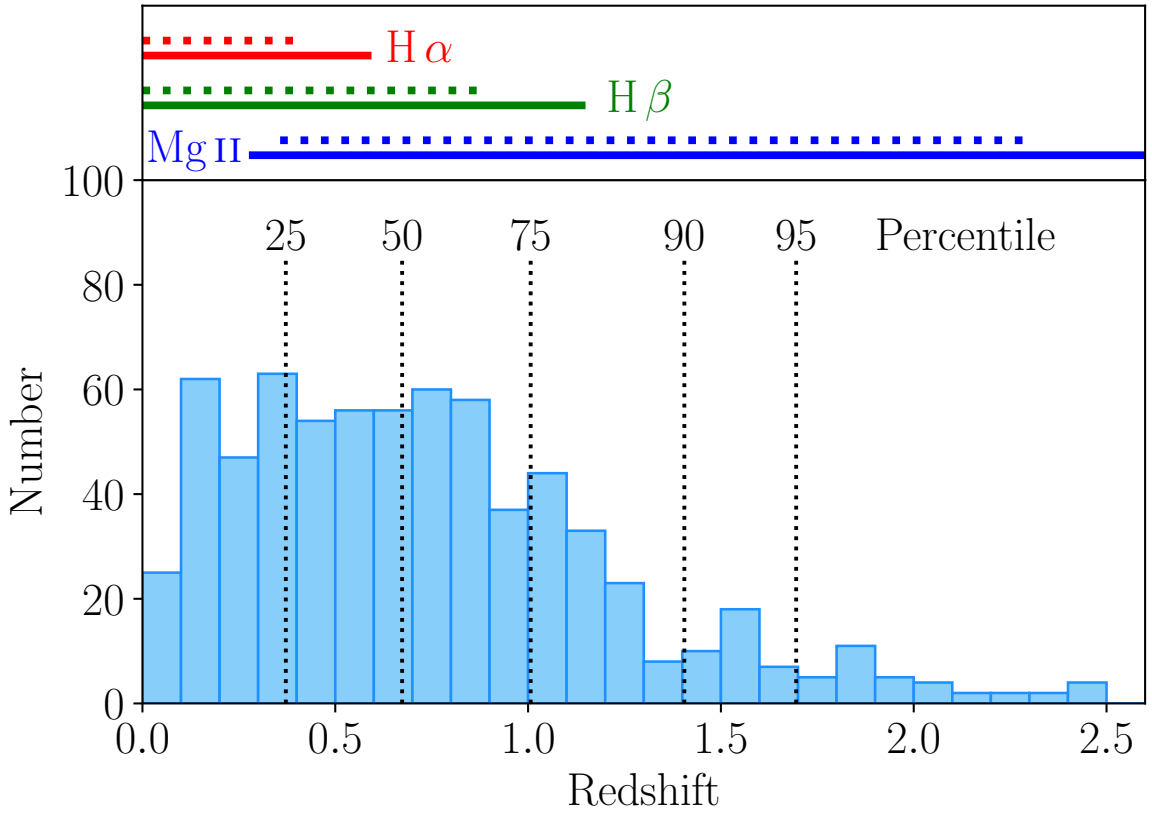

Figure 3.4: The redshift distribution of the 696 AGN in our sample. Black dotted lines indicate the 25$^{\text{th}}$, 50$^{\text{th}}$, 75$^{\text{th}}$, 90$^{\text{th}}$ and 95$^{\text{th}}$ percentiles of the distribution. The top panel shows the redshift range for which the centroids of prominent broad emission lines (H α, H β and Mg II) are just visible in the wavelength coverage of the SDSS (dotted lines) and BOSS (solid lines) spectrographs.

powerful jets (Section 3.4.3). The optical luminosity at 4400 Å and radio luminosity at 5 GHz are required to calculate the radio-loudness. The UV-X-ray energy index ($\alpha_{\rm ox}$: Section 3.6) indicates the relative power of the UV-emitting accretion disc to the X-ray emitting corona. To calculate $\alpha_{\rm ox}$ we require a measurement of the UV luminosity at 2500 Å and the X-ray luminosity at 2 keV. In this Section, we describe how the measurements of these quantities were obtained.

3.4.1 UV and optical luminosities

For sources at $z \gtrsim 0.5$ the rest-frame monochromatic luminosity at 2500 Å is available directly from the SDSS optical spectrum. For nearer sources, this wavelength falls outside the SDSS spectrograph range, but photometry covering it may be available from either *XMM* OM or SDSS. For AGN at lacking spectroscopic coverage of 2500 Å we first attempt obtain an estimate from the *XMM* OM photometry. We take the *XMM* OM UVW1 band flux for AGN at $z < 0.2$; for AGN at $z \geqslant 0.2$ we take the U band flux, or the UVW1 band flux no U band measurement is available. If there is no *XMM* OM coverage of 2500 Å we take the PSF flux from the appropriate Sloan filter (either the u or g band). We deredden the photometry, taking the line-of-sight colour excess from the Schlegel et al. (1998) dust maps as was done for the spectral analysis and convert the fluxes to rest-frame luminosities. Finally, L_{2500} is obtained from the optical spectrum of 421 AGN, from *XMM* OM photometry of 226 AGN and from SDSS photometry of 26 AGN; overall, we have a measurement of L_{2500} for 673 AGN (96 per cent of the full sample). The AGN lacking a L_{2500} measurement are all at the low-redshift end of our sample ($z < 0.25$).

The optical continuum at rest-frame 4400 Å was not measured by R20 or R17 but, similarly to L_{2500}, L_{4400} can be measured for the majority of our AGN either from the SDSS spectrum, or from the Sloan or *XMM* OM photometry. When making a measurement from the optical spectrum, we avoid contamination of the continuum flux measurement by the blended H γ, [O III] and Fe II emission lines by performing a linear interpolation between line-free continuum windows at 4200 and 5100 Å, where possible. In spectra where these two windows are not present (or are too near the end of the spectrum to be reliable), we measure the 4400 Å flux directly. For AGN lacking optical spectral coverage of 4400 Å we took the flux from the photometric band containing rest-frame 4400 Å, corrected for Galactic reddening and converted the flux to a rest-frame luminosity, as was done for the L_{2500} measurements. In our final sample of 696 AGN, 482 have L_{4400} measured by interpolating the spectral contin-

uum; 111 from a direct measurement of the 4400 Å spectral flux; 38 estimates from SDSS photometry; and 65 have no 4400 Å coverage in either the spectrum or photometry. None of the 65 sources lacking L_{4400} from the SDSS spectrum or photometry had coverage in the *XMM* OM data because they were higher-z AGN and the OM filters do not cover long enough wavelengths to contain the rest-frame 4400 Å. The AGN lacking a L_{4400} measurement are all at the high-redshift end of our sample ($z > 1$).

3.4.2 X-ray luminosities

A detailed analysis of the X-ray spectra of our sources is beyond the scope of this Chapter but will be presented in future work. However, it is possible to estimate the X-ray luminosities from data provided in the DR9 catalog. Our sources span a wide range of redshifts and therefore the EPIC fluxes reported for various (observed) energy bands are not directly comparable since they sample different parts of the emitted spectrum.

To calculate the monochromatic X-ray flux at rest-frame 2 keV, we follow a similar approach to that of Lusso & Risaliti (2016) (see section 2.1 of that paper). We determine the EPIC flux in soft (0.5–2 keV) and hard (2–12 keV) energy ranges. An approximate correction for Galactic photoelectric absorption is applied to the X-ray fluxes[2]. From these band fluxes we estimate the monochromatic fluxes at 1 and 3.5 keV, assuming a photon index $\Gamma = 1.7$ in each energy band. We then translate these fluxes at observed energies to luminosities at rest-frame frequencies in $\log(\nu)$-$\log(\nu L_\nu)$ form. Using these two points, the rest-frame monochromatic 2 keV luminosity is calculated by linear interpolation (for sources with $z < 1$) or extrapolation (for sources with $z > 1$). An estimate of the X-ray photon index Γ is also made.

In Fig. 3.5 we show the distributions of $\log(L_{2\,\mathrm{keV}})$ and Γ. Our sample spans approximately 4 dex in X-ray luminosity, with median $\log(L_{2\,\mathrm{keV}}) = 26.2$ and $\sigma = 0.6$. The median estimated photon index is $\Gamma = 1.9$ with $\sigma = 0.4$. 673 of the AGN (97 per cent) have $1 \leqslant \Gamma \leqslant 3$. Less than 2 per cent of sources have $\Gamma < 1$, perhaps indicating absorption of soft X-rays in the AGN or host galaxy. The Γ distribution of the 101 narrow-line AGN is different from that of the 595 broad-line AGN. The narrow-line AGN have a median $\Gamma = 2.3$ and $\sigma = 0.5$, indicating that their X-ray spectra are generally softer than typical quasars. Kolmogorov-Smirnov and Anderson-Darling tests both indicate that the photon indices of the narrow-line and broad-line AGN are drawn from different distributions.

[2]The line-of-sight Galactic neutral hydrogen column density is obtained from the maps of the Leiden/Argentine/Bonn (LAB) Galactic H I Survey (Kalberla et al. 2005) using the Python tool `gdpyc` (`https://pypi.org/project/gdpyc/`).

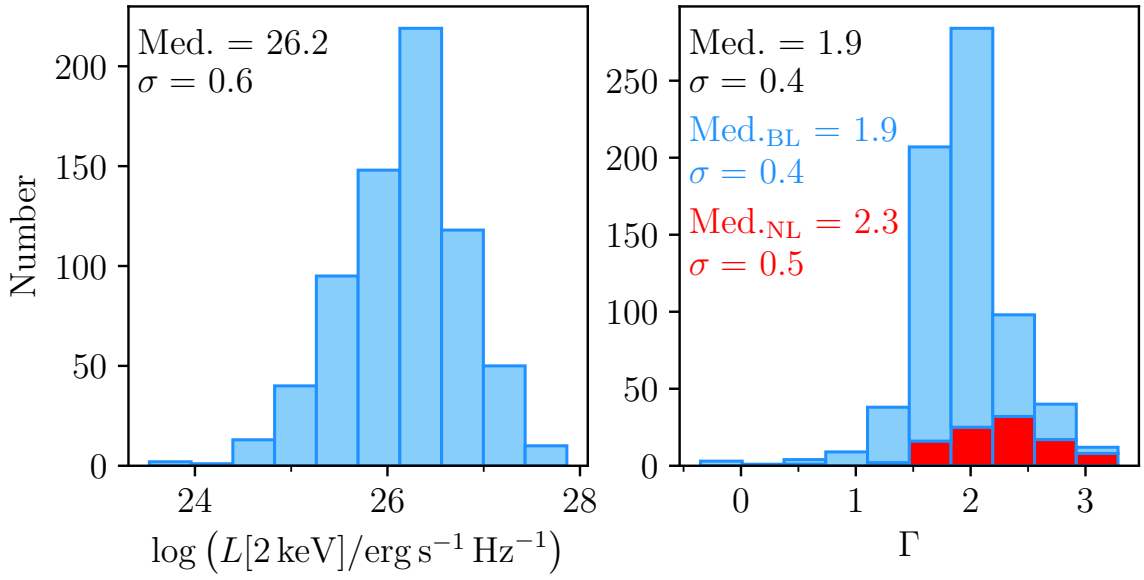

Figure 3.5: Distributions of the monochromatic luminosity at 2 keV in the rest frame (left) and the estimated X-ray photon index Γ (right). In the right-hand panel the subsample of 101 AGN with relatively narrow permitted lines ('NL') are shown in red, and the 595 broad-line sources ('BL') in blue.

3.4.3 Radio properies

Our sample of 696 AGN was cross-matched with both FIRST (Becker et al., 1995) and NVSS (Condon et al., 1998), using a matching radius of 10″ that has been shown to have a false association rate of only 0.2 per cent with FIRST (Lu et al., 2007). Both FIRST and NVSS sample the sky at 1.4 GHz, with beam-widths of 5.6″ and 45″ respectively. 679 of our AGN were within the FIRST footprint, of which we found a matching radio source for 107. 73 AGN also have an NVSS radio detection (all of which were FIRST-detected). We converted the 1.4 GHz fluxes to rest-frame 5 GHz luminosities using the method of Alexander et al. (2003), assuming a common radio spectral index ($\alpha_{\rm R}$) of 0.6. A comparison of the NVSS and FIRST luminosities for the 73 of objects matched with both catalogues can be seen in Fig. 3.6. It is clear that for the majority of sources the FIRST and NVSS luminosities match to within a factor of 2. The population is skewed towards higher luminosities in NVSS due to its larger beam width; this is most pronounced in objects which possess extended morphologies as can be seen by the radio map of the FR-II source SDSS J151443.07+365050.4 (Fig. 3.6, inset).

For the 101 radio-detected AGN with a L_{4400} measurement we calculated the radio loudness according to the common definition of $L_{5\,\rm GHz}/L_{4400}$ (Kellermann et al., 1989). Fig. 3.7 displays the distribution of the 5 GHz luminosities of the radio-detected AGN. The distribution has been separated into distinct populations: the left-hand panel shows the radio-loud and

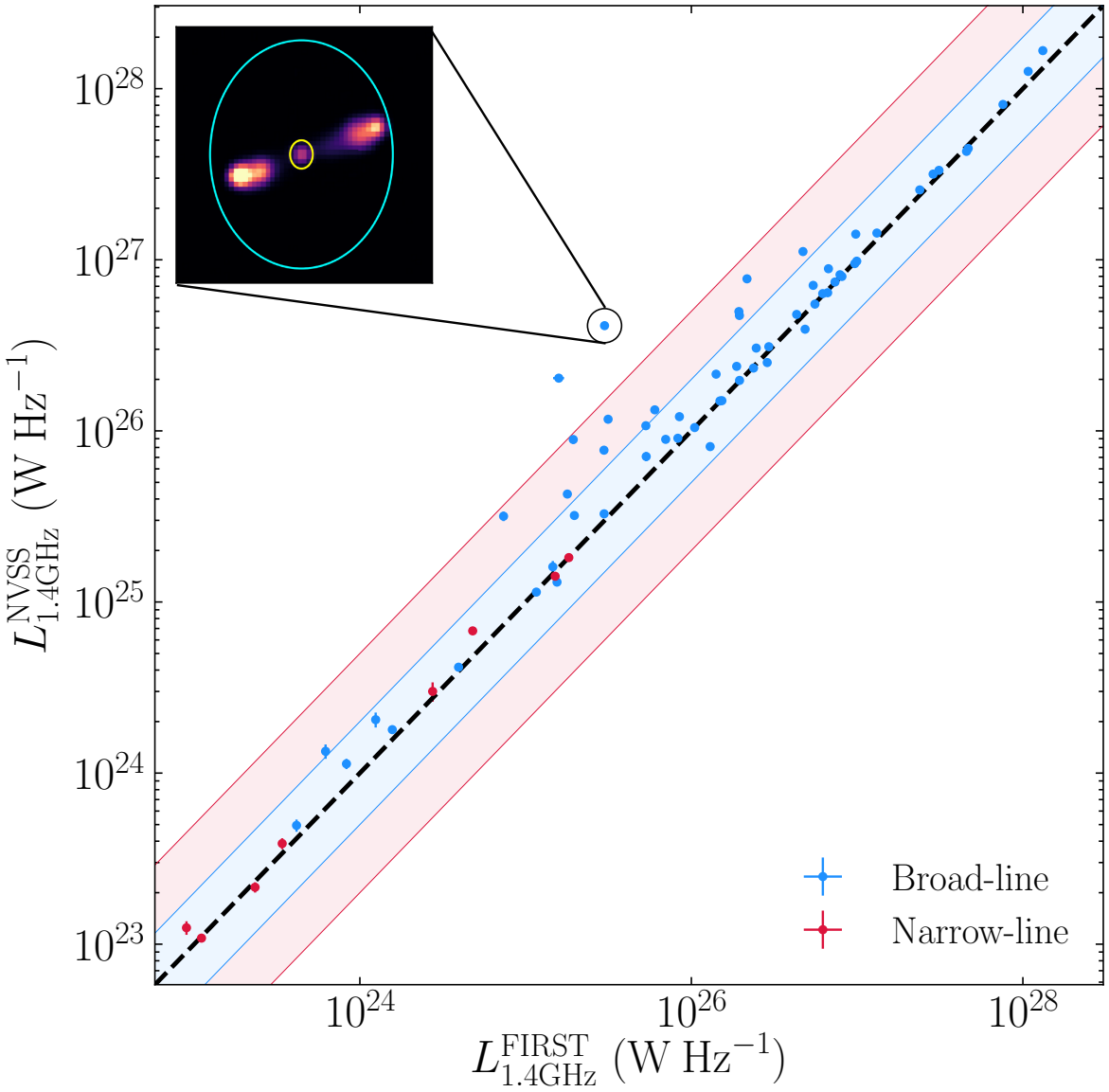

Figure 3.6: A comparison of the rest-frame luminosities taken from the FIRST and NVSS catalogues at 1.4 GHz, as described in the text. The dashed line marks the one-to-one relation, the shaded blue region being a factor of 2 deviation, and the pink shaded region a factor of 5. The top-left inset figure shows the radio map for the FR-II source SDSS J151443.07+365050.4, which has a much higher luminosity in NVSS than FIRST, as a result of the larger beam size of the former survey. The 45" beamwidth of NVSS is in blue, and the 5.6" beamwidth of FIRST is in yellow.

Table 3.1: Radio-loudness of the sample

Subset	N	Per cent	
No radio coverage (Outside FIRST footprint)	20	2.8%	of the full sample
No 4400 Å coverage	65	9.3%	of the full sample
Radio-detected and radio-loud	67	11.0%	of AGN with radio & 4400Å
Radio-detected and radio-quiet	34	5.6%	of AGN with radio & 4400Å
Radio-undetected and radio-quiet	210	34.4%	of AGN with radio & 4400Å
Radio-undetected, undetermined radio-loudness	300	49.1%	of AGN with radio & 4400Å

radio-quiet subsets and the right-hand panel shows the narrow-line and broad-line subsets (see Section 3.3.3). For many undetected radio sources, we can still estimate whether they are likely to be radio-quiet by considering the FIRST detection threshold as an upper limit to their radio flux. Assuming a FIRST survey flux density threshold of 1 mJy (Becker et al., 1995), we calculate a curve corresponding to the lowest rest-frame 5 GHz luminosity we would expect for a source to be detected; we show this curve in Fig. 3.8. For radio undetected sources we plot the radio luminosity corresponding to $R = L_{5\,\mathrm{GHz}}/L_{4400} = 10$ (the luminosity above which they would be considered radio-loud). Radio-undetected AGN with luminosities more than 3σ above the curve (210 AGN) are considered to be radio-quiet, since we would expect them to have been radio-detected if they were radio-loud. However, as Fig. 3.8 illustrates, there remain 300 AGN below the curve which are too faint for us to determine whether they are either radio-loud or radio-quiet. Following this exercise, we determine that our sample contains 67 radio-loud AGN, 244 radio-quiet AGN and a total of 385 AGN for which we are unable to determine the radio-loudness. The radio-loudness classifications are summarised in Table 3.1.

Overall, ~ 10 per cent of our sample with both FIRST and 4400 Å coverage is found to be radio-loud, similar to previous estimates for bright quasars (Kellermann et al., 1989; Miller et al., 1990). However, we caution that this radio-loud fraction is a lower limit because of the large number of sources in our sample without any radio-loudness determination. We tentatively find differences in the radio-loud fractions of broad-line and narrow-line type 1 AGN. Whilst narrow-line AGN make up ≈ 15 per cent of the sample overall, they constitute only ≈ 9 per cent of the radio-loud sources. This is consistent with previous findings that the radio-loud fraction of NLS1s is lower than that of typical broad-line AGN and quasars (Komossa et al., 2006; Zhou et al., 2006; Rakshit et al., 2017). Again, selection biases and incomplete data mean it is not possible to draw firm conclusions from our sample.

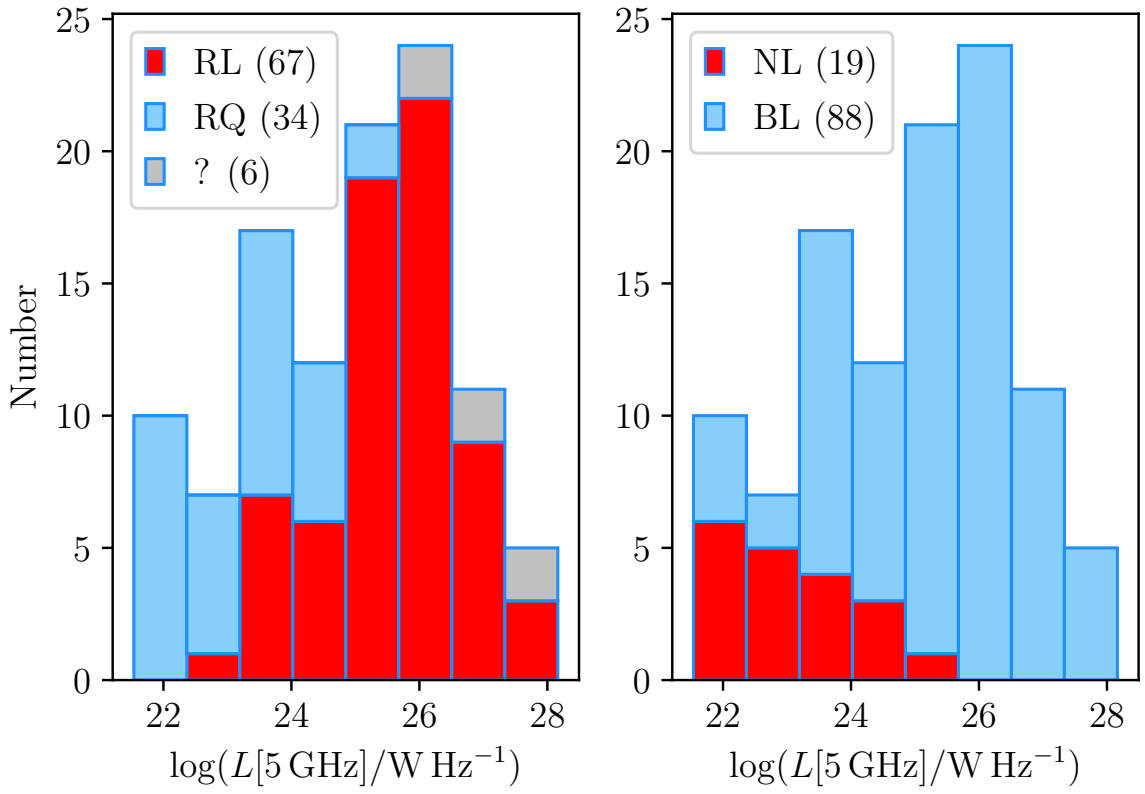

Figure 3.7: For the 107 FIRST radio detected AGN in our sample we show distribution of 5 GHz rest-frame radio luminosities derived from the observed 1.4 GHz fluxes, as described in the text. In the left panel, the distribution is divided into radio-loud and quiet ('RL' and 'RQ') subsamples in addition to 6 sources lacking an optical L_{4400} measurement with unknown radio loudness ('?'). In the right panel the distribution is divided into narrow-line and broad-line ('NL' and 'BL') subsamples on the basis of the relative widths of the permitted optical emission lines.

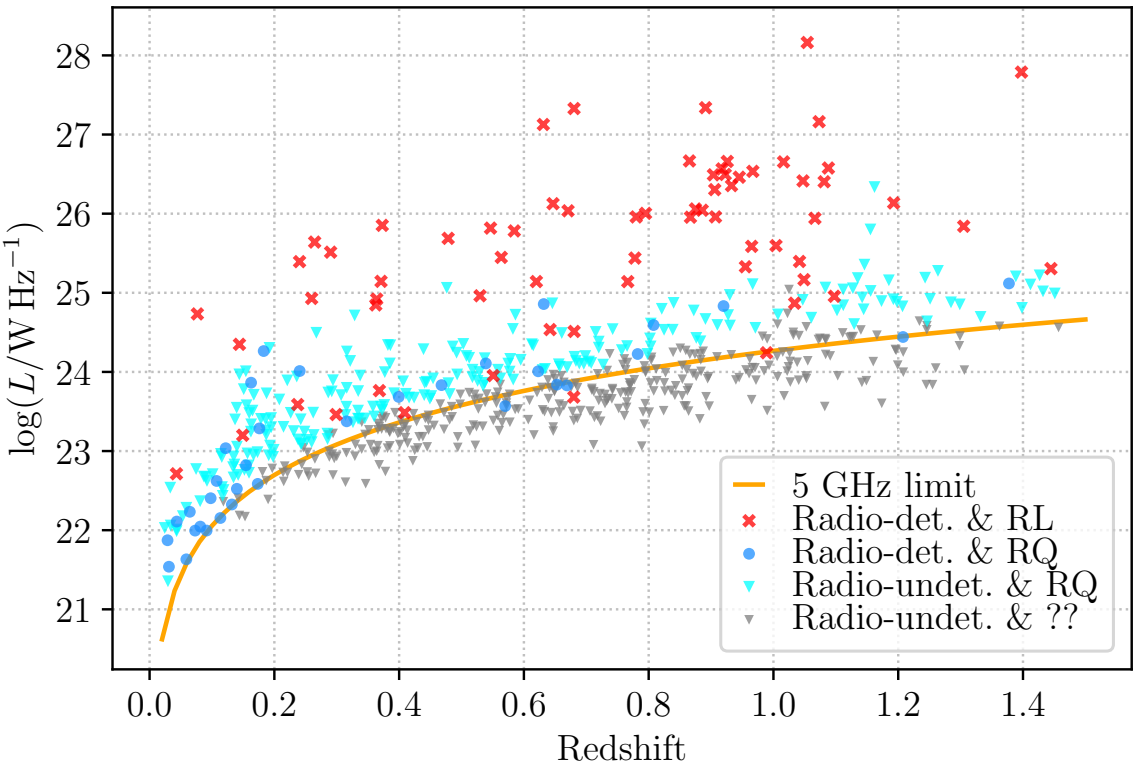

Figure 3.8: The orange curve shows the limiting rest-frame 5 GHz luminosity for a source to be detected by FIRST at 1.4 GHz, assuming a survey flux density threshold of 1 mJy and radio spectral index $\alpha = 0.6$. Radio-detected and radio-loud (radio-quiet) sources are shown with red crosses (blue circles). For radio-undetected sources, we plot the luminosity corresponding to $R = L_{5\mathrm{GHz}}/L_{4400} = 10$. Undetected sources more than 3σ above the curve (cyan triangles) are considered radio-quiet. Undetected sources within 3σ of the curve and below (grey triangles) could be either radio-quiet or radio-loud.

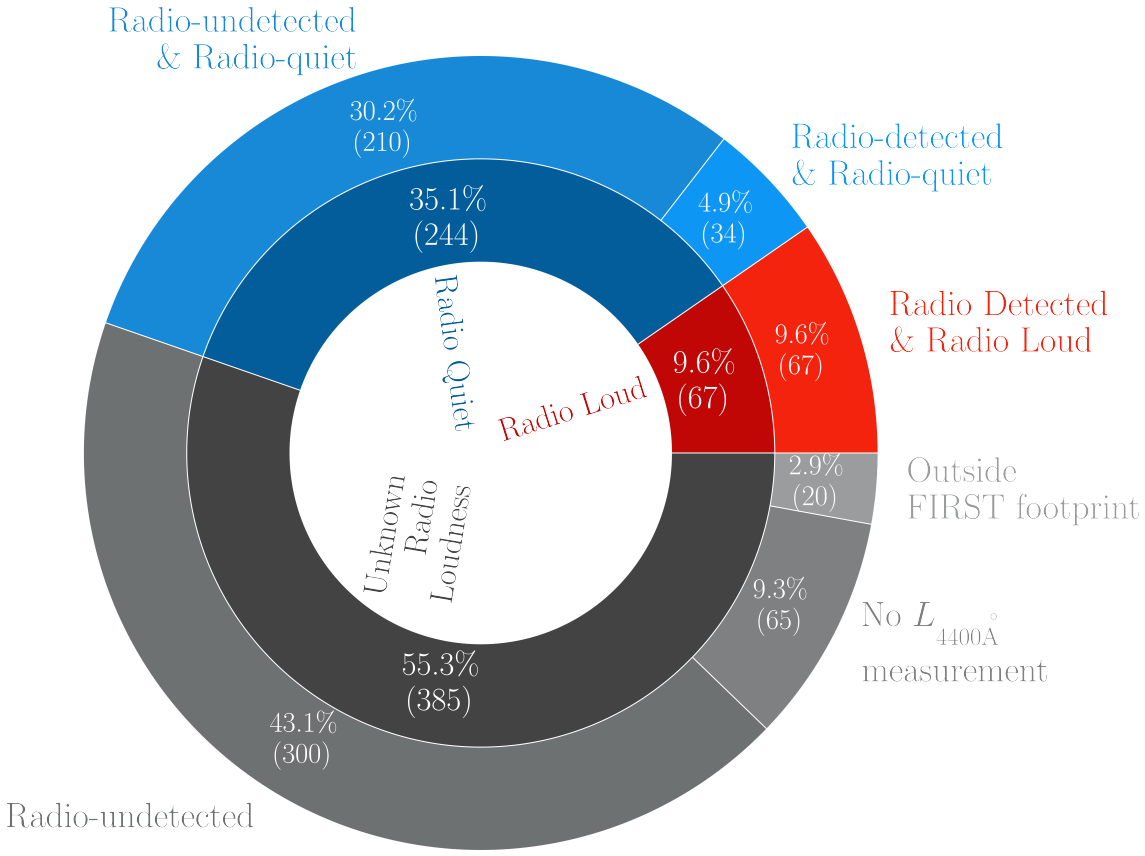

Figure 3.9: Radio detection and radio loudness fractions for the AGN in our sample.

3.5 Black hole masses

The widths and luminosities of the broad permitted emission lines, as well as the continuum luminosities can be used to derive the masses of the black holes in our AGN. In this Section we assess the broad emission line properties and calculate the black hole masses using new scaling relations.

3.5.1 Broad emission line properties

We compare the broad line widths for sources with spectra containing multiple lines. The line width comparisons are shown in Fig.3.10. We find that H β is slightly broader than both H α and Mg II. 207 spectra contain both H α and H β and from the median of the logarithmic distribution we find that H β is on average ≈ 10 per cent broader than H α. There is greater scatter between the widths of H β and Mg II, and we do not find strong evidence of a systematic difference in the line widths (H β is on average only ≈ 2 per cent broader than Mg II) as measured in 302 spectra. Qualitatively, these results somewhat are similar to those of previous studies, e.g. Osterbrock & Shuder (1982) found H β to be ≈ 16 per cent broader than H α in Seyfert galaxies; Mejía-Restrepo et al. (2016) also found H β to be broader than both H α and Mg II.

In Fig.3.11 we show the ratio of the H α to H β luminosities (the Balmer decrement). We find $\langle \mathrm{H}\,\alpha/\mathrm{H}\,\beta \rangle = 3.9$ which is higher than the Case B value 2.74 (Osterbrock & Ferland 2006) and the intrinsic value 2.72 ± 0.04 suggested by Gaskell (2017), based on observations of blue AGN. This Balmer decrement is also greater than the average found in previous studies of type 1 AGN: $\langle \mathrm{H}\,\alpha/\mathrm{H}\,\beta \rangle = 3.16$ (Lu et al. 2019); $\langle \mathrm{H}\,\alpha/\mathrm{H}\,\beta \rangle = 3.06$ (Dong et al. 2008); $\langle \mathrm{H}\,\alpha/\mathrm{H}\,\beta \rangle = 3.45$ (La Mura et al. 2007). From the Shen et al. (2011) Catalog of SDSS DR7 Quasar Properties, we determine $\langle \mathrm{H}\,\alpha/\mathrm{H}\,\beta \rangle = 3.55$ with $\sigma = 0.14$ dex. Taken as a reddening indicator, this large average Balmer decrement implies a high degree of dust extinction in our sample of AGN: on average the colour excess $E(B-V) \approx 0.37$ mag. However, if large a Balmer decrement is due to reddening, we would expect the optical continuum to be similarly affected such that AGN with a higher Balmer decrement will have a correspondingly greater value of L_{5100}/L_{3000}. The solid red line in Fig. 3.11 shows the expected trend; as can be seen in the Figure, the data do not follow this trend.

We have computed the best-fitting linear regressions between using the package LINMIX (Kelly, 2007). The results are shown in Fig. 3.12 and the values are quoted in Table 3.2.

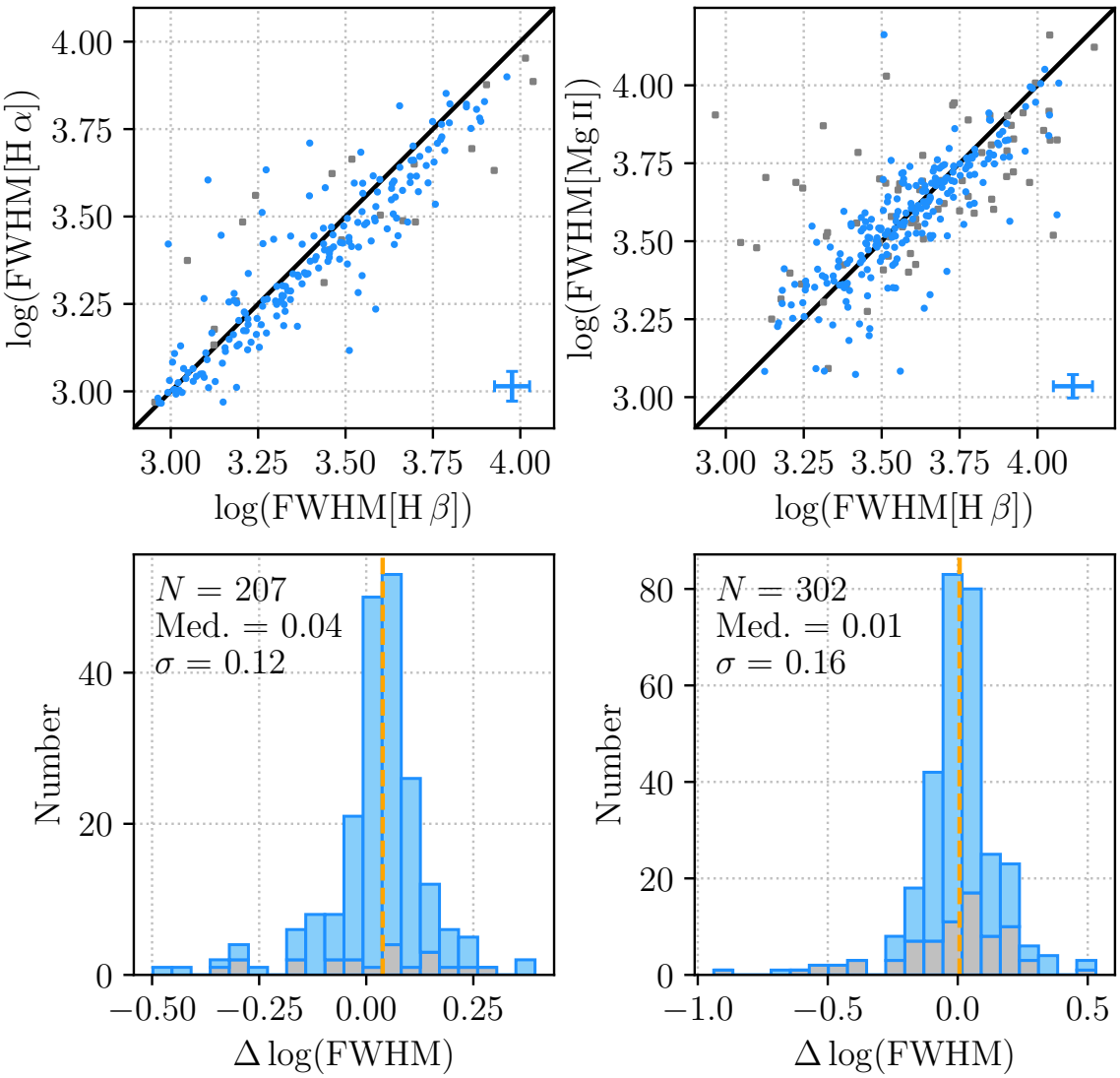

Figure 3.10: A comparison of the full widths at half maxima of H α, H β and Mg II. Upper panels show the values for sources with at least two lines in their spectra. Representative errorbars are shown in the lower right corners of the plots. The lower panels show the logarithm of the line width ratios.

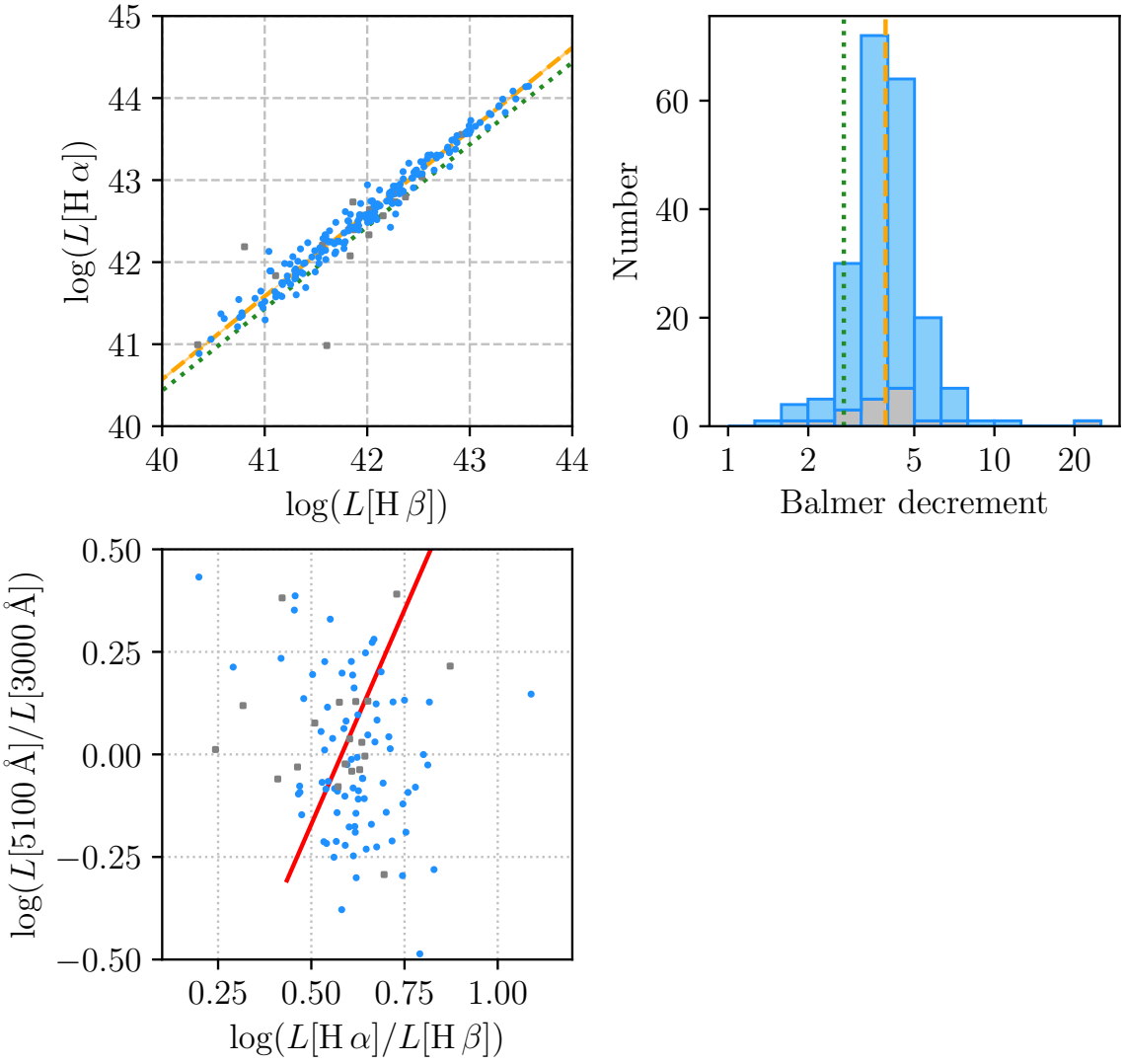

Figure 3.11: *Top left:* A comparison of the H α and H β broad line luminosities for low-z sources with spectra containing both lines. *Top right:* The distribution of the Balmer decrements of these sources. The orange dashed lines indicate the best-fitting linear regression, corresponding to $L_{\mathrm{H}\,\alpha}/L_{\mathrm{H}\,\beta} = 3.9$. The green dotted lines correspond to $L_{\mathrm{H}\,\alpha}/L_{\mathrm{H}\,\beta} = 2.72$ (Gaskell, 2017). *Bottom:* Balmer decrements versus ratios of the monochromatic continuum lumionsities at 5100 and 3000 Å. The red line shows the expected trend from a standard reddening curve. Data with good quality measurements of all quantities are shown as blue circles; the grey squares represent data in which at least one measurement has a quality flag raised.

Table 3.2: Correlations between emission line and continuum luminosities. We give the independent and dependent variables (x and y) and the slope and intercept (m and c) of the best-fitting linear regression. σ is the dispersion of the data about the linear regression.

x	y	m	c	σ
$\log L_{5100}$	$\log L_{\mathrm{H}\,\alpha}$	1.08 ± 0.02	-5.0 ± 1.0	0.20
$\log L_{5100}$	$\log L_{\mathrm{H}\,\beta}$	1.06 ± 0.02	-4.4 ± 0.7	0.18
$\log L_{5100}$	$\log L_{\mathrm{Mg\,II}}$	0.87 ± 0.03	4.4 ± 1.3	0.21
$\log L_{3000}$	$\log L_{5100}$	0.82 ± 0.01	7.8 ± 0.5	0.11
$\log L_{3000}$	$\log L_{\mathrm{H}\,\beta}$	0.89 ± 0.02	2.8 ± 0.9	0.20
$\log L_{3000}$	$\log L_{\mathrm{Mg\,II}}$	0.85 ± 0.02	5.1 ± 0.7	0.19

The Balmer lines show an almost 1:1 correspondence with the 5100 Å continuum luminosity, whereas the relation with Mg II is slightly shallower with a gradient $m = 0.87 \pm 0.03$. The relationships of both H β and Mg II luminosities with the 3000 Å continuum luminosity are shallower compared with those calculated for the 5100 Å luminosity. The strongest correlation we find is between the two continuum luminosities: the best-fitting linear regression has a slope $m = 0.82 \pm 0.01$ and a dispersion of only 0.11 dex (around half that of the emission lines).

3.5.2 Black hole mass estimates from broad emission lines

Assuming that the gas in the BLR is virialised, an estimate of the mass of the black hole (M_{BH}) can be made by applying the virial theorem:

$$M_{\mathrm{BH}} = f\frac{R_{\mathrm{BLR}}V^2}{G} \tag{3.5.1}$$

where R_{BLR} is the radius of the BLR, V is the velocity of the gas and G is the gravitational constant. R_{BLR} is often determined from radius-luminosity relations derived for reverberation-mapped AGN (e.g. Peterson et al. 2004; Bentz et al. 2006; Landt et al. 2011b). The virial factor f accounts for the geometry and inclination to the line of sight of the BLR and cannot be easily determined for individual sources. Where the broad line FWHM is used as a proxy for V then $f \sim 1$.

We first make estimates of the black hole masses of the AGN in our sample using recently-calibrated relations for virial black hole masses based on the broad H α, H β and Mg II emission lines. The relations used are listed in Table 3.3. It is common to use the AGN continuum luminosity as a proxy for the BLR radius, however Mejía-Restrepo et al. (2016), Greene et al. (2010) and Woo et al. (2018) provide relations using instead the broad emission line luminosity (of H α, H β, and Mg II, respectively). Both Mejía-Restrepo et al. (2016) and Woo

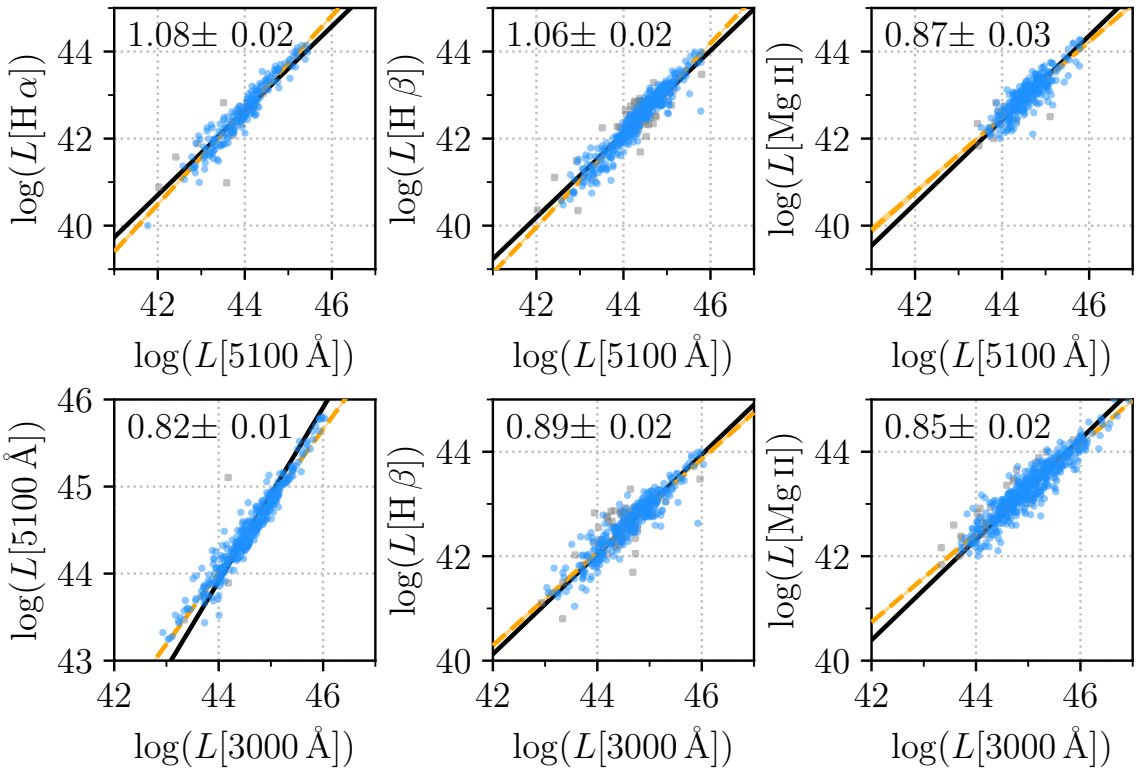

Figure 3.12: Correlations between emission line and continuum luminosities. Dashed orange lines show the best-fitting linear regression to the data, and the value of the slope (m) is given in the inset text. Relations with a gradient $m = 1$ are shown for comparison (solid black lines). Data with good quality measurements of all quantities are shown as blue circles; the grey squares represent data in which at least one measurement has a quality flag raised.

Table 3.3: The virial black hole mass relations from literature used.

FWHM	L	Relation ref.	$N_{\rm spectra}$	$N_{\rm spectra}^{\rm good}$
H α	5100 Å	MR16	219	189 (86%)
H α	H α	MR16	220	194 (88%)
H β	5100 Å	MR16	481	401 (83%)
H β	H β	G10	482	406 (84%)
Mg II	3000 Å	MR16	517	479 (93%)
Mg II	Mg II	W18	521	483 (93%)

We list the emission-line FWHM and luminosity L used to calculate the black hole masses. $N_{\rm spectra}$ is the number of spectra in which the relevant quantities were measured. $N_{\rm spectra}^{\rm good}$ is the subset of $N_{\rm spectra}$ (and the percentage) for which no quality flags were raised on the relevant emission line or luminosity measurements. *References*: MR16 = Mejía-Restrepo et al. (2016); G10 = Greene et al. (2010); W18 = Woo et al. (2018).

et al. (2018) provide relations to estimate the black hole mass directly from a luminosity and broad emission line FWHM, assuming $f = 1$ and $f = 1.12$, respectively[3]. For the Mejía-Restrepo et al. (2016) relations we use the values calibrated from the global fit to the continuum (the middle column of Table 7 in that paper); these are the most appropriate relations for our sample since PyQSOFit performs a global fit to the broadband AGN continuum, rather than a local fit under each separate emission line. The Woo et al. (2018) relation we adopt for Mg II was derived with free β and γ, which has least scatter with respect to fiducial black hole masses estimated using FWHM(H β) and L_{5100}. Greene et al. (2010) provide a formula to calculate $R_{\rm BLR}$ from the H β emission line luminosity; with an estimate of $R_{\rm BLR}$we then calculate the black hole mass using Eqn. (3.5.1) assuming $f = 1$.

3.5.3 Comparisons of black hole mass estimates

In Fig. 3.13 we compare the mass estimates obtained using the relationships given in Table 3.3. Generally, we find good agreement between the different estimates. For estimates using the continuum luminosity, we see excellent agreement between H α and H β masses; on average the difference between the mass estimates are negligible ($\Delta M_{\rm BH} \approx 5$ per cent) and the scatter is $\sigma = 0.22$ dex. There is slightly greater scatter between the H β and Mg II estimates ($\sigma = 0.26$), as has been noted in previous studies (e.g. Mejía-Restrepo et al. 2016). The average agreement is less good, also, with the Mg II masses being systematically higher than the H β masses by ≈ 60 per cent. The greatest discrepancy we see is between the two H α relations: masses calculated using the broad H α line luminosity are systematically lower

[3]This isn't stated explicitly in Woo et al. (2018), they say they use $f = 4.47$ and refer to the Appendix of Woo et al. (2015); here it says $f = 4.47$ is used for estimates using $\sigma_{\rm line}$ and $f = 1.12$ is used with FWHM$_{\rm line}$.

Table 3.4: Recalibrated virial black hole mass relations

FWHM	L	$\log(K)$	α	Median $\Delta\log(M_{\rm BH})$	σ	$N_{\rm spec}$
H α	H α	7.56 ± 0.04	0.47 ± 0.02	0.00	0.21	182
H α	5100 Å	6.91 ± 0.02	0.55 ± 0.03	0.02	0.22	182
H β	H β	7.83 ± 0.01	0.494 ± 0.007	−0.01	0.10	401
Mg II	Mg II	7.70 ± 0.05	0.46 ± 0.04	−0.03	0.29	224
Mg II	3000 Å	6.86 ± 0.03	0.51 ± 0.03	−0.02	0.26	224

The relations are of the form $\log(M_{\rm BH}) = \log(K) + \alpha\log(L) + 2\log(\mathrm{FWHM})$. Values of K and α have been determined which minimise the systematic offset and scatter with respect to masses estimated from the relation involving the FWHM of broad H β and the 5100 Å luminosity. The relations are calculated with the luminosity L in units 10^{44} erg s^{-1} and FWHM in units 1000 km s^{-1}. For each relation we have calculated we give the resulting systematic offset and dispersion with respect to the H β-5100 Å mass estimates. $N_{\rm spec}$ is the number of spectra used to establish the relation.

than those made with the continuum luminosity by almost a factor of two. Mass estimates using the broad H β line luminosity are also systematically lower than those involving the continuum luminosity, but the discrepancy is less than between the H α estimates (≈ 60 per cent). In the case of Mg II, there is a very minor systematic in the opposite sense with estimates using the continuum luminosity being ≈ 20 per cent smaller than those using the line luminosity.

To obtain better consistency between the various mass estimates, we have calculated new virial mass relations. We have assumed that the relation using the FWHM of broad H β and the 5100 Å luminosity provides the best and most reliable estimate of the mass. Then, assuming a virial mass relation of the form

$$\log\left(\frac{M_{\rm BH}}{\mathrm{M}_\odot}\right) = \log(K) + \alpha\log\left(\frac{L}{10^{44}\ \mathrm{erg\,s^{-1}}}\right) + 2\log\left(\frac{\mathrm{FWHM}}{1000\ \mathrm{km\,s^{-1}}}\right), \tag{3.5.2}$$

we determine the quantities K and α which reduce the systematic offset and scatter in the relations with respect to the H β-5100 Å mass estimates. These values are given in Table 3.4 along with the median $\Delta\log(M_{\rm BH})$ with respect to the H β-5100 Å mass estimates, and the dispersion σ. It can be seen that the systematic offsets between different relations are greatly reduced.

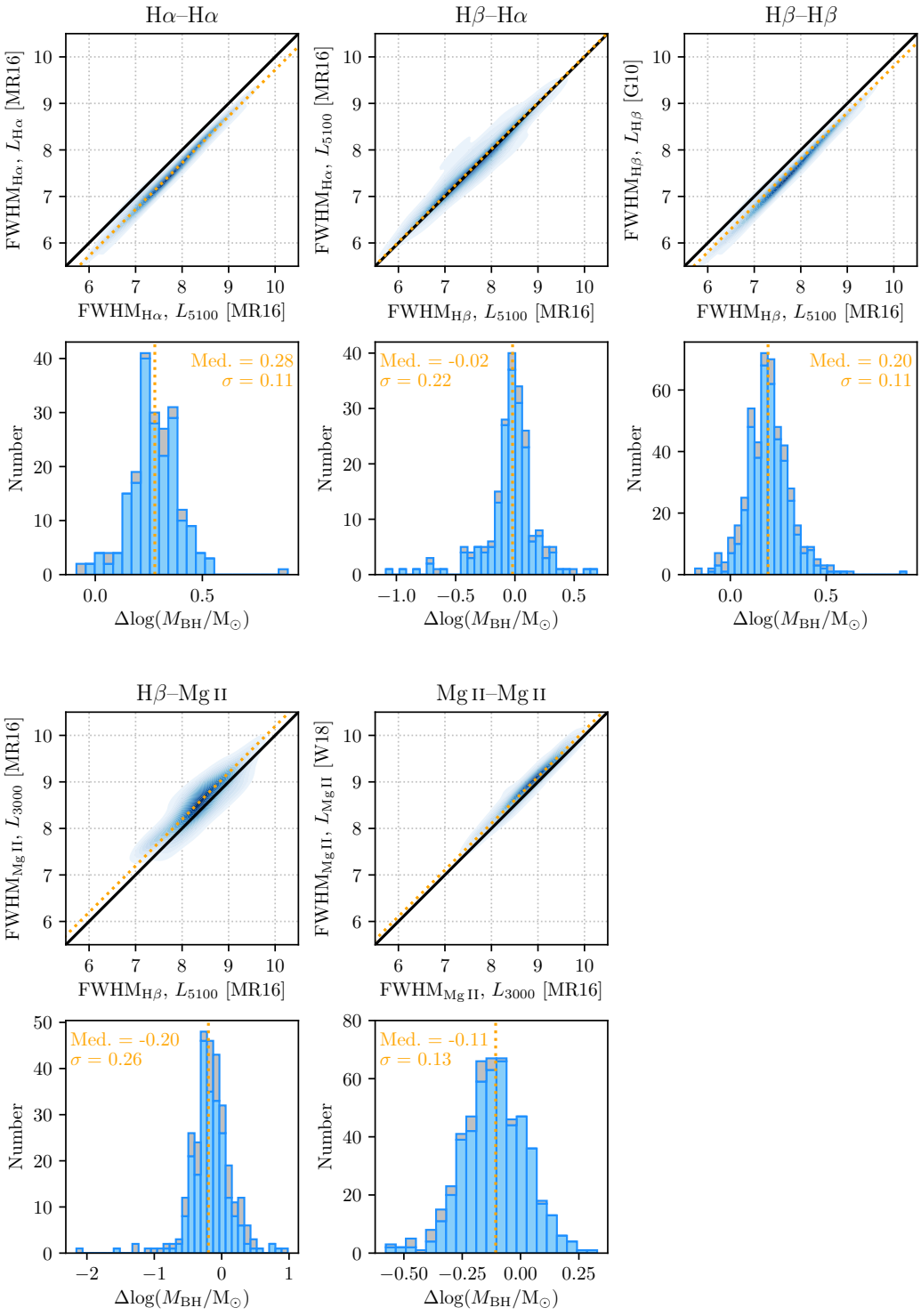

Figure 3.13: Comparisons of black hole mass estimates made using single-epoch relations taken from literature. The contour plots in the upper panels show $\log(M_{\rm BH}/{\rm M}_\odot)$, calculated using the broad emission line FWHM and the luminosity indicated on the axis label. The literature reference for the relation used is also given in square parentheses (see Table 3.3 in the text). The lower panels show the difference in $\log(M_{\rm BH}/{\rm M}_\odot)$. Reliable mass differences (with no quality flags raised on any of the quantities used to calculate the two masses) are shown in blue and the quality-flagged values are shown in grey. The median and standard deviations of the distributions for the reliable differences are given in the inset text. The orange dotted lines mark the median in both lower and upper panels.

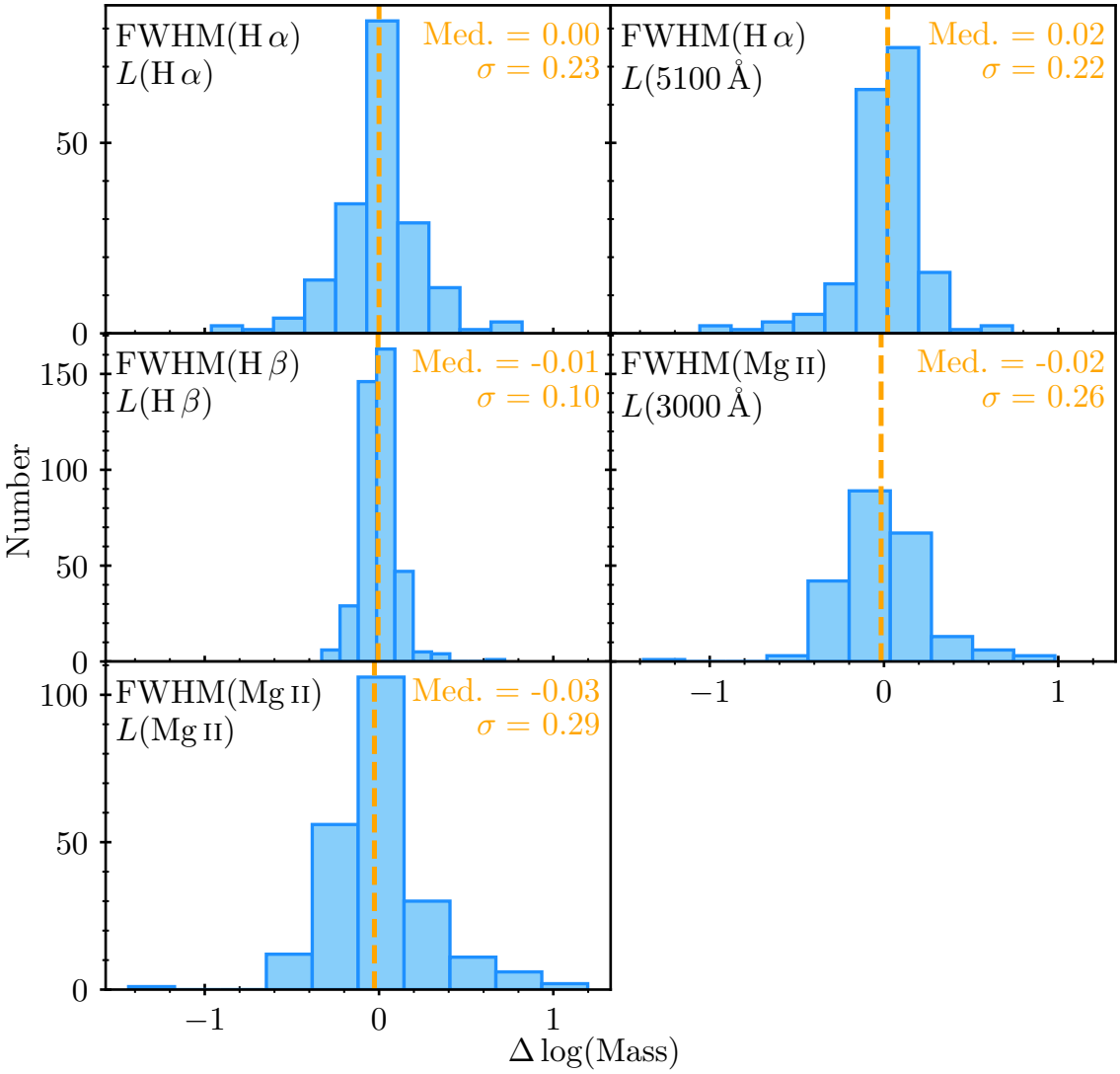

Figure 3.14: Comparisons of BH mass estimates from broad emission lines. $\Delta\log(\mathrm{Mass}) = \log\left(\mathrm{Mass}|^{\mathrm{FWHM(H\,\beta)}}_{L(5100)}\right) - \log\left(\mathrm{Mass}|^{\mathrm{FWHM}(x)}_{L(y)}\right)$.

3.6 α_{ox}: The UV/X-ray relationship

We calculate the UV/optical-X-ray energy index

$$\alpha_{ox} = -0.384 \log \left(\frac{L_{2\,keV}}{L_{2500}} \right) \tag{3.6.1}$$

for the AGN in our sample. An α_{ox} determination is possible for 673 AGN in our sample (97 per cent), with the remaining AGN having no spectroscopic or photometric coverage of 2500 Å.

We find a mean $\alpha_{ox} = 1.39$ with standard deviation $\sigma = 0.16$. These values are in good agreement with those found by Lusso et al. (2010) for a sample of quasars in the XMM-COSMOS field (mean 1.37 and dispersion 0.18). Lusso & Risaliti (2016) also show that there is a correlation of α_{ox} with L_{2500}. Fig. 3.15 (upper, left-hand panel) shows this for the 673 AGN in our sample. Plainly there is a good correlation of these quantities in our data. We use LINMIX, the Python version of the package LINMIX_ERR (Kelly, 2007), to perform a linear regression analysis. We find that

$$\alpha_{ox} = -(2.37 \pm 0.19) + (0.126 \pm 0.006) \times \log L_{2500}. \tag{3.6.2}$$

The dispersion about this line is 0.12 dex, showing that including the luminosity dependence does indeed reduce the scatter, as required if this is to be used as a cosmological probe (Lusso & Risaliti 2016). By contrast, there is much larger scatter of α_{ox} with the monochromatic X-ray luminosity, $L_{2\ keV}$. For these quantities we find

$$\alpha_{ox} = (0.56 \pm 0.27) + (0.032 \pm 0.010) \times \log L_{2\ keV}. \tag{3.6.3}$$

and the correlation is much weaker than that of the α_{ox}-L_{2500} relation and its significance is marginal (a little over 3σ).

3.6.1 α_{ox} related to emission line width and radio-loudness

In addition to values for the full sample, we have also calculated median α_{ox} and UV/X-ray luminosities for the narrow- and broad-line and radio-loud and radio-quiet subsets. We have also determined the luminosity-α_{ox} relations for these subpopulations. The resultant values are listed in Table 3.5. It is clear that the narrow-line AGN are typically lower in luminosity (both UV and X-ray) than the broad-line AGN. A greater difference is seen in the UV luminosities (median $\Delta L_{2500} \approx 1.0$ dex) than the X-ray luminosities (median

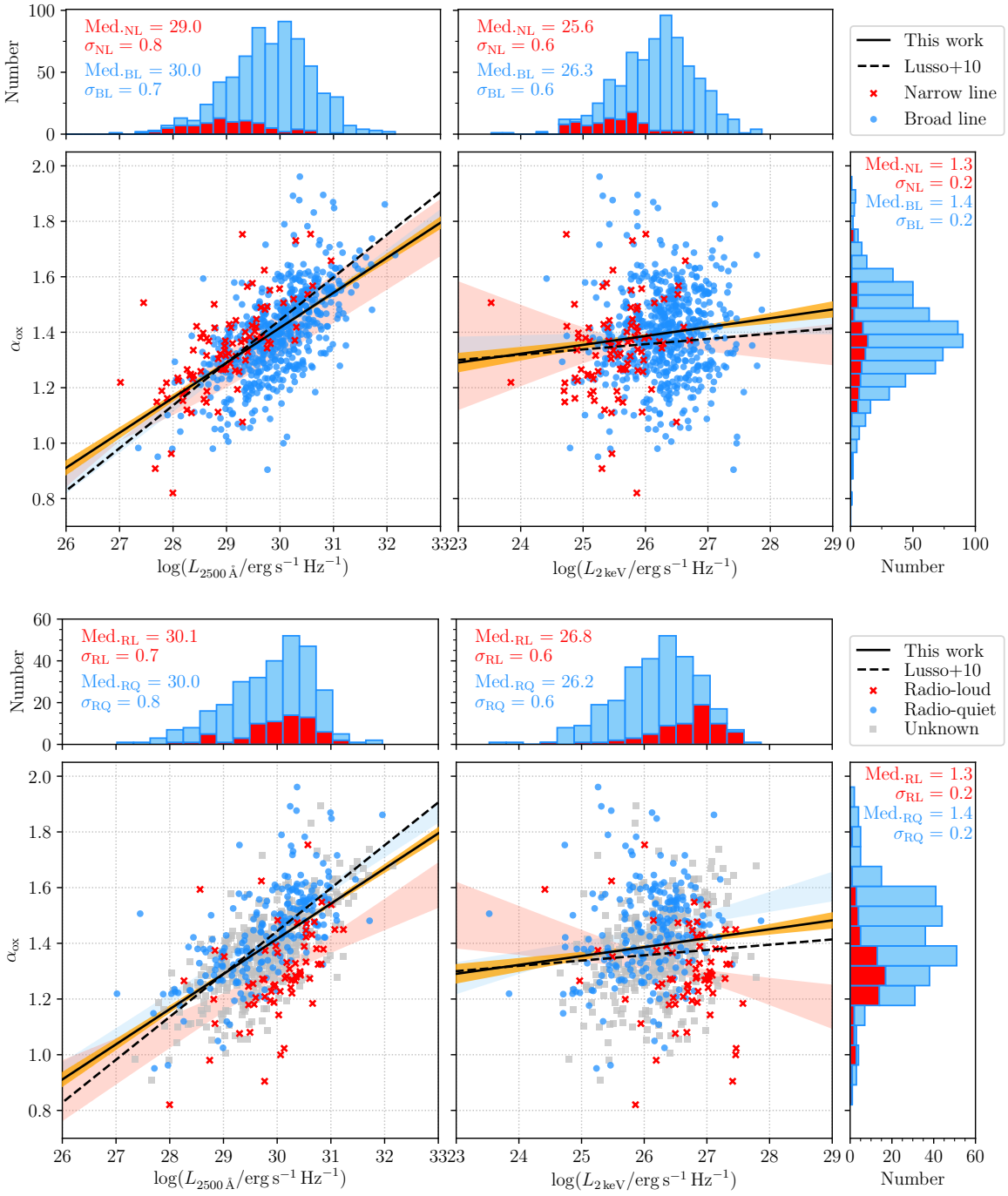

Figure 3.15: The relationship between α_{ox} and L_{2500} (left) and between α_{ox} and $L_{2\ keV}$ (right). We show the best fit relation found in this work for all 673 AGN with an α_{ox} determination (solid black line) and its 1σ confidence region (in orange). For comparison we show the best fit relation found by Lusso et al. (2010) (dashed black line). Stacked histograms show the distributions of the three quantities. In the top figure, narrow-line (FWHM $< 2000\ km\,s^{-1}$) and broad-line sources are coloured red and blue, respectively; the pale red and blue areas in the scatter plots are the linear regression 1σ confidence regions for the respective subsets. These uncertainties appear much smaller than the dispersion as although the sample dispersion and the slope uncertainty are related they are not the same thing. Despite the high scatter we are still able to determine the slope with decent precision, this is largely due to the small uncertainties on the individual data points themselves. In the bottom figure data pertaining to the radio-loud and radio-quiet sources are similarly coloured red and blue, respectively.

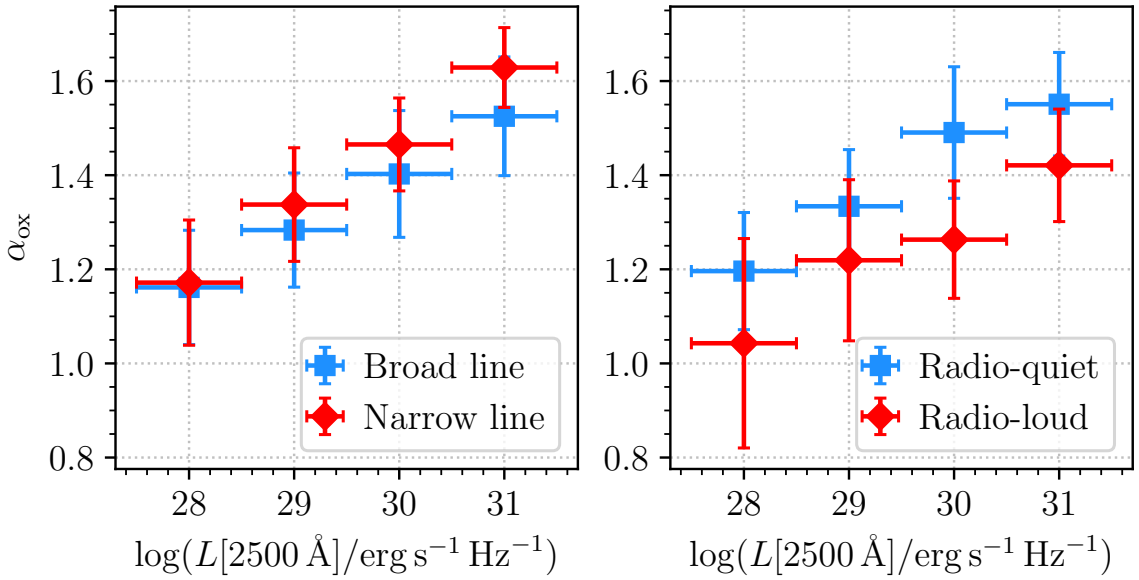

Figure 3.16: α_{ox} versus UV luminosity for (left) narrow-line / broad-line and (right) radio-loud / radio-quiet subsamples. In each luminosity bin, narrow-line sources have higher α_{ox} values (softer SEDs) than broad-line sources whereas radio-loud sources have lower α_{ox} values (harder SEDs) than radio-quiet sources.

$\Delta L_{2\,\text{keV}} \approx 0.7$ dex) between the narrow- and broad-line AGN. We therefore see a difference in the α_{ox} distributions of the subsets, with the narrow-line AGN having, on average, slightly lower α_{ox}. Fig. 3.16 shows that this is a consequence of the narrow-line sources having lower UV luminosities with $\alpha_{ox} \propto L_{2500}$. If the AGN are binned by L_{2500} then it can be seen that the α_{ox} values of the narrow-line AGN are systematically *higher* than those of the broad-line AGN across the luminosity bins, and progressively so with increasing L_{2500}. We do not observe any strong differences in the luminosity-α_{ox} relations for the narrow- and broad-line subsets.

The differences in luminosities between the radio-loud and radio-quiet populations are less marked than those seen in the narrow- and broad-line subsets. The radio-loud and radio-quiet AGN have similar UV luminosities, although the radio-loud subset is on average more luminous in X-rays (median $\Delta L_{2\,\text{keV}} \approx 0.6$ dex). The radio-loud AGN have, on average, lower values of α_{ox} than the radio-quiet AGN. Fig. 3.16 shows that the difference in α_{ox} between radio-loud and radio-quiet AGN holds across the different UV luminosity bins.

Table 3.5: Median, standard deviation and linear regression values from α_{ox} analysis.

Set	N	$M_{\alpha_{ox}}$	$\sigma_{\alpha_{ox}}$	$M_{\log L_{2500}}$	$\sigma_{\log L_{2500}}$	$M_{\log L_{2\,keV}}$	$\sigma_{\log L_{2\,keV}}$	$m(\log L_{2500})$	$c(\log L_{2500})$	$\sigma(\log L_{2500})$	$m(\log L_{2\,keV})$	$c(\log L_{2\,keV})$	$\sigma(\log L_{2\,keV})$
All	673	1.38	0.16	29.84	0.78	26.24	0.61	0.126 ± 0.006	-2.37 ± 0.19	0.12	0.032 ± 0.010	0.56 ± 0.27	0.15
NL	90	1.34	0.17	29.01	0.77	25.57	0.56	0.151 ± 0.017	-3.03 ± 0.49	0.12	0.056 ± 0.032	-0.09 ± 0.83	0.16
BL	583	1.39	0.16	29.97	0.71	26.32	0.55	0.137 ± 0.008	-2.71 ± 0.22	0.12	0.017 ± 0.012	0.96 ± 0.31	0.15
RL	67	1.28	0.16	30.15	0.67	26.82	0.60	0.105 ± 0.029	-1.86 ± 0.87	0.15	-0.056 ± 0.033	2.79 ± 0.89	0.16
RQ	233	1.45	0.17	30.00	0.84	26.16	0.63	0.133 ± 0.010	-2.53 ± 0.30	0.12	0.055 ± 0.018	0.01 ± 0.46	0.16

Median, standard deviation and linear regression values from α_{ox} analysis. For the whole sample of AGN with an α_{ox} determination (673 AGN) we list the median (M) and standard deviation (σ) of α_{ox}, the logarithms of the 2500 Å UV luminosity and 2 keV X-ray luminosity as well as the slope (m) and intercept (c) of the best-fitting linear regression of the luminosities with respect to the dependent variable α_{ox} and the dispersion (σ) of the data about the best fit. These quantities are also given for the narrow-line ('NL'), broad-line ('BL'), radio-loud ('RL') and radio-quiet ('RQ') subsets of the sample. N is the number of AGN in each set. Luminosities are in units $\mathrm{erg\,s^{-1}\,Hz^{-1}}$.

3.6.2 α_{ox} related to black hole mass and accretion rate

We also test the correlation between α_{ox} and black hole mass. There is substantial scatter but LINMIX determines a significant relation with mass:

$$\alpha_{ox} = (0.703 \pm 0.068) + (0.079 \pm 0.008) \times \log\left(M_{BH}/M_{\odot}\right). \quad (3.6.4)$$

The relation is shown in Fig.3.17.

We now investigate how much of the scatter in α_{ox}- L_{2500} can be explained by a two-parameter dependence on both mass and mass accretion rate (or equivalently, mass and L/L_{Edd}). The left panel of Fig. 3.18 shows the range of $\log L_{2500}$ (a tracer of mass accretion rate) and black hole mass spanned by our sample, binned onto a grid with two squares per decade. Within each grid square, we calculate the average α_{ox}, where this should be a more reliable estimate as we are using only AGN with approximately the same mass and mass accretion rate. We remove squares containing fewer than 3 AGN, so the dispersion is dominated by the intrinsic scatter. This shows a subtle but systematic trend in the data, where α_{ox} depends on both mass and L_{2500}, rather than L_{2500} alone.

For comparison, the right panel of Fig. 3.18 shows the predicted α_{ox} over this parameter space from the quasar SED model QSOSED (Kubota & Done 2018). This model was built from extrapolating trends seen in a much smaller sample of AGN than used here (Jin et al. 2012a,c). These earlier studies showed that optical-X-ray AGN SEDs are typically fit by three components: an outer standard disc, emitting thermal blackbody down to some radius, where it transitions to incomplete thermalisation (warm Comptonisation). This optically thick structure truncates below the radius R_{hot} so that the gravitational power within some $R_{hot} > R_{isco}$ (where R_{isco} is the radius of the innermost stable circular orbit) is dissipated instead in heating optically thin plasma which produces the power law tail by Comptonisation. The QSOSED model assumes that the total X-ray luminosity is fixed at $L_X = 0.02L_{Edd}$, i.e. the maximal ADAF limit. This captures the main trends seen in the data: at low mass accretion rates, only just above $L_{bol} = 0.02L_{Edd}$, then almost all of the accretion power is required to power the X-rays, and R_{hot} is large. Conversely, at high mass accretion rates, $L_{bol} \approx L_{Edd}$ then almost all of the accretion power is dissipated in the disc, and the X-rays at $L_X = 0.02L_{Edd}$ are weak in comparison and $R_{hot} \rightarrow R_{isco}$. The changing size of R_{hot} and changing luminosity of the disc also sets the spectral slope of the power-law tail self consistently, so that the model predicts the entire SED (with some additional assumptions about the warm Compton region).

We use this model, with spin fixed at 0 and the inclination at 30° to generate a set of SEDs from black hole masses in the range $\log(M_{\rm BH}/{\rm M}_\odot) = 5.75$–$9.75$ and adjust the mass-normalised accretion rate between $\dot{m} = 0.022 - 2.5$ (the limits of the QSOSED model) to match $\log L_{2500}$ to the midpoint of each grid square. We extract $\alpha_{\rm ox}$ from these model SEDs, and colour code each grid point using the same colour scale as from the real data. We also overplot lines of constant $L/L_{\rm Edd}$ on both panels. At high Eddington ratio, the UV is always dominated by the standard disc over this mass range, so the monochromatic flux $L_{2500} \propto (M_{\rm BH}\dot{M})^{2/3}$, giving rise to the straight line on the $\log(M_{\rm BH}/{\rm M}_\odot) - \log L_{2500}$ plot. Importantly this is not a linear relationship even for a standard disc. $L_{2500} \propto \dot{M}^{2/3} \propto L_{\rm bol}^{2/3}$ rather than being directly proportional to $L_{\rm bol}$. Thus using monochromatic flux (or flux in a narrow band) with a constant K-correction to convert to bolometric flux and hence derive Eddington ratio is subtly incorrect even for standard disc models. The solid grey line in Fig. 3.18 shows the Eddington luminosity assuming a bolometric correction of $L_{\rm bol} = 2.75 \times L_{2500}$ (Krawczyk et al. 2013).

The QSOSED model assumes the disc extends outwards only to the self-gravity radius, which may be only a few hundred $R_{\rm g}$ at the lowest accretion rates for the highest masses (see Chapter 4). This causes the very sharp drop in L_{2500} for these objects. If instead we set the outer disc radius to 10^5 $R_{\rm g}$ (as is usually done in the near-infrared spectroscopic modelling of AGN, e.g. Landt et al. 2011a), the $0.022L_{\rm Edd}$ curve more closely resembles those of higher Eddington ratios (the grey dot-dashed line in Fig. 3.18).

It is very clear that the data sit between the Eddington limits hardwired into the QSOSED code. The model assumes that below $L/L_{\rm Edd} \sim 0.02$, the flow is completely dominated by the optically thin, geometrically thick, hot plasma (ADAF/RIAF flow). This transition is seen directly in the 'changing-look' AGN (Noda & Done 2018, Ruan et al. 2019). Above $L/L_{\rm Edd} \sim 2.5$, the disc photosphere becomes brighter than its local Eddington flux (Kubota & Done 2019), so its structure is probably modified by strong winds and/or radial advection of the excess radiation. Both these probably lead to vertical structure so the radiation escaping preferentially only in a narrow funnel, making these objects rare (as well as probably short lived).

Our data also show that high Eddington fraction flows are rare at the highest black hole masses in our sample. These objects are very rare in the local Universe, and are only likely to be present at the peak of the quasar epoch at $z \sim 2 - 4$, beyond our sample redshift limit. Conversely, we only see the lowest mass black holes ($M \sim 10^6 M_\odot$) at the highest Eddington

ratios as these are otherwise diluted by their host galaxy (e.g. Done et al. 2012).

In Fig. 3.19 we show the difference between α_{ox} as predicted by QSOSED and as determined in our sample. This shows that there is generally reasonable agreement between the model predictions and α_{ox} estimated from our data, In the majority of bins (97 per cent) $|\Delta\alpha_{ox}| < 0.2$ and in 63 per cent of the bins $|\Delta\alpha_{ox}| < 0.1$. A detailed exploration of the differences between the data and the model predictions as a function of black hole mass, luminosity and accretion rate will be performed in future work (Chapter 4).

3.7 Discussion

The SOUX sample is a medium-sized selection of ≈ 700 AGN up to $z = 2.5$ with quality multiwavelength data. Although it is not a complete or representative AGN sample, it has been assembled with the intention of exploring a wide range of parameter space. This first study has focussed on making key measurements from the data (the broad optical emission line properties, multiwavelength luminosities etc.) and on determining a consistent set of black hole masses from which we will be able to determine Eddington ratios. We have taken a preliminary look at changes in the shape of the AGN SED (characterised by the parameter α_{ox}) as a function of black hole mass and AGN luminosity. To first order, α_{ox} is found to vary with both mass and luminosity as broadly expected from a theoretical model. However, the systematic differences seen at second-order will require a deeper exploration. Our main findings are discussed below.

3.7.1 Broad emission line properties

In Section 3.5.1 we showed that the luminosities of the three broad emission lines H α, H β and Mg II are strongly correlated with the AGN continuum luminosities. Of all emission line-continuum relations we tested, the least scatter is found between $L_{H\,\beta}$ and L_{5100}, although the relationship between the two continuum luminosities L_{5100} and L_{3000} had least scatter overall. Whilst the luminosities of the two Balmer lines are approximately linearly proportional to L_{5100}, the $\log(L_{Mg\,II})$-$\log(L_{5100})$ relation has a gradient noticeably lower than 1 ($m = 0.87 \pm 0.03$). These results are similar to those reported by Shen & Liu (2012), who also found least scatter between L_{5100} and L_{3000}. Shen & Liu (2012) report a $\log(L_{H\,\alpha})$-$\log(L_{5100})$ relation with a gradient consistent with 1 and a $\log(L_{Mg\,II})$-$\log(L_{5100})$ proportionality consistent with the one we find here (0.86 ± 0.07). However, in their work they find a much steeper relation

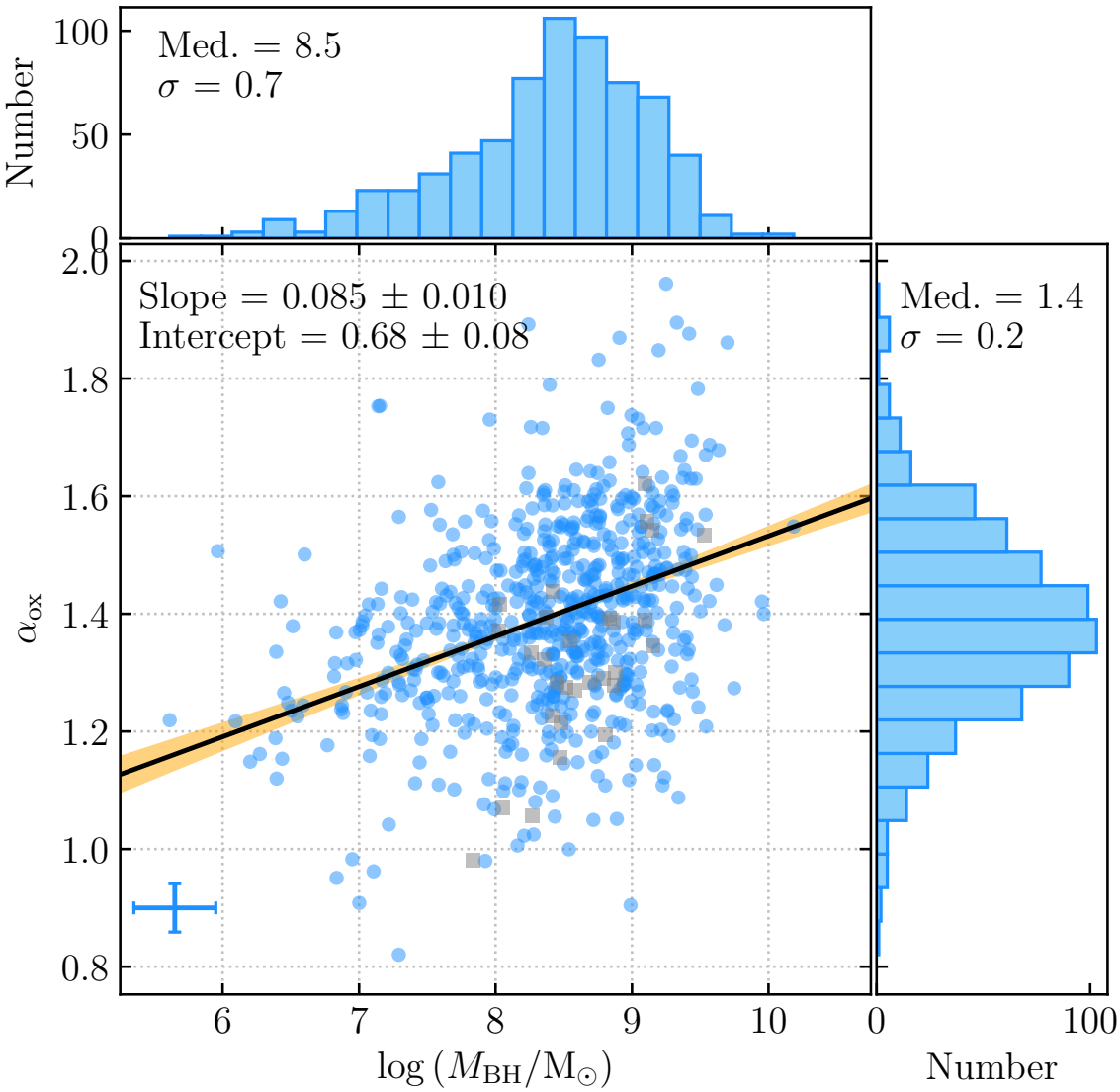

Figure 3.17: The relationship between α_{ox} and the logarithm of black hole mass. We show the best fit relation found (solid black line) and its one sigma confidence region (in orange). Histograms show the distributions of the two quantities. A representative error bar is shown in the bottom left.

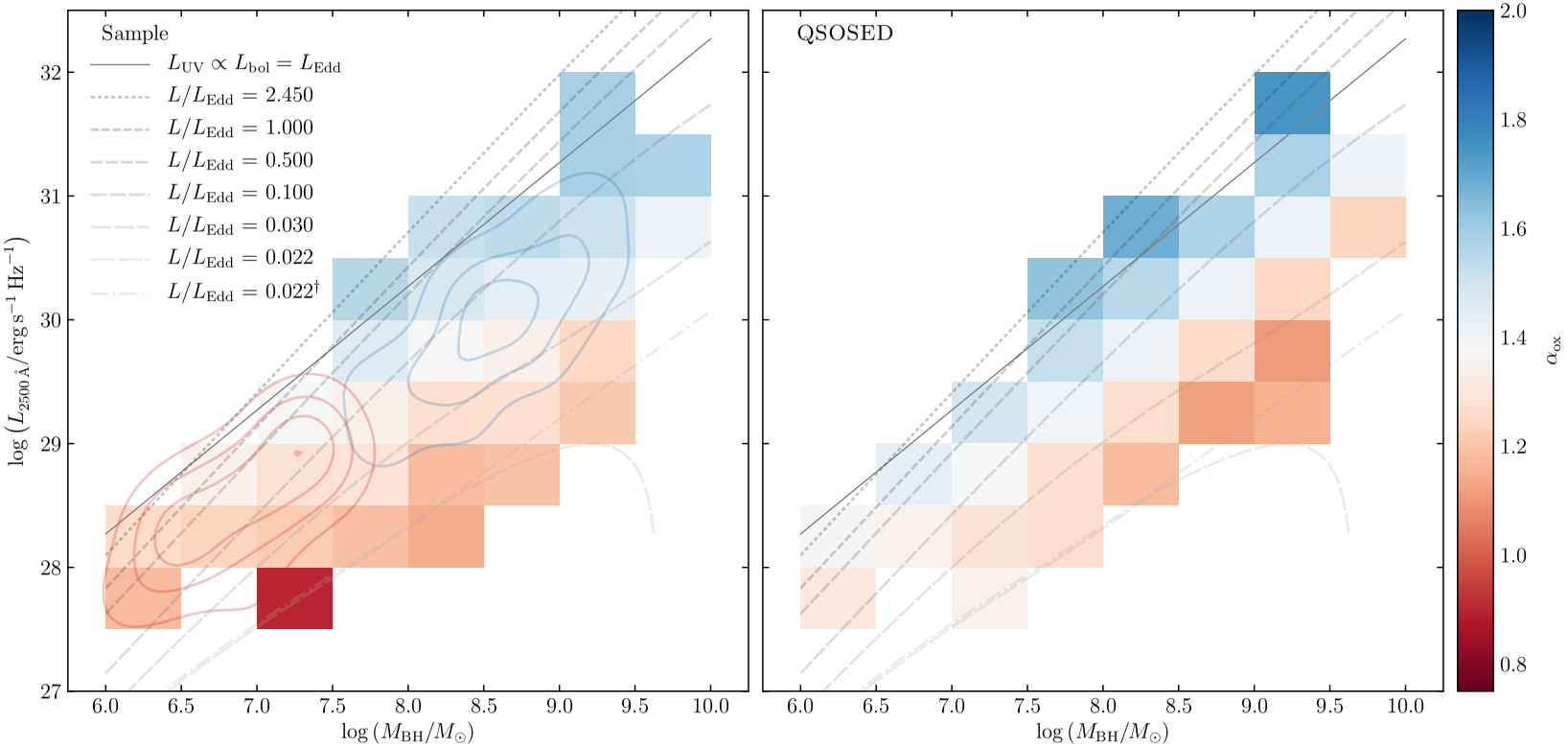

Figure 3.18: The relationship between L_{2500}, black hole mass and $\alpha_{\rm ox}$. The left panel shows the average values of $\alpha_{\rm ox}$ for sources in our sample binned by mass and UV luminosity (only bins containing three or more AGN are drawn). The right panel shows the value of $\alpha_{\rm ox}$ taken from an SED generated by the Kubota & Done (2018) model QSOSED, for mass and UV luminosity at the centre of each grid point. Dashed grey lines correspond to curves of constant $L/L_{\rm Edd}$ as determined by QSOSED. †The dot-dashed grey line is computed for a model in which the outer disc radius $\log(R_{\rm out}/R_{\rm g}) = 5$, whereas in all other models the outer radius is the disc self-gravity radius. The solid line shows a curve of constant $L/L_{\rm Edd} = 1$ assuming that $L_{2500} \propto L_{\rm bol} \propto M_{\rm BH}$; clearly this relation is shallower than that predicted by QSOSED in which $L_{2500} \propto M_{\rm BH}^{4/3}$. Blue and red contours in the left-hand panel indicate the parameter space spanned by AGN drawn from the SDSS quasar catalogue and the supplementary sources added from the R17 NLS1 catalogue, respectively.

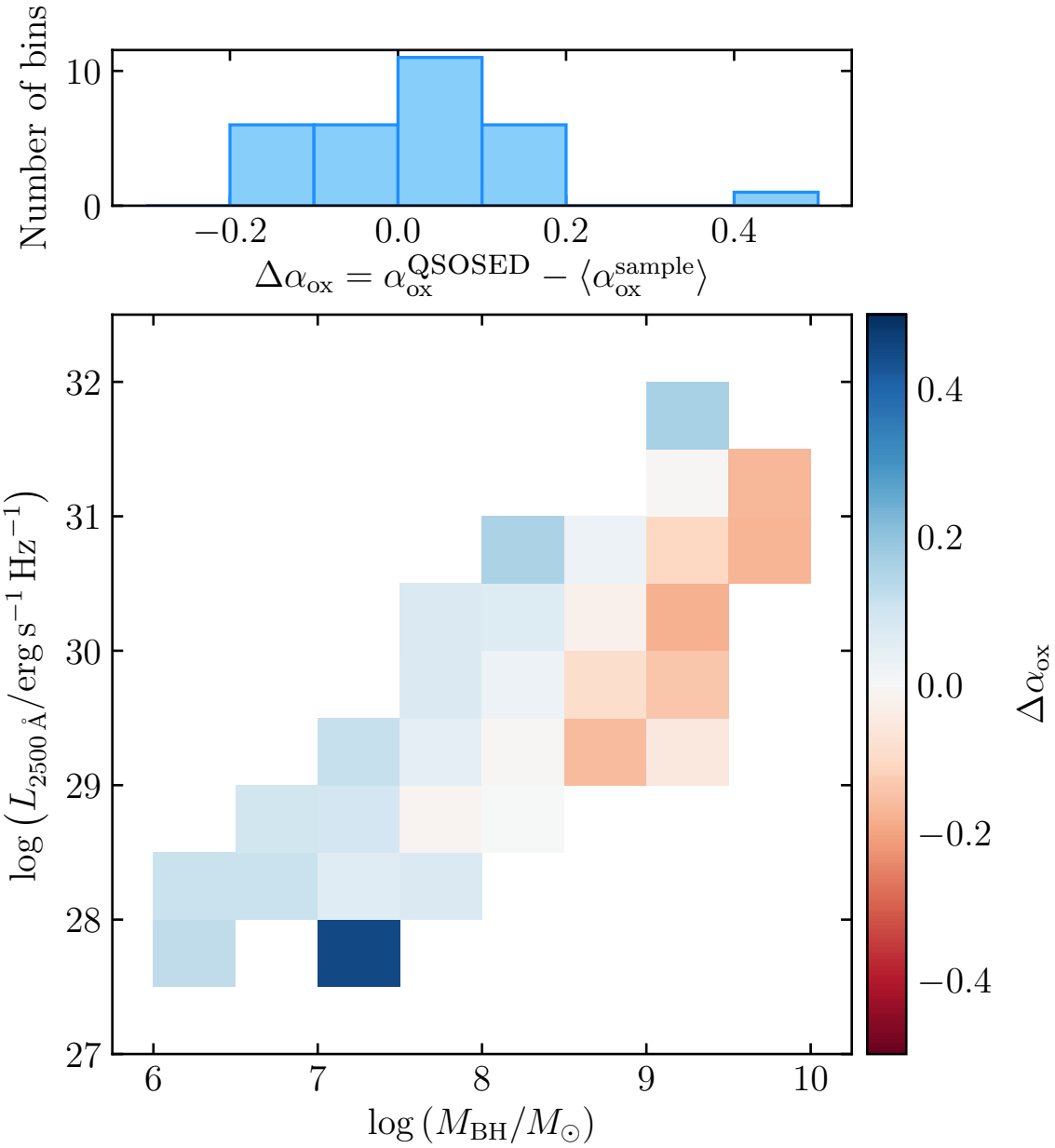

Figure 3.19: The difference between $\alpha_{\rm ox}$ as predicted by the QSOSED model and as found from the data in our sample. The bottom panel shows the difference in each mass-luminosity bin, the top panel shows the distribution of $\Delta\alpha_{\rm ox}$ values.

between $\log(L_{\mathrm{H}\,\beta})$-$\log(L_{5100})$ (1.25 ± 0.07). Collinson et al. (2017) also find a shallower gradient between $\log(L_{\mathrm{Mg\,II}})$ and $\log(L_{5100})$ than between the Balmer lines and $\log(L_{5100})$. In their work, the gradients of all three emission line luminosites with respect to $\log(L_{5100})$ were shallower than 1, although they were working with a much smaller sample (11 AGN) and much more limited range in luminosities (~ 1 dex). It has been shown both in observations of reverberation-mapped and 'changing-look' AGN (e.g. Sun et al. 2015; Zhu et al. 2017; Kynoch et al. 2019a) and photoionisation models (e.g. Guo et al. 2020) that Mg II responds more weakly to changes in the continuum than does H β. Here, we appear to see a similar trend in a statistical sense across several hundred snapshots of individual AGN. Overall, the strong correlations between the emission line and continuum luminosities demonstrate that the former are also reasonable proxies for the BLR radius, which we discuss below.

3.7.2 Relations for virial black hole masses

In Section 3.5 we computed the black hole masses of the AGN in our sample from the broad emission line widths and either a nearby continuum luminosity or the broad emission line luminosity. We used several single-epoch virial black hole mass relations taken from recent literature (Mejía-Restrepo et al. 2016; Woo et al. 2018; Greene et al. 2010). Mass estimates from the FWHM of the Balmer lines and L_{5100} are generally in very good agreement, with a minor systematic difference $|\Delta\log(M_{\mathrm{BH}})| = 0.2$ (i.e. a 5 per cent difference in M_{BH}) and $\sigma = 0.22$ dex.

We noticed that estimates made using the emission line luminosities were systematically lower than those made using the continuum luminosities. The masses were ≈ 60 per cent lower in the case of H β and a factor two lower in the case of H α. Discrepancies of a similar magnitude have been observed in previous studies calculating the black hole mass of single objects (e.g. Landt et al. 2017; Kynoch et al. 2019b) and also in large samples (e.g. Shen & Liu 2012). Shen & Liu (2012) compared masses calculated with FWHM(H α)-$L_{\mathrm{H}\,\alpha}$ (Shen et al., 2011) and FWHM(H β)-$L_{\mathrm{H}\,\beta}$ (Greene & Ho, 2005) with fiducial masses calculated with FWHM(H β)-L_{5100} (Vestergaard & Peterson, 2006) and again estimates using the line luminosities were lower, although to a lesser extent than we find here (Shen & Liu 2012 report $\langle \Delta \log[M_{\mathrm{BH}}] \rangle = -0.08$ and -0.04 dex for H α and H β, respectively). The cause of these systematic discrepancies is not apparent. Since, as we noted above, there is generally very good correspondence between the emission line and continuum luminosities, in principle it should be possible to determine a reliable mass estimate using either luminosity as a proxy for

the BLR radius. It is not simply the case that we have taken scaling relations from different papers in which the relations were derived for different samples of AGN: both relations involving H α are taken from Mejía-Restrepo et al. (2016). Such large systematic difference in mass are difficult to explain since by Eqn. (3.5.1) $M_{\rm BH} \propto L^{1/2}$, a factor four difference in luminosity is required to reconcile a factor two difference in mass. It is very unlikely that the 5100 Å luminosity is contaminated by stellar light from the AGN host galaxies, resulting in masses determined using L_{5100} to be overestimates. PyQSOFit performs a subtraction of the host galaxy spectrum where necessary, so in principle the stellar contamination has been removed from our measures. Also, we see a systematic offset even for the highest-mass sources: highly luminous quasars with optical luminosities far in excess of their host galaxies. It is possible that the procedure we used to fit the broad emission lines to determine their width and luminosity could differ from those used to determine the $M_{\rm BH}$ relations (Mejía-Restrepo et al. 2016 used only two Gaussians for the Balmer lines, for example). As described in Section 3.3.3, in this work the broad emission lines were fit with three Gaussians. In many cases, the total line profile was decomposed into broad, very broad an intermediate Gaussians. The shallow but very broad component (of H α in particular) which was often present can add a substantial amount of flux to the line without greatly increasing the FWHM. However, this would have the effect of increasing the emission line luminosities, whereas what we find is that the Balmer lines are apparently *underluminous* compared to the continuum for a given FWHM, resulting in the lower mass estimates. Differences in the assumed virial factor f will also result in equivalently larger or smaller estimates of the mass. However, the relations we test here use $f \approx 1$ so the mass differences ought to be very minor.

To avoid problems introduced by systematic differences in black hole masses, we have calculated new mass scaling relations tailored to our sample. As described in Section 3.5.2 these minimise the discrepancy with respect to estimates made from FWHM(H β) and L_{5100} (which we judge is the most reliable estimate). The new scaling relations are presented in Table 3.4. We note here that all of our relations have a slope α consistent with 0.5, the expected value for an inverse-square scaling of radius with luminosity. For additional quality assurance, we visually inspected all of the optical spectra of the AGN in our sample and selected the best broad line from which to estimate the black hole mass, based on the line profile shape and S/N in the vicinity of the line. For each AGN in our sample we list the 'preferred' mass estimate indicating our choice of broad emission line and luminosity scaling relation for that source. We have adopted, wherever possible, the black hole mass from estimates using FWHM(H β) and L_{5100}: the majority of the AGN in our sample (52 per cent) have their

adopted masses derived using this relation from Mejía-Restrepo et al. (2016). 45 per cent of the sample have their preferred mass estimate taken from our FWHM(Mg II) and L_{3000} relation; these were mostly the high-z AGN in our sample, lacking optical spectral coverage of H β. Our newly-derived FWHM(Mg II)-L_{3000} mass relation has a very minor systematic offset with respect to the fiducial masses calculated via the FWHM(H β)-L_{5100} relation of Mejía-Restrepo et al. (2016): the median difference in mass is negligible, although there is substantial scatter ($\sigma = 0.26$ dex). Following these recalibrations, we do not expect that using different mass relations for different AGN will have an adverse impact on our results.

3.7.3 $\alpha_{\rm ox}$

In Section 3.6 we explored the relationship between the UV-X-ray energy index $\alpha_{\rm ox}$ with the UV and X-ray luminosities and with black hole masses and accretion rates. We found a highly significant correlation of $\alpha_{\rm ox}$ with the monochromatic UV luminosity $\log L_{2500}$ for all AGN as well as the four subsets (narrow- and broad-line; radio loud and quiet) taken separately. The correlation between $\alpha_{\rm ox}$ and $\log(L_{2\,\rm keV})$ was very much weaker, but marginally significant at $> 3\sigma$. This appears to be driven by the inclusion of AGN from the R17 NLS1 catalogue; in a previous study of quasars, Lusso et al. (2010) found no significant relation between $\alpha_{\rm ox}$ and $\log(L_{2\,\rm keV})$. We did not find any highly significant correlations between $\alpha_{\rm ox}$ and $\log(L_{2\,\rm keV})$ for the full sample or any of the subsamples taken individually,

In Section 3.6.2 we calculated the mean $\alpha_{\rm ox}$ for 30 subsets of AGN, binned in both $\log(M_{\rm BH})$ and $\log(L_{2500})$, ranging over four orders of magnitude in each quantity (Fig. 3.18). Fig. 3.18 illustrates the diversity of our sample - blue and red contours show the location in parameter space of AGN drawn from the quasar and NLS1 catalogues, respectively. The inclusion of AGN from the NLS1 catalogue extends our sample to, on average, lower black hole masses and higher accretion rates than probed by the quasars. This wide coverage of parameter space allows us to track changes in the spectral shape (characterised by $\alpha_{\rm ox}$) with both mass and accretion rate. The narrow-line AGN have slightly higher $\alpha_{\rm ox}$ compared with broad-line AGN in the same L_{2500} bin, likely as a result of the stronger soft X-ray excess (warm corona) component in the former. We also found differences between our radio-loud and radio-quiet AGN: the radio-loud AGN are on average more X-ray luminous than the radio-quiet sources although both sets have similar UV luminosities (Fig. 3.18). Even when binning by L_{2500}, radio-loud AGN have systematically lower $\alpha_{\rm ox}$ (harder SEDs) than the radio-quiet AGN. Often this is attributed to the radio jet contributing to the hard X-ray band, resulting in an

increased X-ray flux and lower UV/X-ray ratio. However, Zhu et al. (2020) argue that for steep radio spectrum sources the excess X-ray luminosity of radio-loud AGN is likely *not* due to contamination of the X-ray emission by the jet, but rather some evolution of the disc-corona configuration. In their study of radio-loud and radio-quiet quasars, Zhu et al. (2020) proposed that the best-fitting relation between $\alpha_{\rm ox}$ and $\log L_{2500}$ for radio-quiet quasars represented an approximate 'jet line' for AGN, in analogy to the jet line of accreting stellar-mass black holes (e.g. Fender et al. 2004): jets are quenched for AGN on the soft-SED side of the line and the most powerful jets are found in sources with the hardest SEDs for a given L_{2500}. Qualitatively, we see a similar trend in our data: radio-loud AGN are systematically offset to lower $\alpha_{\rm ox}$ with respect to the radio-quiet sources, and very few radio-loud AGN are found to have higher $\alpha_{\rm ox}$ than the average for radio-quiet AGN at a given $\log L_{2500}$. The best-fitting $\alpha_{\rm ox}$-$\log L_{2500}$ relation for radio-quiet quasars found by Zhu et al. 2020 is entirely consistent with our relation for radio-quiet AGN (Table 3.5). We compared the mean $\alpha_{\rm ox}$ for each bin to the prediction made from the AGN SED model QSOSED. The correspondence between the model predictions and our sample data is remarkably good: nearly two thirds of bins have $|\Delta\alpha_{\rm ox}| < 0.1$ and in all but one bin $|\Delta\alpha_{\rm ox}| < 0.2$. Fig 3.19 shows that the largest differences between the data and the model are at the extremes of the parameter space: for high-mass/high-luminosity AGN, the model predicts SEDs which are too red and for low-mass/low-luminosity AGN the predictions are too blue. Of course, if there is a systematic error in the calculation of the black hole masses then the QSOSED predictions are not being compared with the appropriate AGN (however, as described previously, we have made efforts to ensure our mass estimates are reliable). Alternatively, whilst QSOSED makes good predictions for moderate masses and luminosities, it is possible that the actual disc-corona coupling is more complex than assumed by the model. We will explore these issues in more detail by modelling the full broadband SEDs of our sample (Chapter 4).

3.8 Summary and conclusions

We have assembled a sample of 696 AGN with SDSS optical spectra which also have X-ray spectra and simultaneous UV photometry recorded by *XMM-Newton*. Although this is not a complete or representative sample of AGN, the objects in our sample cover a wide range in black hole mass, accretion rate, spectral hardness and redshift. The majority of sources in our sample are luminous AGN found in the SDSS-DR14 quasar catalog; these are supplemented selection of lower-luminosity, narrow-line type 1 AGN in the local Universe (taken from R17).

We have performed new optical spectral fits so that we have comparable measurements of the AGN continuum and at least one broad emission line (H α, H β, or Mg II) for the narrow-line AGN (and broad-line AGN that lacked a measurement of the Mg II spectral region). We computed single-epoch virial black hole masses for the AGN in our sample using these broad emission lines with both the local continuum luminosity and the emission line luminosity. However, we found some large systematic differences between mass estimated made using different relations (masses calculated with FWHM[H α] and $L_{\mathrm{H}\,\alpha}$ are on average a factor two lower than those made using L_{5100}). We have calculated new virial black hole mass relations which minimise the systematic offsets with respect to fiducial black hole masses calculated using FWHM(H β) and L_{5100}.

From the X-ray catalogue fluxes, we have estimated rest-frame 2 keV luminosities and X-ray photon indices for the full sample. Combining these with out UV measurements, we have calculated the energy index α_{ox}, an indicator of the broadband SED shape. We found that narrow-line sources have steeper X-ray spectra (higher Γ) and, when controlling for L_{UV}, have softer broad-band SEDs (higher α_{ox}) than typical broad-line AGN. Differences were also found with respect to radio properties with radio-loud AGN being on average more X-ray luminous and having harder SEDs (lower α_{ox}) than the radio-quiet subset. Fig. 3.18 illustrates that the majority of the AGN in our sample lie between $0.02 < L/L_{\mathrm{Edd}} < 2$, with narrow-line type 1 offset to lower masses and higher accretion rates compared with the broad-line quasars. We have shown that the model QSOSED makes very good predictions of α_{ox} for a given black hole mass and accretion rate. However, there are systematic differences between the data and model predictions at the extremes of the mass-accretion rate parameter space. In Chapter 4 we will model the broad-band SEDs, testing the physical assumptions made in the QSOSED model and further explore the relationships between the UV/optical and X-ray spectra.

Chapter 4

The SOUX sample: Optical/UV/X-ray SEDs and the Nature of the Disc

4.1 Introduction

Active Galactic Nuclei (AGN) are powered by mass accretion onto a Supermassive Black Hole (SMBH). They emit radiation over a large swathe of the electromagnetic spectrum. This emission is often represented in the form of a spectral energy distribution (SED), which shows the power emitted as a function of frequency. Therefore, SEDs can be used as a powerful diagnostic tool, allowing us to probe the physical structures and emission mechanisms in these objects. These show clearly that the AGN emission cannot be solely explained by a standard (Shakura & Sunyaev, 1973) disc model. The reality is much more complex.

AGN span a very wide range in mass (from 10^{5-10} $M_{\odot}$) and luminosity, and display distinctly different SEDs across this parameter space. The effect of orientation along our line of sight with respect to an equatorial obscurer (i.e. the dusty torus) is at the core of the 'unified model' (Antonucci, 1993). Nonetheless, whilst orientation is surely a key parameter, the systematic differences in SED shape, seen with changing mass and luminosity (e.g. Vasudevan & Fabian 2007; Jin et al. 2012a; Lusso & Risaliti 2016), strongly indicate an intrinsic change in the broad band SED as a function of both mass and/or luminosity.

Intrinsic changes in the SED are well documented in Black Hole Binary (BHB) systems (Done et al., 2007). There is a strong spectral transition which occurs at $\sim 0.02 L_{\rm Edd}$ as the source

slowly dims down from a soft, thermal state to a hard, Comptonised state. These different emission mechanisms most likely signal a fundamental change in the nature of the accretion flow, from a geometrically thin, cool, optically thick disc similar to the standard disc models (Shakura & Sunyaev, 1973) to an optically thin, hot, geometrically thick flow such as the Advection Dominated accretion flow models (ADAFs), (Narayan & Yi, 1995).

A single model for AGN would then take this state transition in BHB and scale it to the higher mass SMBH. Indeed, a strong spectral change is seen in some AGN which vary across the transition luminosity of $\sim 0.02 L_{\rm Edd}$ (Noda & Done, 2018; Krumpe et al., 2017; Ruan et al., 2019). These 'changing-look' AGN show hard X-ray spectra below a few percent of Eddington which are dominated by the hard X-ray power law, similar to the hard state in BHB. However, above a few percent of Eddington, i.e. when the BHB show predominantly disc dominated spectra, the AGN instead show spectra with a substantial amount of non-disc emission. This is shown explicitly in Kubota & Done (2018) (hereafter KD18), who used AGN of similar mass around $10^8 M_{\odot}$ but changing $L_{\rm bol}$, (NGC5548, Mrk509 and PG1115+407) to demonstrate the systematic change in SED shape for $L/L_{\rm Edd} \sim 0.03 - 0.5$. These AGN are all above the transition value, but at the lower end of the luminosity range they show spectra which have strong hard X-ray emission, with similar power to that seen in the UV, while the spectra become systematically more UV disc dominated with increasing $L/L_{\rm Edd}$.

The amount of disc to X-ray luminosity is often characterised by $\alpha_{\rm ox}$, the spectral index of a power law connecting from the UV (2500Å) to the X-ray band at 2 keV. This index becomes progressively more negative with increasing $L_{\rm bol}$, indicating that the spectra become more disc dominated (Lusso et al., 2010; Lusso & Risaliti, 2016). However, given the range of SMBH masses, it is not clear whether this trend (which does have scatter) is driven by $L_{\rm bol}$ or by $L_{\rm bol}/L_{\rm Edd}$ or a separate factor. Understanding the origin of this correlation would reduce the scatter, and give more accuracy as well as more confidence in its use as a cosmological probe (Lusso et al., 2018).

Another key difference between the BHB and SMBH at luminosities above the transition is the appearance of an additional component in the spectrum, between the disc and X-ray tail. There is a downturn in the UV which appears to connect to an upturn at soft X-ray energies above the 2-10 keV power law (soft X-ray excess). The origin and nature of this are not well understood, but it can be fairly well fit as a warm Comptonisation component, in addition to a separate hot Comptonisation component producing the X-ray tail (Porquet et al., 2004; Gierliński & Done, 2004; Jin et al., 2012a).

These differences show that there must be something that breaks the scaling between the SMBH and BHB above $0.02L_{\rm Edd}$. One key theoretical difference is that the disc temperature in AGN is lower, $T_{\rm peak} \propto (\dot{m}/M)^{1/4}$, where $\dot{m}{=}L/L_{\rm Edd}$, at the innermost stable circular orbit, giving a predicted peak luminosity in the UV for AGN rather than in the soft X-rays as seen in BHB. This means that atomic physics should be very important in AGN discs, whereas the soft X-ray BHB discs are mostly dominated by plasma physics. This change in opacity could drive turbulence/convection, or even mass loss via UV line driven disk winds (e.g. Laor & Davis 2014).

Another expected theoretical difference is that radiation pressure is much more important in SMBH discs. The typical density of the SMBH disc is lower as well as its temperature, so the gas pressure in the disc is lower for a given temperature. So radiation pressure within the disc can dominate over gas pressure inside the disc over a much wider radial range compared to the BHB at similar Eddington fraction, $\dot{m}$ (Laor & Netzer, 1989). This again could lead to turbulence/convection, and all hydrodynamic turbulence couples to the MRI dynamo which is the source of the viscosity, enhancing the heating towards the disc surface and potentially producing the warm Comptonisation region (Jiang & Blaes 2020).

KD18 built a phenomenological model (QSOSED) to describe the changing SED in their very small sample of very well studied AGN of fixed mass ($M_{\rm BH} \sim 10^8 M_{\odot}$). This model is based on the expected Novikov-Thorne heating rate from a disc at a given $M_{\rm BH}$ and $L/L_{\rm Edd}$, but incorporates a phenomenological prescription for how the energy is emitted, either as a standard disc (outer radii), warm Comptonisation (mid radii) or hot corona (inner flow). Kynoch et al. (2022) (hereafter K22) have assembled a much larger sample of AGN with good quality spectral data and thus fairly well defined SEDs spanning the optical/UV and X-ray bandpass where most of the accretion energy should be emitted (detailed in Section 4.2.1). We use this new sample to critically test the QSOSED model across a wide range in mass and $L/L_{\rm Edd}$ with the aim of understanding and characterising the accretion disc structures in AGN.

4.1.1 This Chapter

We first present an overview of the SOUX sample as defined in K22 along with a description of the models used throughout this work. We perform stacked fitting on the SOUX sample, binning on $M_{\rm BH}$ and L_{2500}.

We first use the sample to test the assumptions made by QSOSED by fitting only to the

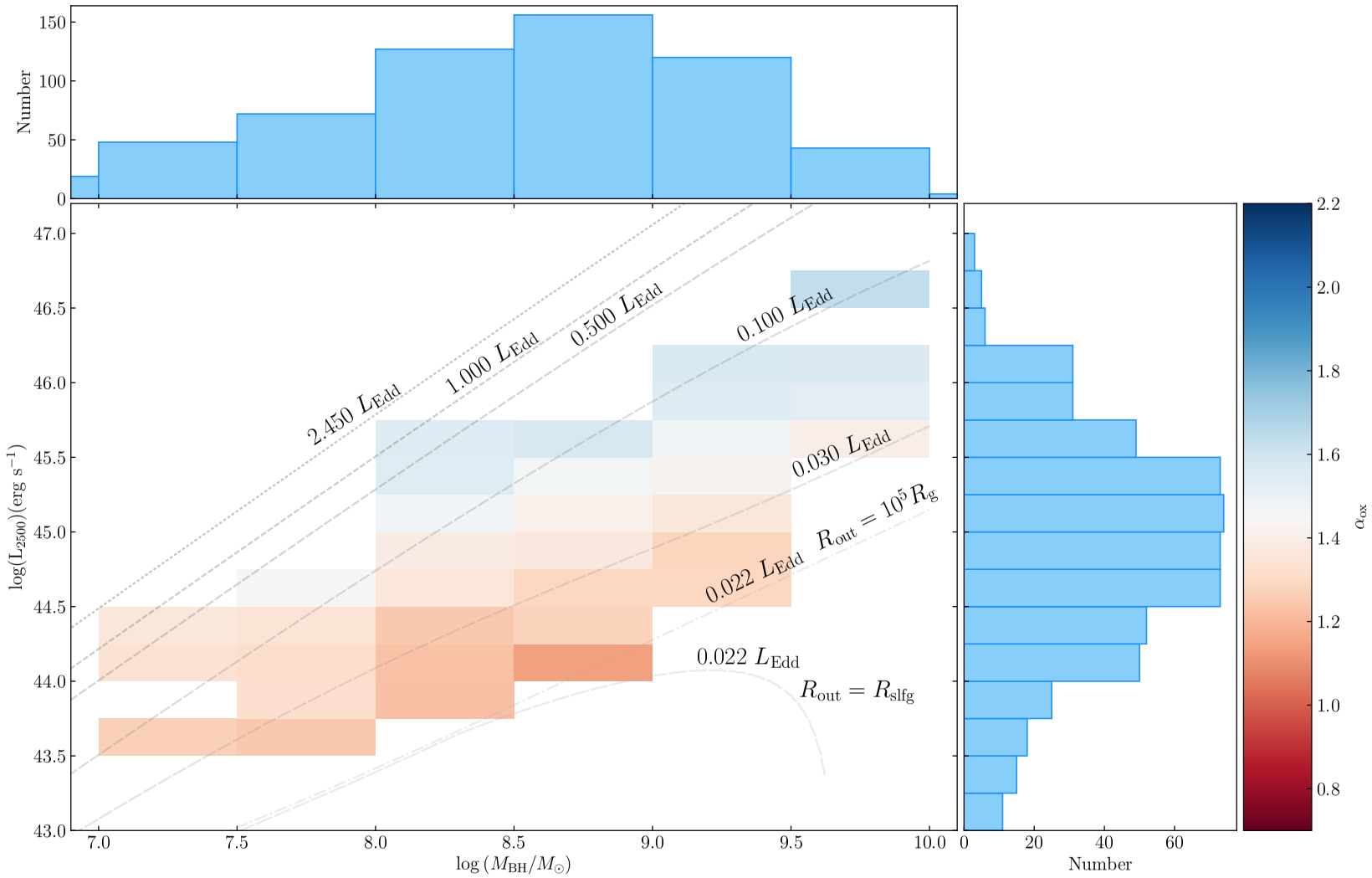

Figure 4.1: Black hole mass against $\log(L_{2500})$ both measured from SDSS optical spectra. Each grid point bin of 0.5 dex in mass and 0.25 dex in $\log(L_{2500})$ contains at least 5 objects. K22 show a similar plot for the full sample, but here we show only the curated sample, i.e. removing 60 very radio loud sources (Radio Loudness > 100) to avoid potential contamination from the jet, and removing an additional 54 sources which display clear signs of intrinsic cold or warm absorption. Lines of constant $\dot{m}$ calculated from QSOSED are overplotted.

UV data and extrapolating to the optical and X-rays, moving on we utilise AGNSED to fit a fully Comptonised model again only to the UV data. We then use AGNSED to fit a fully Comptonised model to all of the data. We compare the complex change in resultant $\dot{m}$'s between these fitting paradigms and note interesting divergences between the models and the data, particularly in the high mass range.

Using the insights gained from these fits we investigate the shape of the optical-UV continuum in the wider parameter space by constructing wider bandpass spectra from all of SDSS. These show no change in the optical-UV continuum at constant L_{2500} for changing mass by 2 dex. This is not compatible with *any* current accretion disc model. Even if the highest black hole masses are overestimated, the amount of finetuning required to match this looks contrived. Instead, we favour solutions where either the accretion disc is completely covered by a warm comptonising layer whose properties change systematically with $L/L_{\rm Edd}$, or the accretion flow structure is fundamentally different to that of the standard disc models. Throughout this Chapter we use $\dot{m}$=$L/L_{\rm Edd}$.

4.2 The Sample and Data

K22 includes a detailed description of the sample selection, spectral fitting procedures, and calculation of important parameters such as black hole mass and radio loudness. We include a short summary here for completeness.

4.2.1 Sample Selection

Our sample is primarily composed of sources taken from the Quasar Catalog of the Fourteenth SDSS Data Release (SDSS-DR14Q, Pâris et al. (2018)) cross matched with the fourth source catalogue of *XMM-Newton* (4XMM-DR9 Webb et al. (2020)). In order to have good quality X-ray spectra we only select sources with ≥ 250 counts in *XMM-Newton* without flags for high background, diffuse emission or poor source properties. The simultaneous OM optical/UV data are taken from the fourth *XMM-Newton* Serendipitous Ultraviolet Source Survey (XMM-SUSS4.1 Page et al. (2012)).

We consider all AGN with z ≤ 2.5 so as to have black hole mass from Mg II. We perform a visual inspection of each optical spectrum and remove any BAL's, Seyfert 2 sources and any object with an unreliable black hole measurement due to poor or contaminated line profiles, i.e. line profiles displaying obvious absorption. Through this process we obtain 633 sources.

We then supplement this sample with a population of Narrow line Seyfert 1 (NLS1) sources taken from (Rakshit et al., 2017), (R17). By repeating the same selection process with this catalogue we obtain 63 NLS1 not present in our previous sample, bringing the total number of sources to 696. Where z $\geq$ 0.4 we discard any OM filter contaminated by the strong Lyα λ1216 UV emission line.

4.2.2 Spectral fitting and Black Hole Mass estimates

Rakshit et al. (2020) (hereafter R20) analyse the SDSS-DR14Q spectra using a modified version of the PYTHON PyQSOFit package (Guo et al., 2018, 2019; Shen et al., 2019). We fit the additional NLS1 sources from R17 with a version of PyQSOFit modified to match the version used in R20. For the 84 sources in both R17 and R20 we compared our fits with those detailed in R20 and found good agreement.

All black hole masses were calculated from FWHM measurements described above, using the scaling relations detailed in Mejía-Restrepo et al. (2016); Woo et al. (2018); Greene et al. (2010).

As our sources extend out to a redshift of 2.5, we are able to measure Black Hole masses from either the H α, H β or Mg II broad emission lines. Where several of these are present we prioritise H β, then Mg II, followed by H α. However, we also visually inspect each optical SDSS spectrum and choose a different line in cases where the 'preferred' option was clearly of lower quality.

4.2.3 Radio Properties

Our 696 sources were cross matched with both the Very Large Array (VLA) Faint Images of the Radio Sky at Twenty-Centimeters (FIRST) Becker et al. (1995) and the National Radio Astronomy Observatory (NRAO) VLA Sky Survey (NVSS) Condon et al. (1998) catalogues using a matching radius of 10$''$ following Lu et al. (2007). Both surveys sample the sky at 1.4GHz, FIRST has a beam width of 5.6$''$ and NVSS of 45$''$. We matched 124 and 83 objects in FIRST and NVSS, respectively.

We converted $F_{1.4\mathrm{GHz}}$ to $F_{5\mathrm{GHz}}$ using the scaling relation $(\nu_{5\mathrm{GHz}}/\nu_{1.4\mathrm{GHz}})^{-\alpha_\mathrm{R}}$ where α_R is a common spectral index of 0.6. These $F_{5\mathrm{GHz}}$ values were converted to rest frame $L_{5\mathrm{GHz}}$ using the method of Alexander et al. (2003). In objects where we have an $L_{4400\text{Å}}$ and a FIRST detection, we calculated the radio loudness parameter, canonically defined as $L_{5\mathrm{GHz}}/L_{4400\text{Å}}$ (Kellermann et al., 1989).

Throughout the following analysis, 60 very radio loud sources (Radio Loudness > 100) were removed as there is a strong possibility that the jet emission dominates over that of the accretion flow itself. The most obvious example of this in the original sample is PMN 0948+0022, a NLS1 where the X-rays are clearly dominated by the jet emission which extends up to Fermi GeV energies (Foschini et al., 2012).

4.2.4 Sample Pruning

K22 did not perform a detailed inspection of the broadband SEDs or X-ray spectra in their analysis. Since in this work it is our goal to model the SEDS, we visually inspected each SED individually and removed any object with strong indicators of intrinsic (host) absorption. This could be due to cold/dusty gas in the molecular torus, easily seen as both the X-ray and UV data points are strongly attenuated at the lowest/highest energies respectively. We also excluded objects with partially ionised, warm absorption from nuclear winds, identified as a sharply concave soft X-ray shape (Reynolds & Fabian, 1995; Chakravorty et al., 2009).

This removed 54 sources from the original SOUX AGN sample of 696, described in K22, leaving 642 sources available for detailed analysis, this shall henceforth be referred to as the SOUX AGN sample. The 54 sources that have been removed from the sample before the fitting process are detailed in Table B.1.

4.3 The AGNSED / QSOSED models

4.3.1 Review of SED models for the accretion flow

We use several different SED models in this work, so here we outline each one in turn. All of them are based on the emissivity of an optically thick, geometrically thin accretion disc in full general relativity (Novikov-Thorne, $\epsilon_{NT}(r)$) (Novikov & Thorne, 1973). They give the luminosity emitted over a disc annulus as $L(r) = 2 \times 2\pi r \epsilon_{NT} n(r) dr$. Throughout this Chapter we use $\dot{m}{=}L/L_{\rm Edd}$.

4.3.2 Standard disc

The standard disc models assume that the disc luminosity at each annulus is emitted as a blackbody spectrum in which the spectral radiance density $B_\nu(T_{\rm eff})$ is a function of the effective temperature $T_{\rm eff}$ and the luminosity at a given radius $L(R) = \sigma_{\rm SB} T_{\rm eff}^4$. Pure standard

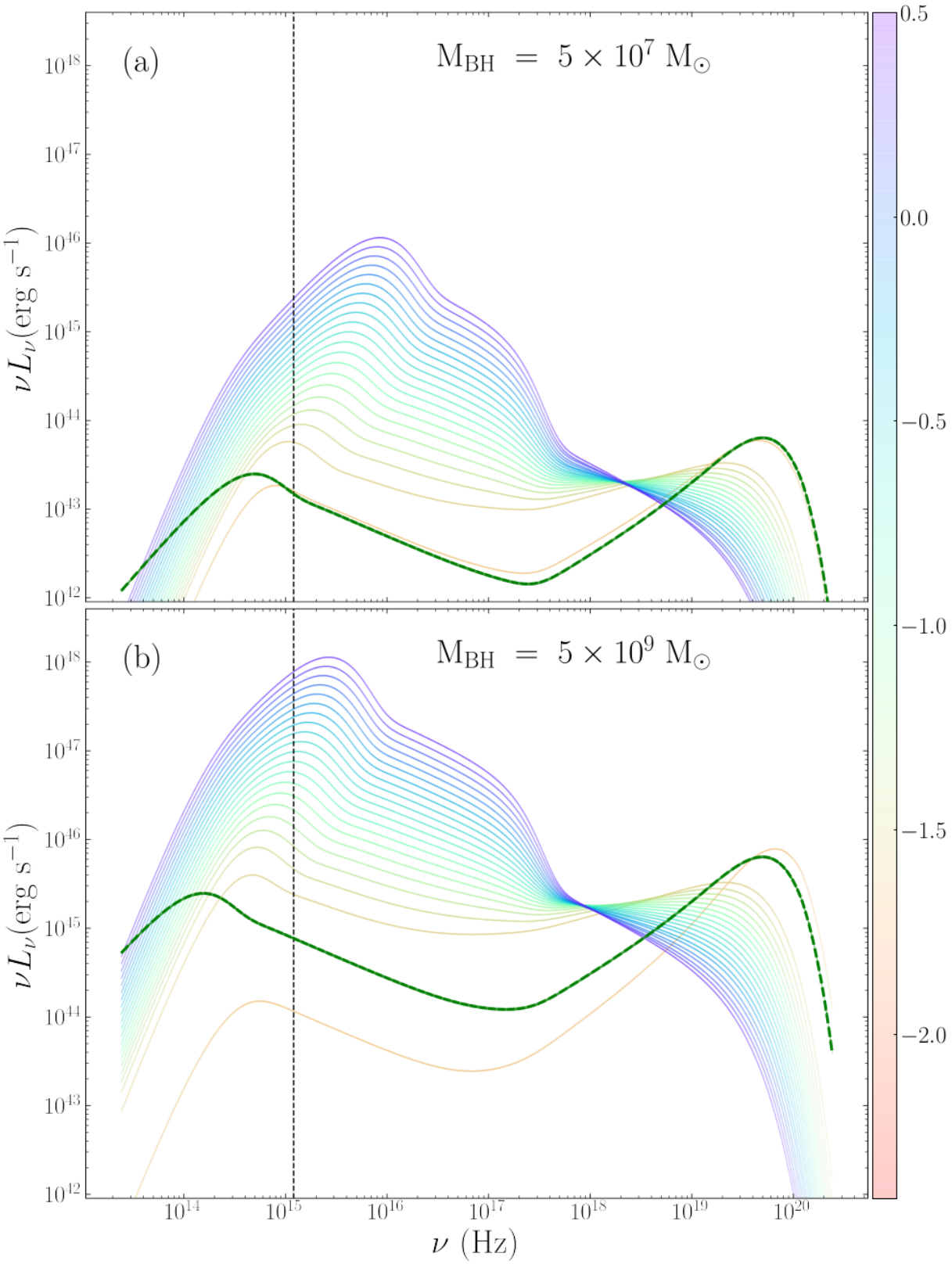

Figure 4.2: QSOSED predictions for the SED for black holes of mass $5 \times 10^7 M_\odot$ (a) and $5 \times 10^9 M_\odot$ (b) for a range of mass accretion rates assuming spin 0. The 2500Å point is indicated by the vertical dashed line on each panel. The green line in the right hand (high black hole mass) panel shows the effect of increasing the outer disc radius from the (small) self gravity radius to a value of $R_{\rm out} = 10^5 R_{\rm g}$ at the lowest $L/L_{\rm Edd}$.

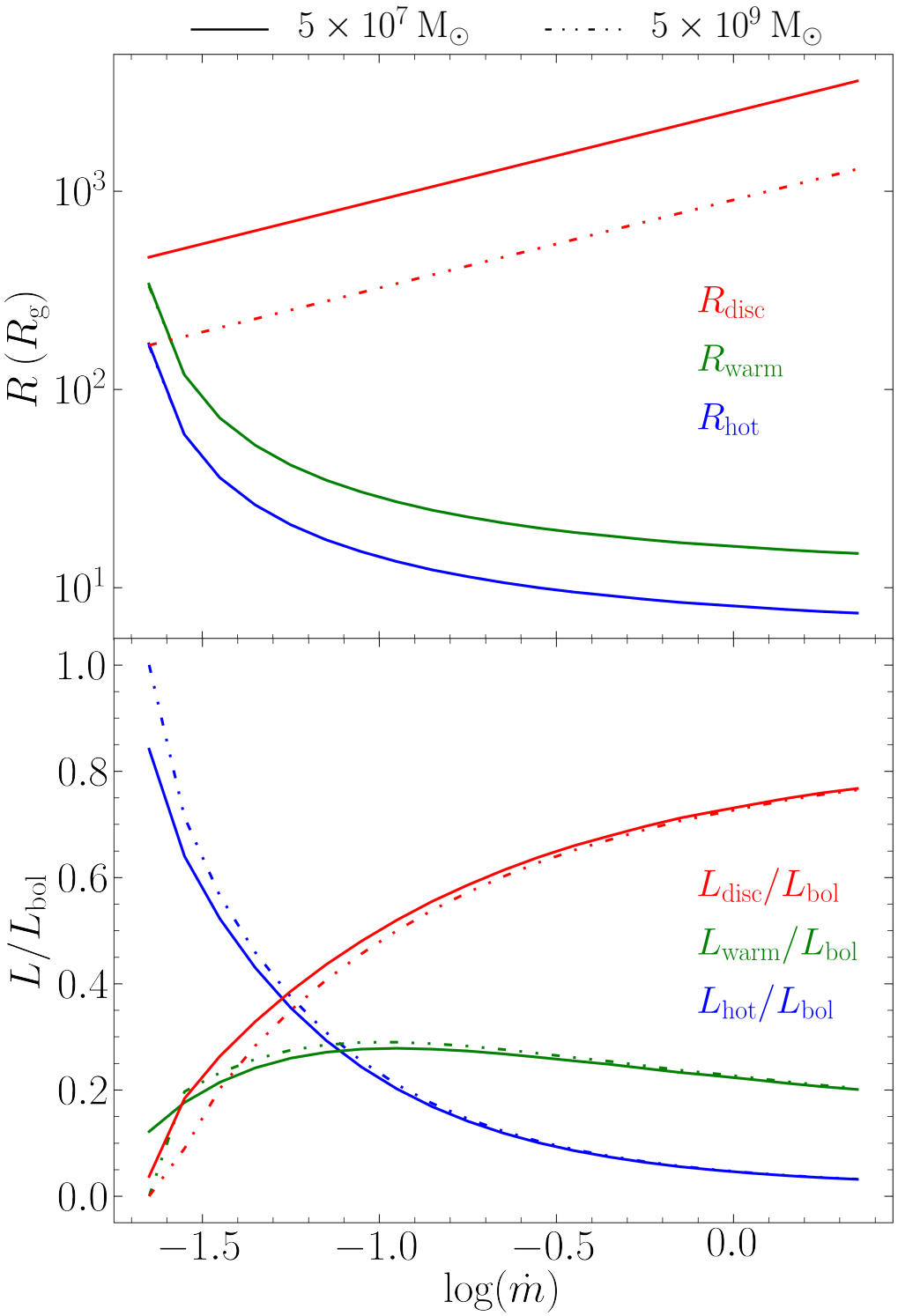

Figure 4.3: Radii and luminosities of each emission region of the QSOSED models shown in Figure 4.2. Solid lines show results for $5 \times 10^7 \mathrm{M}_\odot$, dashed show $5 \times 10^9 \mathrm{M}_\odot$. **Top Panel:** The hot corona extends from R_isco to R_hot (blue), the warm corona extends from R_hot to R_warm (green), and the standard disc extends from R_warm to R_out (red). **Bottom Panel:** Corresponding luminosities in each component: $L_\mathrm{disc}/L_\mathrm{bol}$ (red), $L_\mathrm{warm}/L_\mathrm{bol}$ (green) and $L_\mathrm{hot}/L_\mathrm{bol}$ (blue).

disc models are where this holds over the entire disc, from $R_{\rm out}$ to $R_{\rm isco}$. This is the model which is most often used to fit the optical/UV spectra, but it cannot produce the soft and hard X-ray emission observed from AGN.

4.3.3 Standard disc with colour temperature correction

The standard disc is the best physically understood model for the underlying accretion flow structure, but the best detailed calculations incorporate radiative transfer through the standard disc vertical structure. These calculations predict a shift of the observed disc emission to higher temperatures (Hubeny et al. 2001, D12). The standard disc energy is mostly dissipated close to the midplane, so it diffuses outwards, setting up a vertical temperature gradient through the optically thick material until it reaches the photosphere and escapes to the observer from and effective optical depth $\tau_{\rm eff} = \sqrt{\tau_{\rm a}(\tau_{\rm a}+\tau_{\rm s})} \sim 1$ where $\tau_{\rm a}$ and $\tau_{\rm s}$ are the optical depth to true absorption and scattering, respectively. True absorption opacity depends on the density and temperature of the material, and decreases with frequency, while electron scattering is a constant. Thus, a single radius in the disc has a spectrum which follows the expected blackbody emission only for $h\nu \ll kT_{\rm eff}$, where $\tau_a \gg \tau_s$, but the higher frequencies are dominated more by scattering, moving the photosphere deeper into the disc vertical structure, where it samples higher temperature emission. This effect becomes apparent for photosphere temperatures $T_{\rm eff} \geq 3\times10^4$ K, and is fairly well approximated by a changing colour temperature correction, $\rm f_{col}$ so that each radius emits a spectrum which can be approximated as $B_\nu(\rm f_{col}T_{\rm eff})/{\rm f_{col}}^4$ where $\rm f_{col} = 1$ for $T_{\rm eff} < 3\times10^4$ K. This produces a slight flattening of the spectrum in the UV, as the radii which would have emitted at this energy are shifted to higher apparent temperature, making the disc spectrum peak at higher energies. However, it still cannot produce the soft or hard X-ray tail. This model is utilised in section 4.5.3.

4.3.4 AGNSED, including soft and hard Comptonisation

The model AGNSED, described in KD18, is more flexible as it allows the accretion energy to be emitted as soft and hard Comptonisation as well as disc blackbody. These are tied together by the underlying assumptions that the energy is from accretion and that the flow is in a radially stratified structure such that the emission only thermalises to a blackbody for $R > R_{\rm warm}$. For $R_{\rm warm} < R < R_{\rm hot}$ the power is instead emitted as an optically thick, warm Comptonisation component, largely in the soft-X-rays (described by its photon index $\Gamma_{\rm warm}$

and electron temperature $kT_{\rm e,warm}$, with seed photon temperature set by the underlying disc), while below $R_{\rm hot}$ the spectrum switches to hot, optically thin Comptonisation (parameters $\Gamma_{\rm hot}$ and $kT_{\rm e,hot}$) from the corona, largely emitting in the hard-X-rays.

This has enough flexibility to fit the entire SED of the accretion flow from optical-UV and X-ray, but has a number of free parameters describing the non-standard disc sections of the flow.

4.3.5 The QSOSED phenomenological model

KD18 fit AGNSED to the SED of three AGN of similar mass, but spanning a range of $L/L_{\rm Edd}$. Their analysis showed that the power dissipated in the hot Comptonisation region was approximately constant at $0.02L_{\rm Edd}$. This is the maximum ADAF luminosity, defining the hard-soft transition luminosity, but its persistence above the transition is not what is expected from BHB. The BHB can show very disc dominated spectra above the spectral transition i.e. with hot Comptonisation power $\ll 0.02L_{\rm Edd}$.

Another difference between the AGN and BHB SEDs is that AGN at $L < 0.1L_{\rm Edd}$ show hard X-ray spectra, with $\Gamma_{\rm hot}< 1.9$, whereas the BHB always show soft spectra, with $\Gamma_{\rm hot}> 2.0$ above the transition. This difference is important as it is difficult to produce hard spectra. An isotropic X-ray corona above a disc must have $\Gamma_{\rm hot}> 1.95$ due to reprocessing. Not all the photon energy illuminating the disc from the corona can be reflected, especially for hard spectra where the luminosity peaks at $\sim$ 100 keV. Compton downscattering on reflection means that at least 30% of the illuminating flux goes instead into heating the disc, producing a reprocessed thermal component which is re-intercepted by the corona. Even if the disc is completely dark (passive), so that all the accretion energy is dissipated in the corona, these reprocessed photons give a lower limit to the seed photons such that $L_{\rm seed} \sim L_{\rm x}$, which ties the X-ray spectral index to $\Gamma_{\rm hot}\sim$ 2 (Haardt & Maraschi, 1991, 1993; Stern et al., 1995; Malzac et al., 2003). Thus the AGN with spectra $\Gamma_{\rm hot}< 1.9$ most likely still have a truncated disc geometry, unlike the BHB above the transition (KD18).

The QSOSED model folds all these results into a predictive model for AGN SED. Firstly, it calculates the truncation radius assuming that the hot X-ray plasma replaces the inner disc from $R_{\rm isco}$ out as far as required for the accretion disc power to reach $0.02L_{\rm Edd}$. Thus $R_{\rm hot} \gg R_{\rm isco}$ for an AGN at $L_{\rm bol}$=$0.03L_{\rm Edd}$, leaving very little power for the outer UV disc emission. Conversely, for an AGN at $L_{\rm bol}\sim L/L_{\rm Edd}$, $R_{\rm hot} \sim R_{\rm isco}$ so its SED is dominated by a UV bright disc component. This geometry and energy balance is used to calculate the

self consistent power law index of the hot Comptonisation region, $\Gamma_{\rm hot}$, fixing its spectrum in both normalisation and index assuming that the electron temperature remains fixed at 100 keV (KD18).

The warm Comptonisation region is even less well understood than the hot Comptonisation, but it can be fit with optically thick material (with optical depth $\tau \sim$ 10–20) which makes it look like it is associated with the disc, perhaps due to the dissipation region moving upwards toward the photosphere, as opposed to the standard Shakura-Sunyaev disc where the dissipation is mainly in the equatorial plane. In this case, Compton reprocessing again sets the energy balance between the seed photons and electron heating. The difference is that for the observed low temperature of $\sim$ 0.2 keV, this gives $\Gamma_{\rm warm}\sim$ 2.5 (Petrucci et al. 2018, KD18). The QSOSED model fixes these parameters, but needs also the transition radius, $R_{\rm warm}$ at which the disc stops emitting as a blackbody standard disc, and instead produces the warm Comptonisation region. Since this is not understood theoretically, KD18 took it from the data which could be described by $R_{\rm warm}= 2R_{\rm hot}$.

Thus the QSOSED model contains a number of assumptions in addition to those of AGNSED (emissivity given by Novikov-Thorne thin disc, but the emission mechanism is radially stratified)

- The hard X-ray luminosity is $0.02L_{\rm Edd}$ irrespective of $L_{\rm bol}$.
- The temperature of the warm Comptonised disc is $\sim$ 0.2 keV, and it is a passive structure so $\Gamma_{\rm warm}=$ 2.5.
- The warm Compton region extends from $R_{\rm warm}= 2R_{\rm hot}$.

The SOUX sample can test many of these assumptions as we have many more than the 3 SEDs used by KD18, and it spans a larger range in $\dot{m}$ and a much larger range in $M_{\rm BH}$.

Example QSOSED spectra

Figure 4.2 shows the model SED for changing $L/L_{\rm Edd}$ (in steps of $\log(L/L_{\rm Edd})$=0.1) for black hole mass of $5 \times 10^7 M_\odot$ (left) and $5 \times 10^9 M_\odot$ (right), for $\log(\dot{m}) = -1.645 - 0.355$ (corresponding to $0.026<\dot{m}<2.6$). The model is only defined over this range as below the lower limit, AGN should make a changing look transition to be completely dominated by the X-ray hot flow (ADAF), and above the upper limit, the effects of optically thick advection/winds

should become apparent, changing the emissivity to that of a slim disc (Abramowicz et al., 1988; Kubota & Done, 2019).

These figures show the effect of the assumptions of the QSOSED model. For both masses, the $R_{\rm hot}$ and $\Gamma_{\rm hot}$ values are the same at a given $\dot{m}$, with $R_{\rm hot}$ decreasing from (155-9)$R_{\rm g}$ as $\dot{m}$ increases from 0.026 to 2.6 (see Figure 4.3), so that the disc and warm Comptonisation regions progressively dominate more of the emission. However, $M_{\rm BH}$ has an impact on the SED as the outer disc temperature and seed photon temperature for the warm disc region are lower at the higher black hole masses. The 2500Å flux (indicated by the vertical dashed lines in Figure 4.2) is generally mostly produced in the outer standard disc region for the lower mass AGN ($5\times10^7 M_\odot$). However, for the higher mass of $5\times10^9 M_\odot$, the outer disc generally peaks below 2500Å, so the UV flux is instead dominated by the warm Comptonisation up to $\sim 0.1 L/L_{\rm Edd}$. Thus the model predicts a mass dependence to the $\alpha_{\rm ox}$ values due to the difference in UV emission mechanism with mass, as well as a dependence on $\dot{m}$.

These model SEDs also show how to understand the lines of constant $L/L_{\rm Edd}$ overlaid on Figure 4.1. A pure disc spectrum should have $L_{2500} \propto (M_{\rm BH}^2 \dot{m})^{2/3}$. Both high and low masses are dominated by the disc at the highest $\dot{m}$, so the grey lines of constant $\dot{m} \gtrsim 0.5$ have $L_{2500} \propto M_{\rm BH}^{4/3}$, a non-linear relation, unlike the constant bolometric correction which is often assumed, where $L_{2500} \propto M_{\rm BH}$. At lower $\dot{m}$, the highest masses start to bend away from this expected slope as the lower standard disc temperature means the UV is dominated by the warm Comptonisation region rather than the standard disk. Here and throughout L_{2500} refers to the monochromatic luminosity at 2500Å (in units of erg s^{-1}).

There is an especially sharp decrease at the lowest $\dot{m}$ for the highest masses. This is because of the assumption that the outer radius of the disc is set by self gravity. The self-gravity radius ($R_{\rm sg}$) is only a few hundred $R_{\rm g}$ for the most massive AGN at the lowest $\dot{m}$ (see Figure 4.3). The QSOSED model has the accretion flow end before making the transition to a thin disc for masses above $2\times10^8 M_\odot$. Above this mass the warm Compton region gets smaller and smaller, extending across (155-311)$R_{\rm g}$ at $2\times10^8 M_\odot$, but shrinking to (155-167)$R_{\rm g}$ at $5\times10^9 M_\odot$ for $\dot{m}$=0.026 (see Figure 4.3). This dramatically reduces the warm Compton emission, leading to an extremely hard X-ray spectrum.

However, the extent of the disc is very uncertain as the self gravity radius is calculated assuming that the disc surface density is set by the standard Shakura-Sunyaev disc in the radiation pressure dominated regime, yet the SED at low $\dot{m}$ is very unlike a standard disc. The green dashed line in Figure 4.2b shows the effect of increasing the disc outer radius to a

generic value of $10^5 R_{\rm g}$ in the QSOSED model. This gives a dramatic increase in the UV and especially the optical luminosity, with the outer standard disc region now evident in the SED at the lowest frequencies. This shows that the optical/UV spectra of the highest mass AGN are sensitive to the outer extent of the disc.

Figure 4.2b also shows that the QSOSED spectra for the highest mass AGN are predicted to peak in the observable optical/UV for $\dot{m}$<0.1. Thus it is the highest mass AGN at the lowest $\dot{m}$ which are most sensitive to the standard disc region and how (or if) this transitions to the warm Comptonised region. Instead, the lowest mass AGN at the highest $\dot{m}$ are most sensitive to the transition from warm Comptonisation to hot Comptonisation. These low mass, highest $\dot{m}$ are also the ones which are most likely to be at low redshift, maximising the visibility of the soft X-ray excess in the XMM bandpass.

Thus the predictions of the QSOSED model can be directly tested on our new sample of AGN which span a wide range of mass and $\dot{m}$.

4.4 Fitting to the data

Throughout this section we consider the mean SED from the data in each grid point, comparing it to a series of models. For the first two sections (4.1 and 4.2) we only fit to the UV data from the OM, and then extrapolate to the optical SDSS and X-ray bandpasses, forming a true test of the model predictions. We only include the X-ray data in the fit in section 4.3. We never include the SDSS data in the fits as this is not simultaneous, so variability could distort the modelling. However, this is minimised by averaging over all the objects in the bin, and the generally good match between the SDSS and OM data provides a check on intercalibration/aperture issues.

We make the mean SED from the data by fitting a single model for each grid point, with black hole mass fixed at the (logarithmic) centre of the mass gridpoint. This model is fit simultaneously to the relevant spectral range of all objects in that bin, but with parameters set to the individual objects co-moving distance/redshift, and appropriate galactic reddening/absorption. The data of each object is plotted in $\log \nu$ versus $\log \nu L_\nu$, where the data are shifted into the rest-frame of the galaxy and corrected for galactic absorption by considering the dust maps of Schlafly & Finkbeiner (2011) using the ASTROPY extinction module with the (Cardelli et al., 1989) extinction profile. Each instrument for each individual source is rebinned onto a common energy grid (6 bins for *XMM-Newton*, and 4 for the OM). We

calculate a weighted log mean of all the data for each bin, with any data point with an intersecting energy uncertainty being included. The weighted 1σ standard deviation for each bin is displayed either side of the weighted mean.

Similarly the composite SDSS optical spectra for each bin are the geometric mean of the dereddened, redshift corrected individual spectra, again with weighted $\pm 1\sigma$ standard deviation. We bin in log space as this means that the spectral slope of the composite has the mean of the spectral slopes of the individual spectra from which it is composed (see Reichard et al. 2003).

All plots containing SED insets follow the same format as described above, and each inset plot spans 2 dex in luminosity and covers the spectral range between $10^{14.2}$Hz and $10^{19.6}$Hz.

4.4.1 QSOSED in each $M_{\rm BH}$-L_{2500} bin

QSOSED can be used to create an entire SED from a single data point so we test this by fitting *only* to the mean UV data and extrapolating the resultant model down through the optical and up into the X-ray bandpass. We fix black hole spin at $a_* = 0$, so $\dot{m}$ is the only free parameter, essentially fixing it to give the mean L_{2500} of each grid point. We show the results of this fit, hereafter called QSOSED UV-tied, in Figure 4.4.

Each inset plot shows the QSOSED fit (solid line) together with the SDSS optical (red), *XMM-Newton* OM UV (blue) and X-ray (pink) data for all objects located in each respective gridpoint. We stress that the QSOSED UV-tied model is fit only to the UV data, not to the X-ray or optical points on this plot, yet to zeroth order the QSOSED model is in broad agreement with both the optical and X-ray data. The range of resultant $\dot{m}$ match the grey lines plotted in Figure 4.1, as expected given that these fits are forced to match the L_{2500} luminosity. These are spin zero models, but in general the entire SED is well fit. Extreme spin gives a factor 6 more luminosity for the same mass accretion rate through the outer disc. There is no strong evidence for this. Strong wind losses would reduce the mass accretion rate through the innermost regions of the disc. We might expect this to be most important at the highest $\dot{m}$ but there is no clear trend of the X-ray flux being strongly overpredicted. To within factors of a few, the thin disc emissivity which is hardwired into QSOSED is giving the correct bolometric flux for a zero spin black hole.

To first order though, there are some interesting discrepancies. At low masses, $\log(M_{\rm BH}/M_\odot) < 8.0$, the optical data are systematically higher than the model especially at low $\dot{m}$, so that the optical/UV spectra are redder than the standard outer disc spectrum assumed in

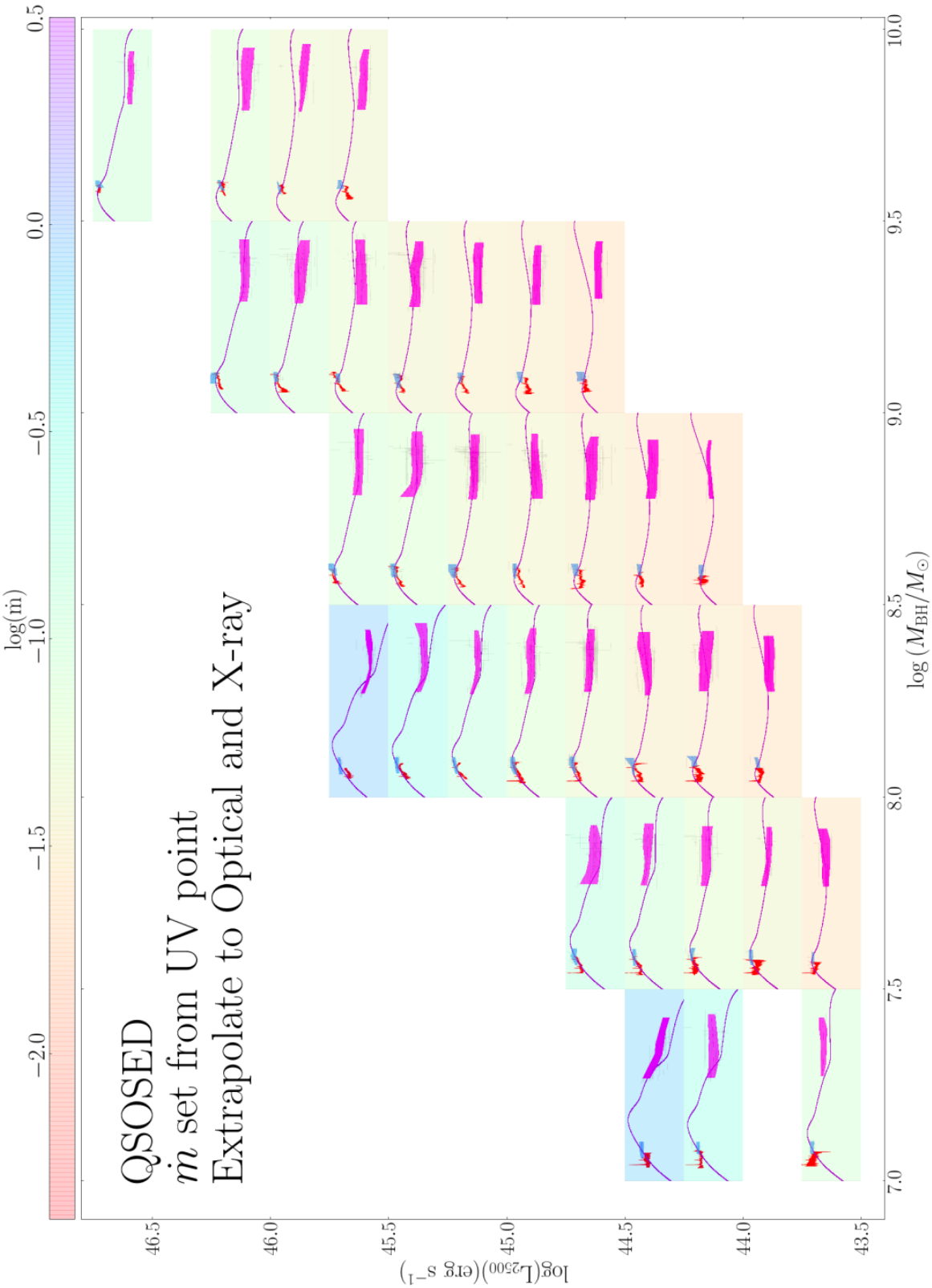

Figure 4.4: The same black hole mass - luminosity plane shown in Figure 4.1, but now with inset in each bin showing the mean data SED from SDSS (red), OM (blue) and X-ray (pink) in that bin, together with the QSOSED UV-tied SED model. The only free parameter in these models is $\dot{m}$, and it is only fit to the UV data and then extrapolated across the whole spectrum. The background colour of each bin indicates the resultant $\log(\dot{m})$.

the model. This could be due to host galaxy contamination, which should become more prominent for low black hole mass, low luminosity AGN (Done et al. 2012, hereafter D12). The X-ray data are also systematically ∼0.5 dex higher than the model for these lower black hole masses at high $\dot{m}$, so require more than the assumed 2% of power dissipated in the hot coronal region (see also Middei et al, in prep).

However, the major discrepancies are all at the highest masses, $\log(M_{\rm BH}/M_{\odot})$ >9.0. The QSOSED model systematically over-predicts the X-ray by ∼(0.3-0.6) dex in luminosity for all values of $\dot{m}$. Most surprisingly though, the optical/UV slope in the data is generally much bluer than the model. This is not likely to be from any aperture difference between SDSS and OM as the galaxy contamination should be negligible at these high luminosities. The optical/UV data simply do not look like a standard outer disc for a black hole of this mass and mass accretion rate. The observed optical/UV spectra in this mass range are still typically rising towards Lyα but the QSOSED models predict that the emission should peak below this energy. As seen in the model spectra shown in Figure 4.2, stronger outer disc emission could easily be produced by increasing the outer disc radius. However, this further decreases the characteristic energy at which the disc components peak, whereas the data show that the QSOSED model (with its very small outer radius from self gravity) already predicts the peak energy being too low. In essence, the optical/UV data do not look like the QSOSED models in this range, as the outer standard disc in the model peaks at too low an energy.

This is very surprising, especially as the outer standard disc is the least controversial of all the QSOSED components. There could be reddening by dust in the host, but this would make the mismatch worse as the data would shift even bluer.

The spectrum rather appears like a standard disc, but shifted over to higher energies than predicted for black holes of this high mass for a wide range in mass accretion rate.

Section 5 considers the optical/UV spectral shape in more detail, and shows that this mismatch at the highest masses is not due to the assumption in QSOSED that the warm and hot Comptonisation power is derived from the disc. Even pure disc models, where all the power is dissipated in blackbody emission down to $R_{\rm isco}$ of a high spin black hole cannot match the far UV emission in the spectra observed in the high mass/low $\dot{m}$ bins.

Interestingly, where the QSOSED models fit the SEDs (highest $\dot{m}$ for $\log(M_{\rm BH}/M_{\odot})$>8), their predicted EUV continuum is able to match the observed 1640Å He II equivalent widths (Temple et al, in prep).

4.4.2 Fully Comptonised outer disc: AGNSED ($R_{\rm hot}$=10$R_{\rm g}$) in each $M_{\rm BH}$-L_{2500} bin

One way to shift the disc spectrum to higher energies is if the whole disc is covered by the warm Comptonised layer. This was suggested by Petrucci et al. (2018) as they had noticed that the spectra of even the $10^8 M_\odot$ AGN seemed to be well fit with just warm and hot Comptonisation, without the need for a standard disc once the host galaxy was subtracted.

We simultaneously fit all the OM data in a given grid point as before, only now we use the more flexible AGNSED and fix $R_{\rm warm}$=$R_{\rm out}$ so that the entire outer disc is covered by the warm Compton layer. We again fix $\Gamma_{\rm warm}$=2.5 and $kT_{\rm warm}$ = 0.2 keV in this region. However, AGNSED also allows freedom in $R_{\rm hot}$ and in $\Gamma_{\rm hot}$, so we first fix these to 10$R_{\rm g}$ and 2.0, respectively in order to find the mean $\dot{m}$ for each grid point. Setting $R_{\rm hot}$ at 10$R_{\rm g}$ is equivalent to fixing the power dissipated in hot corona to $\sim$ 7% of $L_{\rm bol}$ (see Figure 4.3 for log($\dot{m}$)=-0.5).

The results from this fitting procedure (henceforth referred to as AGNSED UV-tied $R_{\rm hot}$=10$R_{\rm g}$) are displayed in Figure 4.5. The background, as in Figure 4.4, is colour coded to the $\dot{m}$ value derived for the bin. These are shifted in a complex way from the values derived before. Figure 4.6a shows the ratio of mass accretion rates derived from QSOSED and the AGNSED UV-tied models. At lower mass/higher L_{2500}, the AGNSED models need a higher $\dot{m}$ than QSOSED, whereas the opposite is true at higher masses.

Figure 4.6b show why this is the case for each of these conditions (bins labelled 1 and 2). At the lowest masses (bin 1) the 2500Å point (dashed line) is in the standard disc region of the QSOSED model (red), as can be seen by comparison to a completely standard disc model for the same $\dot{m}$ (purple: $R_{\rm hot}$=$R_{\rm warm}$=6$R_{\rm g}$). Instead, a model with the same $\dot{m}$ where the entire disc is covered by the warm Compton region (black: $R_{\rm hot}$=6$R_{\rm g}$, $R_{\rm warm}$=$R_{\rm out}$) has much lower luminosity at 2500Å. This is because the warm Comptonisation shifts the disc spectrum to higher energies, acting like a colour temperature correction. Hence to match the data at 2500Å requires a higher $\dot{m}$ (dashed black line).

Instead, for the highest masses, the behaviour is opposite as the QSOSED model (red) already had the 2500Å flux produced by the warm Comptonisation rather than the outer standard disc. A completely standard disc (purple: $R_{\rm hot}$=6$R_{\rm g}$, $R_{\rm warm}$=6$R_{\rm g}$) is again shown for comparison. Completely covering the disc with the warm Comptonisation with the same $\dot{m}$ (black: $R_{\rm hot}$=6$R_{\rm g}$, $R_{\rm warm}$=$R_{\rm out}$) now over-predicts the 2500Å flux, so here the data require a lower $\dot{m}$ to match the data than before (dotted black line).

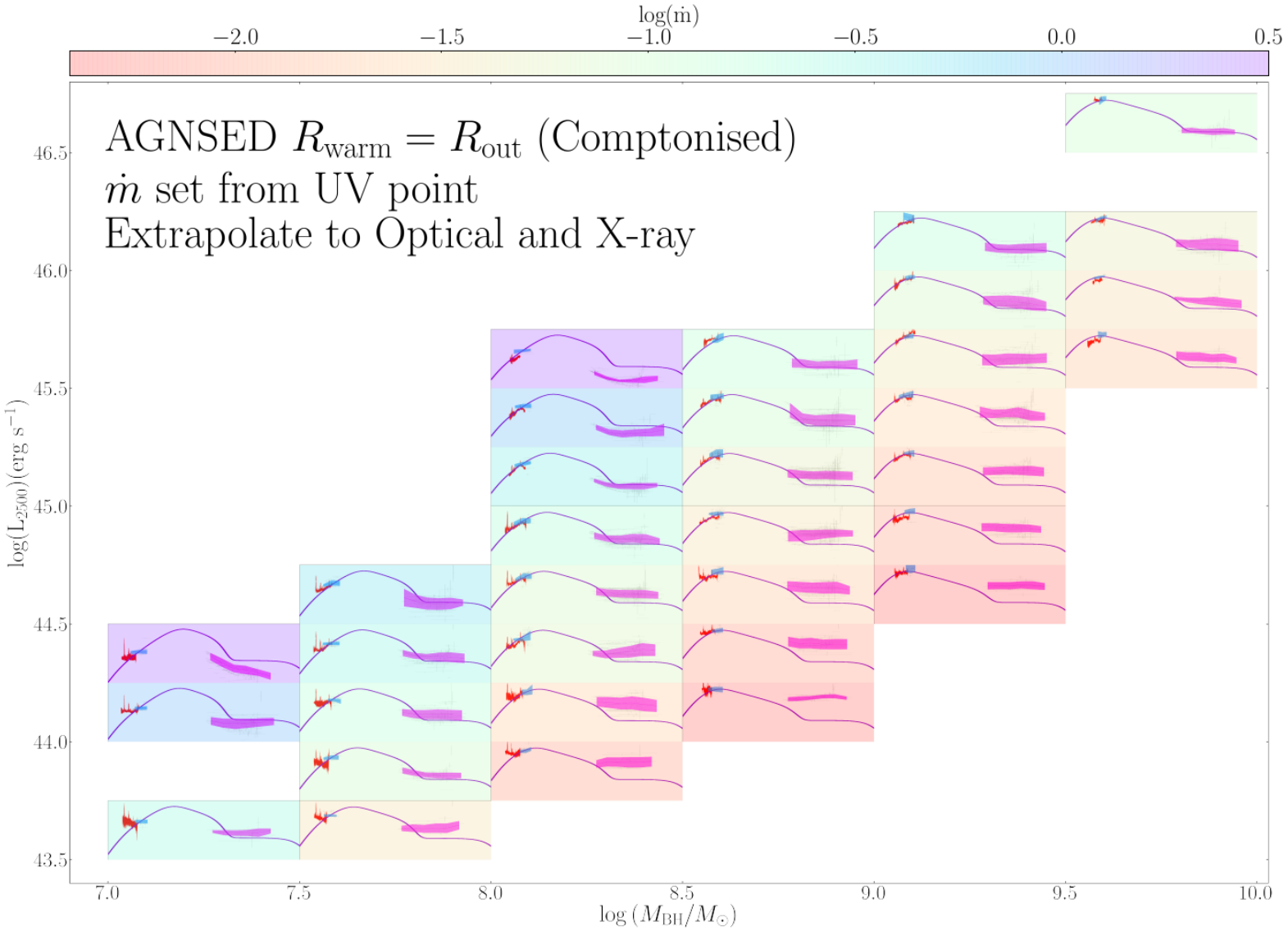

Figure 4.5: As in Fig. 4.4, but with the AGNSED UV-tied model. This has $R_{\rm warm}$=$R_{\rm out}$ so the entire outer disc is Comptonised, fixed $R_{\rm hot}$=10$R_{\rm g}$ and $\Gamma_{\rm hot}$=2.0, so $\dot{m}$ is the only free parameter. Again, this is fit only to the UV data points, and then extrapolated. The background colour in each bin indicates the resultant log($\dot{m}$), which is often quite different to that derived from the QSOSED UV tied models.

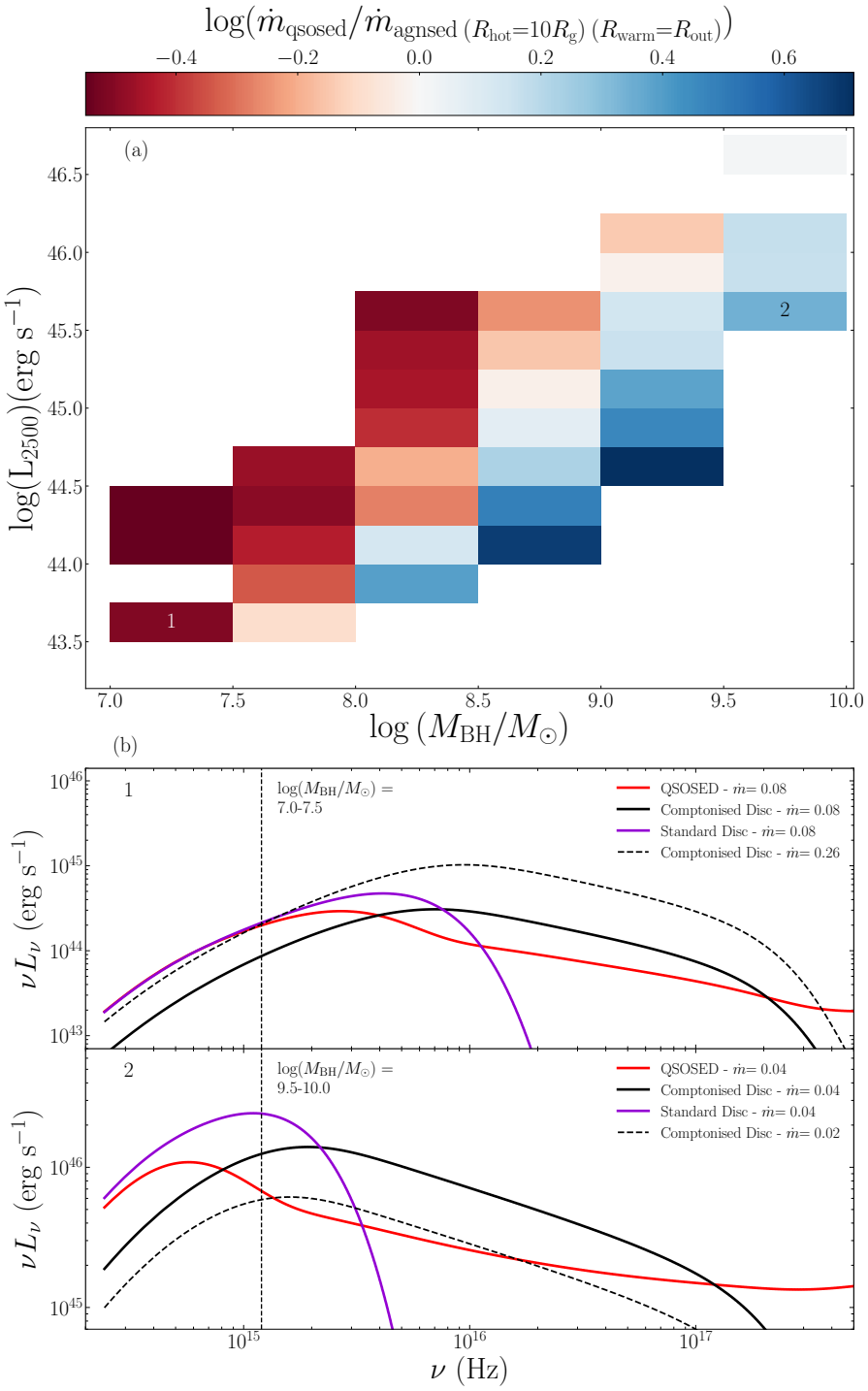

Figure 4.6: **Left-hand Panel:** A comparison of the $\dot{m}$ derived from the QSOSED UV-tied and AGNSED UV-tied models of Figs 4.4 and 4.5. Lower mass bins systematically require higher $\dot{m}$ for the Comptonised outer disc AGNSED model, while the opposite is true for the highest masses. This is explained by looking in detail for spectra in the black hole mass/luminosities bins highlighted as 1 and 2. **Lower Right-hand Panel:** In the low mass bin, the QSOSED UV tied model (red) is dominated by the outer standard disc. A pure standard disc of the same $\dot{m}$ is shown in purple, and has very similar L_{2500}. Comptonising this entire emission shifts the spectrum to the right (black line), so this has much lower L_{2500} than before. Thus a fully comptonised disc requires higher $\dot{m}$ (black dashed line) to recover the same L_{2500} as before. **Upper Right-hand Panel:** In the high mass bin, the outer disc peaks below 2500Å, so the QSOSED UV tied model is dominated by the warm Comptonised emission (red). A pure disc spectrum (purple) at the same $\dot{m}$ has much larger luminosity at 2500Å. Completely comptonising this (black) still has larger luminosity at 2500Å, so to match the same L_{2500} point requires a smaller $\dot{m}$.

This complex change in $\dot{m}$ across the large range in $M_{\rm BH}$ and L_{2500} increases the range of $\dot{m}$ spanned by the sample. The lowest $\dot{m}$ values are now well below $\dot{m}$= 0.02, so this loses the correspondence of the 'changing look' transition in AGN with the soft-hard transition from a disc to ADAF-like state in BHB (e.g. Noda & Done 2018; Ruan et al. 2019).

Nonetheless, Figure 4.5 shows that the new model has the desired effect in giving a much better match to the optical/UV spectra at high black hole masses, but now the optical/UV from the lower mass black holes are not well fit. This could be due to host galaxy contamination, but there are additional tensions in the SEDs. The model now overpredicts the soft X-ray excess, most clearly in the higher luminosity bins with $\log(M_{\rm BH}/M_{\odot})$<8.5. A large soft X-ray excess is an inevitable result of assuming that all the disc power above $R_{\rm hot}$=10$R_{\rm g}$ is Comptonised.

However, it is also clear that a fixed $R_{\rm hot}$=10$R_{\rm g}$ is not compatible with the X-ray data. This is most clear in the $\log(M_{\rm BH}/M_{\odot})$=8.0-8.5 bin, where this model systematically under-predicts the X-ray power at low $\dot{m}$, and over-predicts it at high $\dot{m}$. There is a very clear decrease in the ratio of hard X-ray to bolometric luminosity as a function of $\dot{m}$ (see e.g. Vasudevan & Fabian 2007; Vasudevan 2008). Hence we let the parameters of the hot Comptonisation region be free to see if this can resolve the tensions seen here.

4.4.3 Comptonised outer disc: AGNSED fits to the mean SED using all *XMM-Newton* data

We still assume the outer disc is completely covered by the warm Comptonisation region with $\Gamma_{\rm warm}$=2.5 and $kT_{\rm e,warm}$ = 0.2 keV but now allow freedom in the hot corona parameters of $R_{\rm hot}$ and $\Gamma_{\rm hot}$ in addition to $\dot{m}$. This requires that we fit the X-ray spectra as well as the UV data in each grid point. Figure 4.7 shows the results of this. There is now a clear discrepancy in the normalisation of the optical/UV model at lower masses. This is especially evident for $\log(M_{\rm BH}/M_{\odot})$<8.0 at high $\dot{m}$, where now the UV normalisation does not match the data. This is because the X-rays are now included in the fit as well as the UV, so the model is averaging between them. The model SED now has a strong soft X-ray excess as the entire outer disc is covered by a warm Comptonising layer. The X-ray spectra do not generally support such a strong soft X-ray excess, so the fit compensates by reducing $\dot{m}$ (which itself will reduce the size of the soft excess) and increasing $R_{\rm hot}$. But this in turn overpredicts the observed X-ray luminosity in the 2-10 keV bandpass, so the fit pushes the X-ray spectral index to its hard limit of $\Gamma_{\rm hot}$=1.6 so as to put some of this power above 10 keV where there

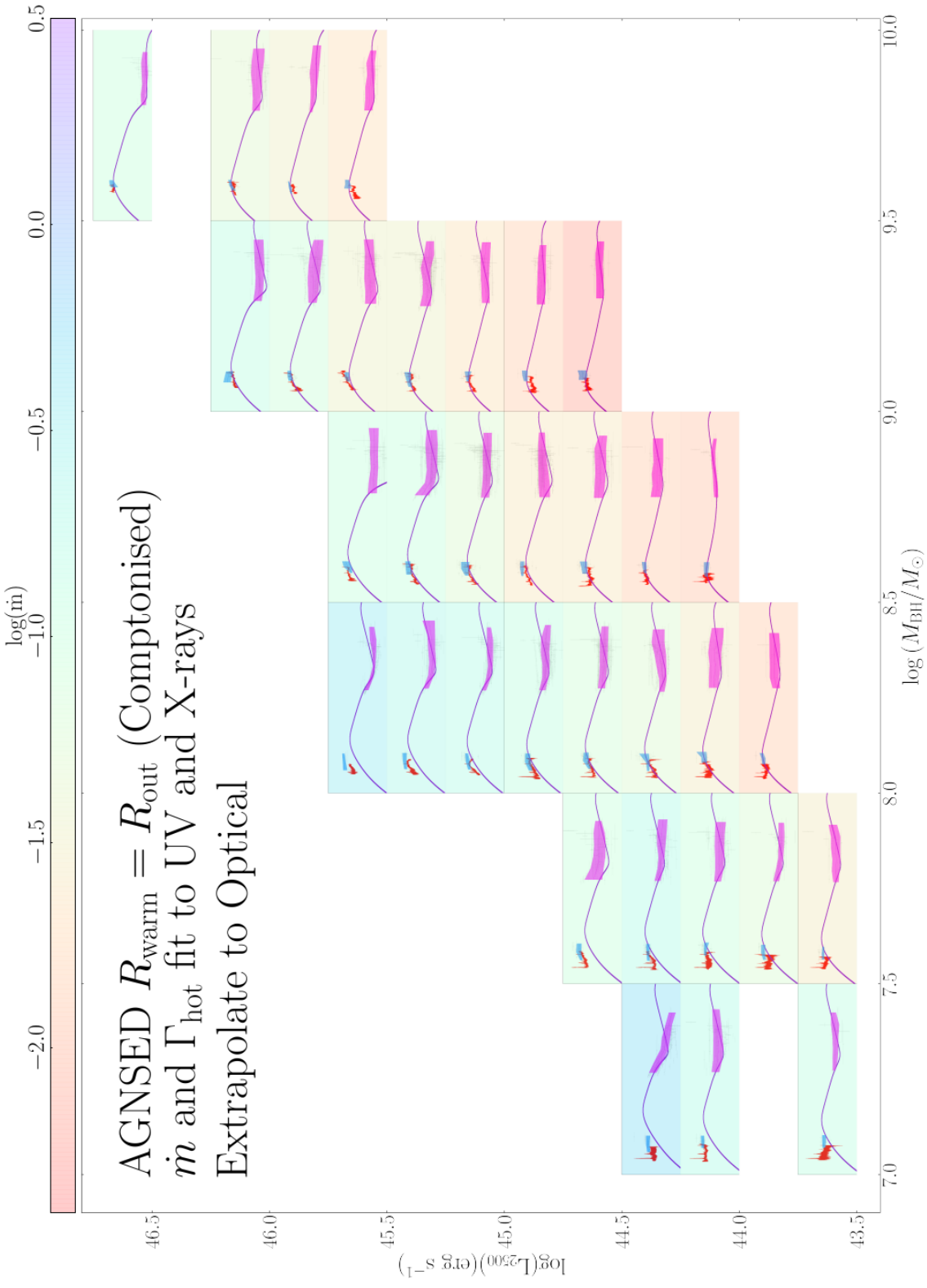

Figure 4.7: As in Fig. 4.5 but with the AGNSED model fit to both the UV and X-ray data, allowing $R_{\rm hot}$ and $\Gamma_{\rm hot}$ to be free parameters in addition to $\dot{m}$.

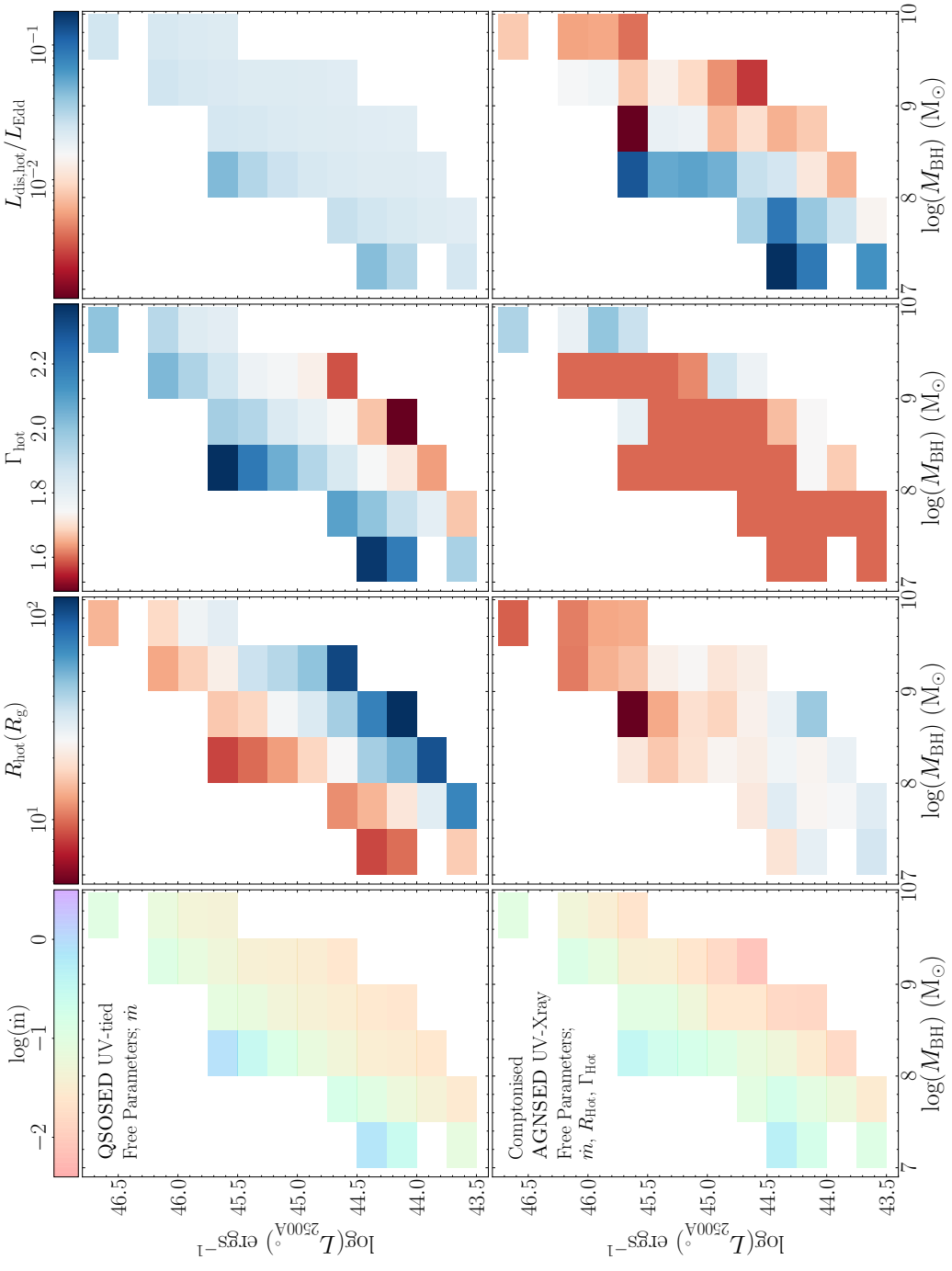

Figure 4.8: A comparison of the parameters of the QSOSED UV-tied fits (upper panel) and those from the AGNSED UV-X-ray fits of Fig. 4.7 (lower panel). Both models have $\dot{m}$ as a free parameter, but the AGNSED UV-X-ray fits additionally have $R_{\rm hot}$, $\Gamma_{\rm hot}$, and X-ray power, $L_{\rm dis,hot}/L_{\rm Edd}$, as free parameters, whereas in QSOSED they are hardwired into the model.

are no data.

Compared to QSOSED, this model gives a much better fit to the high $M_{\rm BH}$, but a worse fit to the lower $M_{\rm BH}$. Figure 4.8 shows the resultant $\dot{m}$, $R_{\rm hot}$, $\Gamma_{\rm hot}$, and $L_{\rm disc,hot}/L_{\rm Edd}$ values for this model, compared to the same parameters for the original QSOSED UV-tied model. The region where $\Gamma_{\rm hot}$ pegs to the minimum value clearly shows the parameter space where the fits are least convincing.

4.5 The shape of the optical/UV spectrum from SDSS

4.5.1 SDSS composites

We can make an independent check of the shape of the optical/UV spectrum using composites from a much larger sample of AGN in the SDSS DR14 quasar catalogue (Pâris et al., 2018). This contains 526356 sources each possessing a publicly available optical spectrum.

Source Selection and composite creation

We derive spectral composites which span a similar range in wavelength to our SDSS-OM bandpass by selecting objects with the same mass and 2500Å luminosity at both low and high redshift. We use a redshift cut at $z \leq 0.8$ and $2.15 \leq z \leq 2.25$ for H β and Mg II respectively. We remove any source with a poor quality flag in z, L_{3000} or $M_{\rm BH}$ in addition to removing any source flagged as a BAL. In order to remove noisy spectra from the composites we carry out a signal to noise cut, removing any source with a continuum signal to noise ratio < 5. This selects only ~ 4 % of the entire SDSS DR14 quasar catalogue

Each spectrum was de-reddened using the ASTROPY extinction module (Astropy Collaboration et al., 2018) assuming a CCM89 extinction profile (Cardelli et al., 1989) and an $R_V = 3.1$. The $E(B-V)$ measurements for each source were derived from the Schlafly & Finkbeiner (2011) dust maps and sourced from the IRSA database. The de-reddened spectra were resampled onto a uniform wavelength grid with a resolution of 5Å, converted to νL_ν and shifted into the restframe.

Here and throughout, as with L_{2500} at 2500Å, L_{3000} refers to the monochromatic luminosity at 3000Å in units of erg s^{-1}. The spectra were binned on L_{3000} and $M_{\rm BH}$ according to the values quoted in R20. The (median$\pm 0.5\sigma$) discrepancy between L_{2500} and L_{3000} measured in

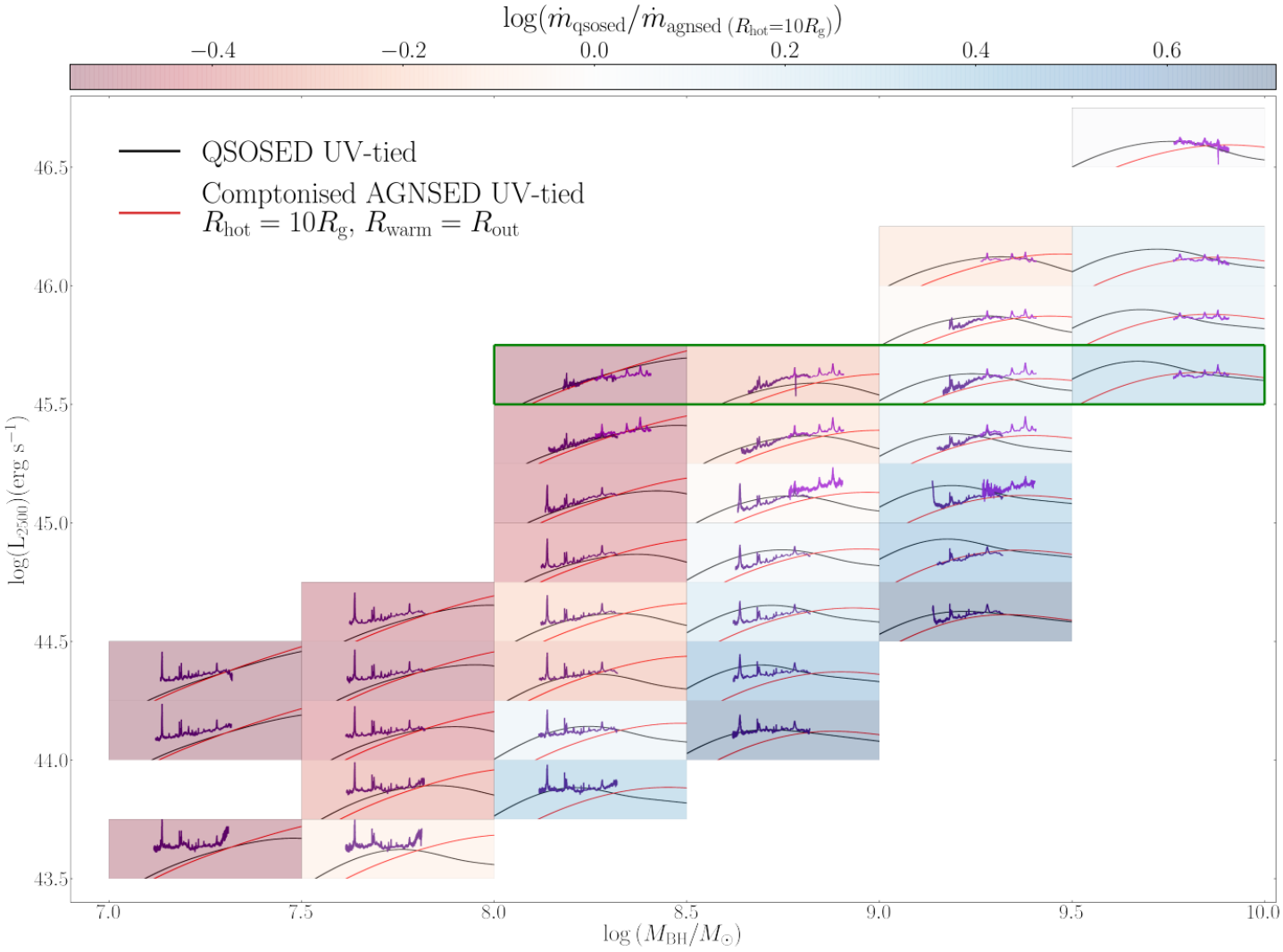

Figure 4.9: SDSS spectral composites from the full Rakshit DR14 SDSS quasar catalogue (Rakshit et al., 2020) with z≤0.8 (indigo) and (2.15≤z≤2.25) (purple). Each bin contains the QSOSED UV-tied SED (black) and AGNSED UV-tied Comptonised SED (red), taken from figures 4.4 and 4.5 respectively. Each bin is coloured on the logarithm of the ratio of $\dot{m}$'s computed by the QSOSED and the Comptonised AGNSED models. Each of the inset plots span 1.5 dex in luminosity and cover the spectral range between 2×10^{14}Hz and 4×10^{15}Hz.

the SOUX AGN sample was (0.05±0.05) dex, so this small shift in wavelength will not affect the findings.

We calculated composite mean spectra for the high and low redshift samples separately, using the weighted geometric mean and weighted standard error at each wavelength element on the uniform wavelength grid. The composites for each of the bins populated by the SOUX AGN sample are shown in Figure 4.9. The whole grid is shown in Appendix: B.2 (Figures B.2 and B.3). Many of the lower mass bins only contain low redshift spectra due to the lack of high redshift sources in this mass range, the opposite effect is true for some of the higher mass bins in which there does not exist any low redshift sources.

Unlike the SOUX sample, Figures 4.4-4.8, we have not attempted to remove very radio loud sources as not all the SDSS spectra have L_{4400}, so this cannot be done systematically, but this will only be a very small fraction of the sources, so their effect on the composites should also be small.

Comparison with QSOSED and AGNSED

Figure 4.9 shows the same grid as seen in Figure 4.1, however each inset plot shows the average QSOSED (UV-tied, black), the average AGNSED model (UV-tied, red), and the SDSS composite spectra (low z: indigo, high z: purple) relevant to each bin.

The models plotted in Figure 4.9 have not been exposed to the SDSS data meaning that the SDSS composite spectra provide a powerful diagnostic tool to assess the accuracy of our fits in both flux level and shape, and their relevance to the wider AGN population.

It is clear that, in general, the composite SDSS data from a wider sample of AGN are similar to the SDSS-OM spectra from the SOUX sample. At lower redshifts we see spectral shapes much redder than the models which is consistent with our own SDSS spectra and UV data points, an effect that we attribute to host galaxy contribution to the spectra. In general, the composites show a better agreement with the QSOSED (black) models over the AGNSED (red) models at low to medium mass, but then shift towards the shape of our AGNSED (red) models at higher masses.

The trends seen in the optical/UV SDSS spectra shown here are similar to those seen in our SOUX sample. This shows that there is no significant selection bias in our sample introduced by the requirement for ≥ 250 counts in *XMM-Newton*, and the removal of extremely loud radio sources (Radio Loudness > 100).

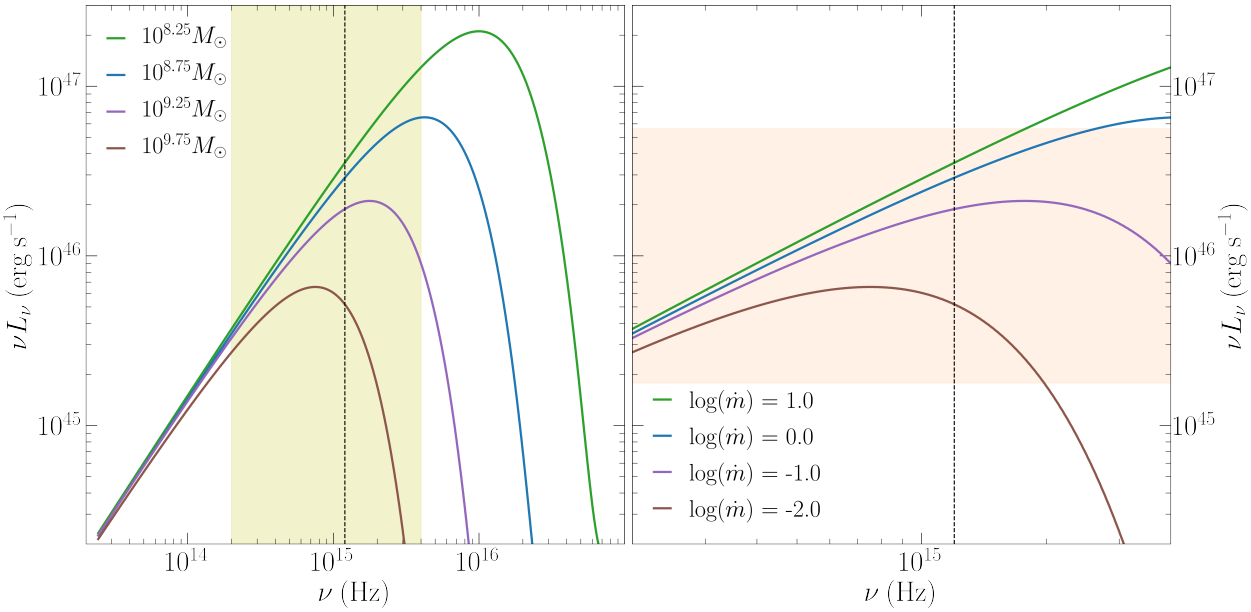

Figure 4.10: **Left-hand Panel:** Pure disc SEDs (spin 0) with $M_{\rm BH}$ centered on the four mass bins enclosed in green in Figure 4.9. The spectra are shown over 3 decades in $\dot{m}$, the necessary range in order to match in Luminosity on the Rayleigh-Jeans tail over a span of 1.5 dex in $M_{\rm BH}$. However, the spectra clearly peak in the SDSS bandpass (highlighted in yellow), so these do not all have the same L_{2500}. **Right-hand Panel:** Zoom-in of the SDSS bandpass, with orange highlighting the range in νL_ν of the insets in Figure 4.9. Disc spectra clearly predict that that there should be strong evolution in the spectral shape in this range.

4.5.2 Changing mass at fixed UV luminosity

The spectral shape of the optical-UV continuum from a pure disc remains constant on the Rayleigh-Jeans part of the disc, with monochromatic luminosity $\nu L_\nu \propto (M^2\dot{m})^{2/3}$. However, we do not expect to be on the Rayleigh-Jeans tail in the rest frame UV for the wide range of masses sampled here. The green box in Figure 4.9 shows we sample a mass range of at least 1.5 dex at a constant monochromatic luminosity, so the standard disc would have to change in $\dot{m}$ by 3 dex. Even if the lowest mass bin was at log($\dot{m}$)=1, the highest would have to be at log($\dot{m}$)=-2, around the changing look transition. The pure disc models at such low mass accretion rates and high masses are predicted to peak at far too low a temperature to make the UV emission seen here. The drop in disc temperature is even more obvious in QSOSED as this has the inner disc progressively replaced by the hot flow as $\dot{m}$ drops (see Figure 4.2). Thus we expect that there should be a clear change in shape for the optical/UV spectra at the same L_{3000} for changing mass (see Fig. 4.10).

We look for this effect by comparing the spectra at constant L_{3000} for different masses where there are 3 or more mass bins in Figure 4.9. These are shown in Figure B.1. Very surprisingly, the optical/UV spectra are almost identical. This is most clearly the case for the highest

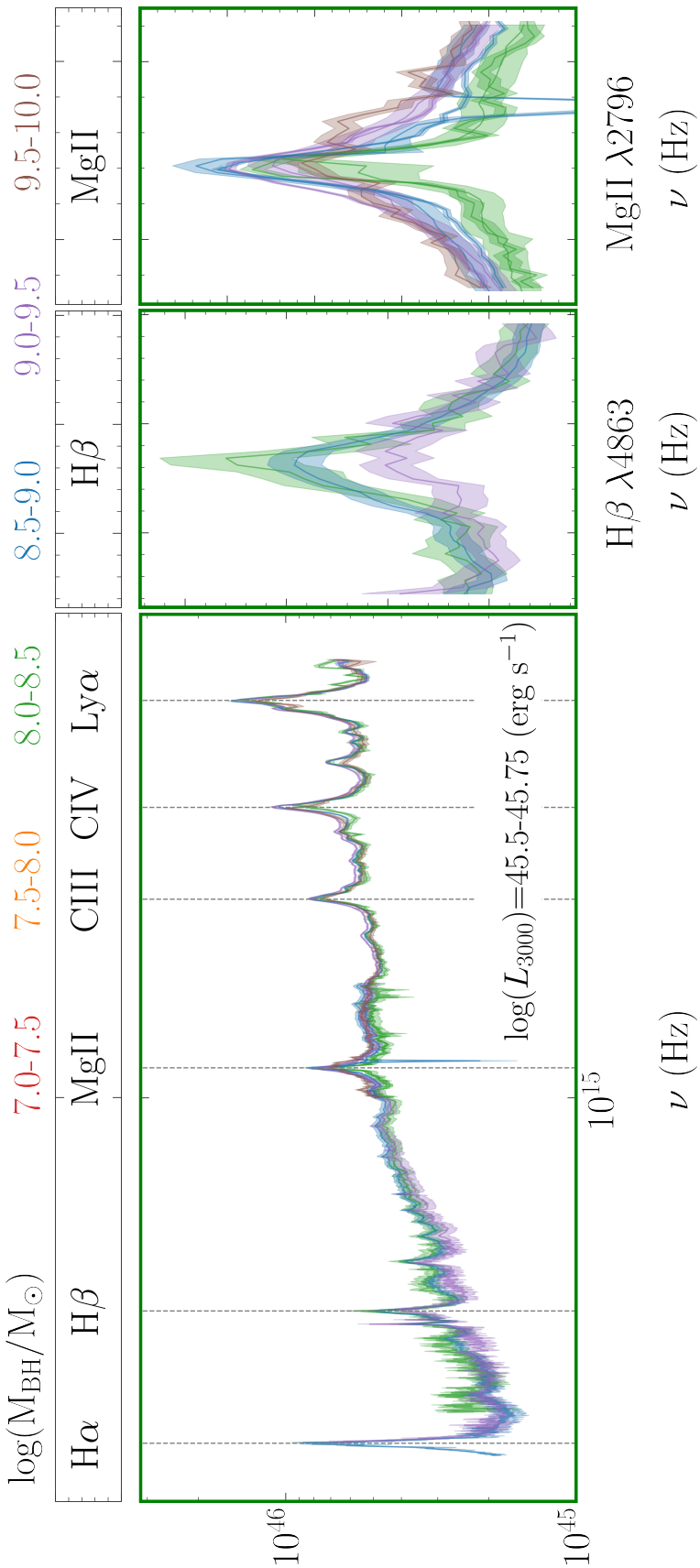

Figure 4.11: The L_{3000} = (45.5-45.75) erg s^{-1} bin from Figure 4.9 showing the unchanging spectral shapes across the mass ranges. The right hand panels show the H β and Mg II emission line regions. A separate Mg II profile is for shown for both the low and the high z composite in any L_{3000} with overlap.

Table 4.1: Mass estimates of composite spectra in the $\log(L_{2500})$=(45.5-45.75)erg s^{-1} bin, obtained using PyQSOFit and the scaling relations of Mejía-Restrepo et al. (2016).

$\log(M_{BH}/M_{\odot})$ Range	FWHM (km/s)	$\log(L_{5100})$	Mass ($M_{\odot}$)	$\log(M_{BH}/M_{\odot})$
8.0-8.5	2597 (H β)	45.36	3.97×10^{8}	8.47
8.5-9.0	3815 (H β)	45.31	5.90×10^{8}	8.77
9.0-9.5	6610 (H β)	45.32	1.80×10^{9}	9.25
9.5-10.0	8161 (Mg II)	45.61	5.53×10^{9}	9.74

luminosity bin, $L_{3000} = (45.5 - 45.75)(\mathrm{erg\,s^{-1}})$, which contains spectra from four mass bins from $\log(M_{BH}/M_{\odot})$=(8.0-10.0) (shown in Figure 4.11 and enclosed in green in Figures 4.9, B.1, B.4 and B.5). The line widths of H β and Mg II clearly show broader line profiles in the higher mass bins, as expected for the change in mass, but the continuum shape shows no significant change.

Appendix: B.2 shows that this trend continues for the whole SDSS sample, not just the range covered by the SOUX sample (Figure B.4). It is clear that even in this wider parameter space, the shape of the optical-UV continuum does not vary significantly with M_{BH} at a fixed luminosity. We also check that the same spectral shape is seen in individual objects by comparing our composite spectra with a sample of individual AGN observed by X-shooter (Capellupo et al. 2015; Fawcett et al. 2022, data by private communication). This instrument covers an extremely wide bandpass, similar to our composites. Figure B.5, shows these composites plotted over the relevant bins from Figure 4.11. Again we see no significant differences between any of the spectra across the entire mass range at a given luminosity.

To check that, when measured, the mass of each composite matches the masses of the sources from which each composite was constructed, we fit the composite spectra presented in Figure 4.11 with PyQSOFit and derive a black hole mass for each using the scaling relations of Mejía-Restrepo et al. (2016). The results from this fitting procedure are presented in Table 4.1. All of the mass estimates fall within the bounds of the mass bin for which each composite was constructed. This allows us to perform detailed fitting of these composite spectra using the central mass of each bin during the fitting procedure.

4.5.3 Detailed fits to the SDSS composites and fine-tuning

Since we have only the optical/UV spectra, we now go back to the standard (and colour temperature corrected) disc models (Section 3.3 and 3.4) to see if these can fit better to this restricted energy range. We fit disc spectra to this unchanging spectral shape across the M_{BH}

range depicted in Figure 4.11. This is the approach often taken with individual objects, but here we use the composites taken by averaging over many objects, which highlight the fine tuning issues.

Fitting with a_* as a free parameter

Firstly we explore fitting pure disc models, allowing black hole spin, a_*, to be a free parameter as well as $\dot{m}$. We take the four mass bins from the $\log(L_{3000})$=(45.5-45.75) (erg s^{-1}) row highlighted in green in Figure 4.9 and fit a pure disc model using AGNSED (f_{col}=1). It is very clear that highest masses require maximal spin in order to make the disk peak at high enough energy to fit the shortest wavelengths sampled here.

We repeat this with the colour temperature corrected disc, ($f_{col}(T)$), as incorporated in OPTXAGNF. The same shift to high spin is apparent, but this model now fits the curvature seen in the data, where there is a systematic flattening of the spectra at the shortest wavelengths. This is because the colour temperature correction from electron scattering onset is when the disc temperature exceeds 3×10^4 K. Annuli above this temperature are shifted relative to annuli at lower temperatures, producing a characteristic bend in the spectrum at a specific wavelength.

It is conceivable that there are real physical mechanisms that cause higher mass black holes to have high spin. However the fine tuning of a_* and $\dot{m}$ that is necessary to allow the UV-optical continuum shape to remain constant across 2.5 dex in black hole mass, as is shown in Figure 4.11, seems contrived.

Additionally, none of these models are self consistent as while they incorporate general relativistic effects on the intrinsic emissivity (Novikov-Thorne, see section 3), they do not include relativistic ray tracing on the observed spectrum. There is a strong gravitational redshift expected for high spin, which is much more important than the Doppler blueshift from the fast orbital motion at low inclinations. All these effects are much weaker for low spin. We use RELAGN (Hagen et al, 2022 in prep) to incorporate both the general relativistic effects on the emissivity and on the radiation transport to the observer. Figure 4.12 shows what happens to the pure disc ($f_{col} = 1$) and to the colour temperature corrected disc, $f_{col}(T)$. There is now no adequate fit for the highest mass bin. Increasing black hole spin apparently fit the data as this gave increased high energy emission by decreasing the inner disc radius but these smallest radii are the most affected by gravitational redshift, offsetting all the gain in far UV emission.

We also fit these four disc regimes allowing for intrinsic reddening as a free parameter using the XSPEC model ZDUST. The resultant $E(B-V)$ values are only significant for the lowest two mass bins of $\log(M_{\rm BH}/M_{\odot})$=(8.0-8.5) and $\log(M_{\rm BH}/M_{\odot})$=(8.5-9.0) and for the pure disc ($f_{\rm col}$=1) for the $\log(M_{\rm BH}/M_{\odot})$=(9.0-9.5). For all other disc regimes and for all of the disc regimes in the highest mass bin $\log(M_{\rm BH}/M_{\odot})$=(9.5-10.0), the resultant $E(B-V)$ values are insignificant. This strongly indicates that intrinsic reddening does not help in fitting disc spectra to the optical-UV continua of high mass AGN. This is as expected as the high mass spectra are far more blue than predicted by the disc models.

Fitting a standard disc with minimal and maximal spin

We illustrate the issues above by plotting the models over a wider energy range for the highest and lowest mass bins. Figure 4.13 shows a comparison of minimal and maximal spin models for the four disc models shown in Figure 4.12 for the $\log(M_{\rm BH}/M_{\odot})$=(8.0-8.5) and $\log(M_{\rm BH}/M_{\odot})$=(9.5-10.0) mass bins at fixed $\log(L_{3000})$=(45.5-45.75) (erg s^{-1}).

In the lower mass bin of $\log(M_{\rm BH}/M_{\odot})$=(8.0-8.5) all four disc models are able to give a reasonable fit to the spectral shape of the optical-UV continuum. In this mass range, with a_*=0, the relativistic correction makes very little difference to the SED shape, the colour temperature correction has far more impact and shifts the peak to higher energies in a similar fashion to Comptonisation. Here the colour temperature corrected fits, dashed lines, provide a better overall fit to the data, and have lower $\dot{m}$ values.

When a_* is fixed at 0.998, maximal spin, the colour temperature corrected models again show better fits to the data. However, the overall spectral shape of the SED is greatly changed, with the relativistic corrections having a much greater effect than the colour temperature correction. The effect of having maximal spin is to drag the peak of emission to much higher energies, this is significantly counter acted however when general relativity is taken into account and the peak of emission is shifted towards lower energies. This occurs due to the increased importance of relativistic redshift on the emission as opposed to blue shift at high spin, whereas at low spin these factors are much more balanced.

At high mass the picture is very different. With $\log(M_{\rm BH}/M_{\odot})$=(9.5-10.0) and a_*=0, the SED shape for all four disc regimes is almost identical, and plainly does not match the data. At these high masses and zero spin, general relativity has almost no effect, other than to slightly increase the $\dot{m}$ value for a given L_{3000}. A common approach at these masses is to assume high spin. The bottom right-hand panel of Figure 4.13 shows that a pure disc at

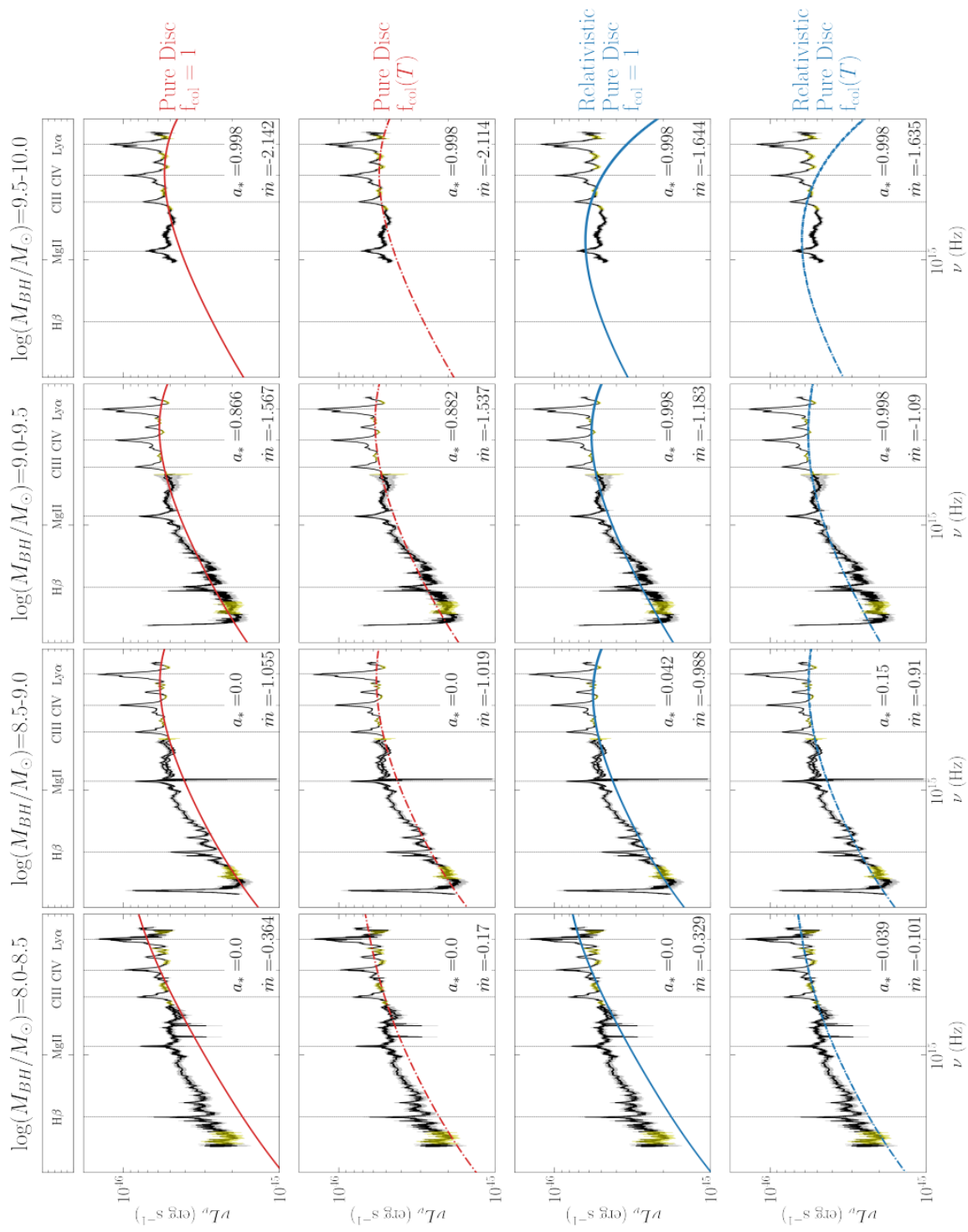

Figure 4.12: Four mass bins, taken from Figure 4.9, of constant mean $\log(L_{3000})$=(45.5-45.75)(erg s^{-1}), spanning 2 dex in $\log(M_{\rm BH}/M_{\odot})$ from 8.0-10.0, fit with AGNSED, OPTXAGNF, RELAGN and RELAGNF, where RELAGNF is RELAGN with a colour temperature correction applied. Each model was fit for $\dot{m}$ and $a_*$to the clean continuum regions highlighted in yellow, the resultant value of which are displayed on each panel.

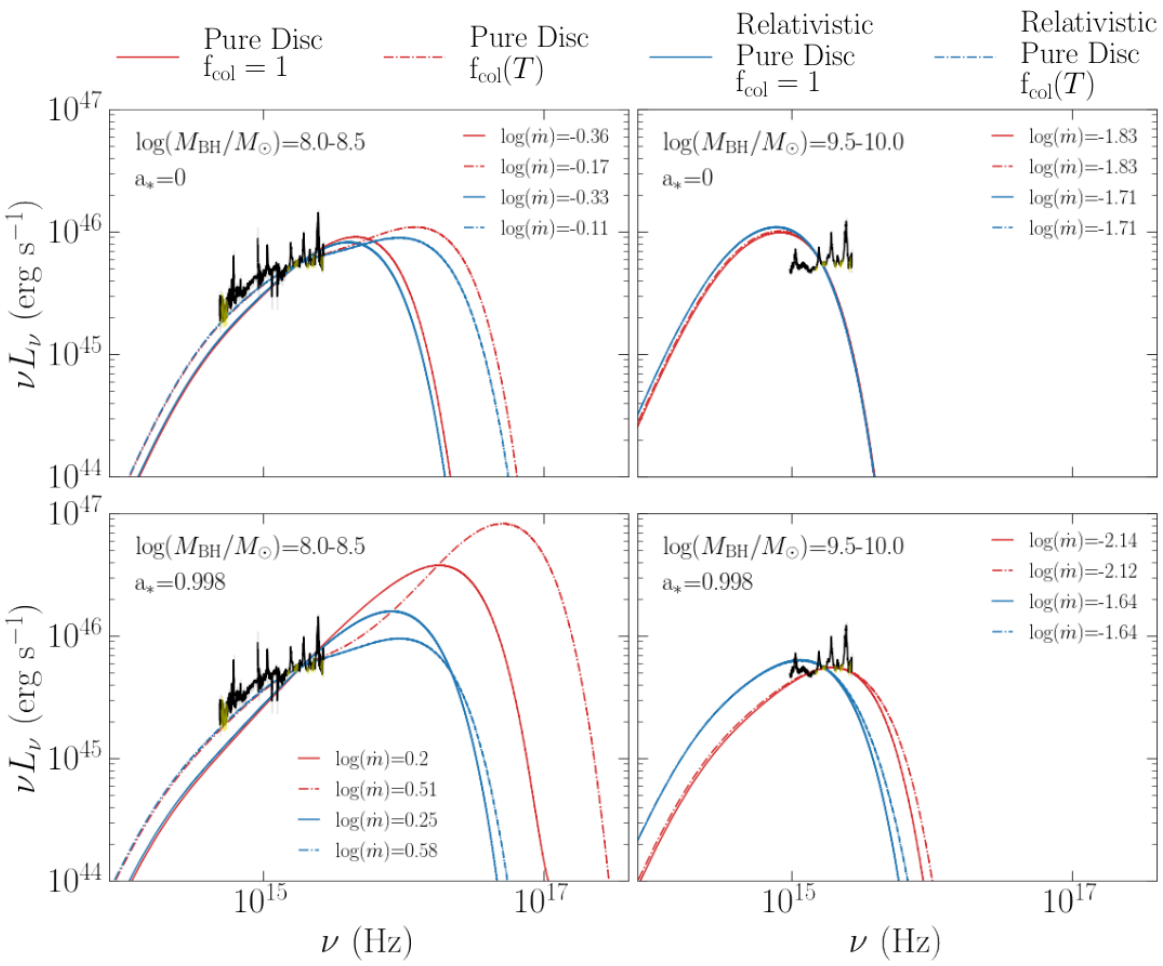

Figure 4.13: The SDSS composites for constant $\log(L_{3000})$=45.5-45.75 (erg s^{-1}) for mass of $\log(M_{BH}/M_\odot)$=(8.0-8.5) (left) and $\log(M_{BH}/M_\odot)$=(9.5-10.0) (right). Yellow points indicate the continuum dominated regions used for the fits. Models are a pure disc (red solid), colour temperature corrected disc (red dashed). The blue solid and dashed lines indicate the same models but with the self consistent ray tracing. This does not have a large effect for spin 0 (upper panels), but is very important for high spin (lower panel), where gravitational redshift means that no disc model with dissipation in the midplane can fit the data.

maximal spin, with or without a colour temperature correction matches the spectral shape of the high mass composite well. However once general relativity is taken into account, all that is gained in shifting the peak to higher energies by maximising spin is lost again, as the peak of emission shift back to lower energies and once again does not fit the data well.

This shows clearly that there is no way to make the optical/UV spectra seen from the highest mass AGN from a standard (or color temperature corrected) disc for these masses and mass accretion rates. Even if all the energy is dissipated in a disc (with the hard and soft X-rays powered by e.g. a separate coronal flow or by tapping the spin energy of the black hole) the UV emission extends to higher energies than predicted for the accretion disc peak in full general relativity. (Figures 4.12 and 4.13).

Fitting a fully Comptonised disc with minimal and maximal spin

A blackbody disc cannot fit the highest masses, even with colour temperature correction as expected from a standard disc photosphere. Instead we explore the viability of fitting fully Comptonised disc spectra, as suggested by Petrucci et al. (2018), which implies that the energy is dissipated away from the midplane, towards the photosphere. Since the SDSS data do not cover the soft X-ray excess, we fix its electron temperature at 0.2 keV.

We fit a fully Comptonised disc ($R_{\rm warm}$=$R_{\rm out}$, $R_{\rm hot}$=$R_{\rm isco}$) for $\dot{m}$ with $\Gamma_{\rm warm}$=2.5, and a Comptonised pure disc fit with both $\dot{m}$ and $\Gamma_{\rm warm}$ as free parameters, to the $\log(M_{\rm BH}/M_{\odot})$=(8.0-8.5) and $\log(M_{\rm BH}/M_{\odot})$=(9.5-10.0) bins, at fixed L_{3000}=(45.5-45.75)(erg s^{-1}). We repeat this process for the same disc models but with the addition of full general relativity. The results of this fitting process are shown in Figure 4.14.

In the low mass regime, both at zero and maximal spin, the optical-UV continuum shape of the composite is too red to fit well to any of the Comptonised disc spectra. The inclusion of general relativity does not make a difference to the fit to the data, however it does change the overall shape of the SED and reduces $L_{\rm bol}$ dramatically. The same is true when fitting to $\Gamma_{\rm warm}$, which in both the zero and maximal spin regimes results in the maximum allowed value of $\Gamma_{\rm warm}$=3.5.

In the high mass bin, Comptonisation provides a better fit to the data. In the spin zero models, there is very little difference between the Comptonised pure discs with and without general relativity, and the best fit $\Gamma_{\rm warm}$ pegs at the minimum value of 2.1 though there is more dispersion for the Comptonisation shape at high spin.

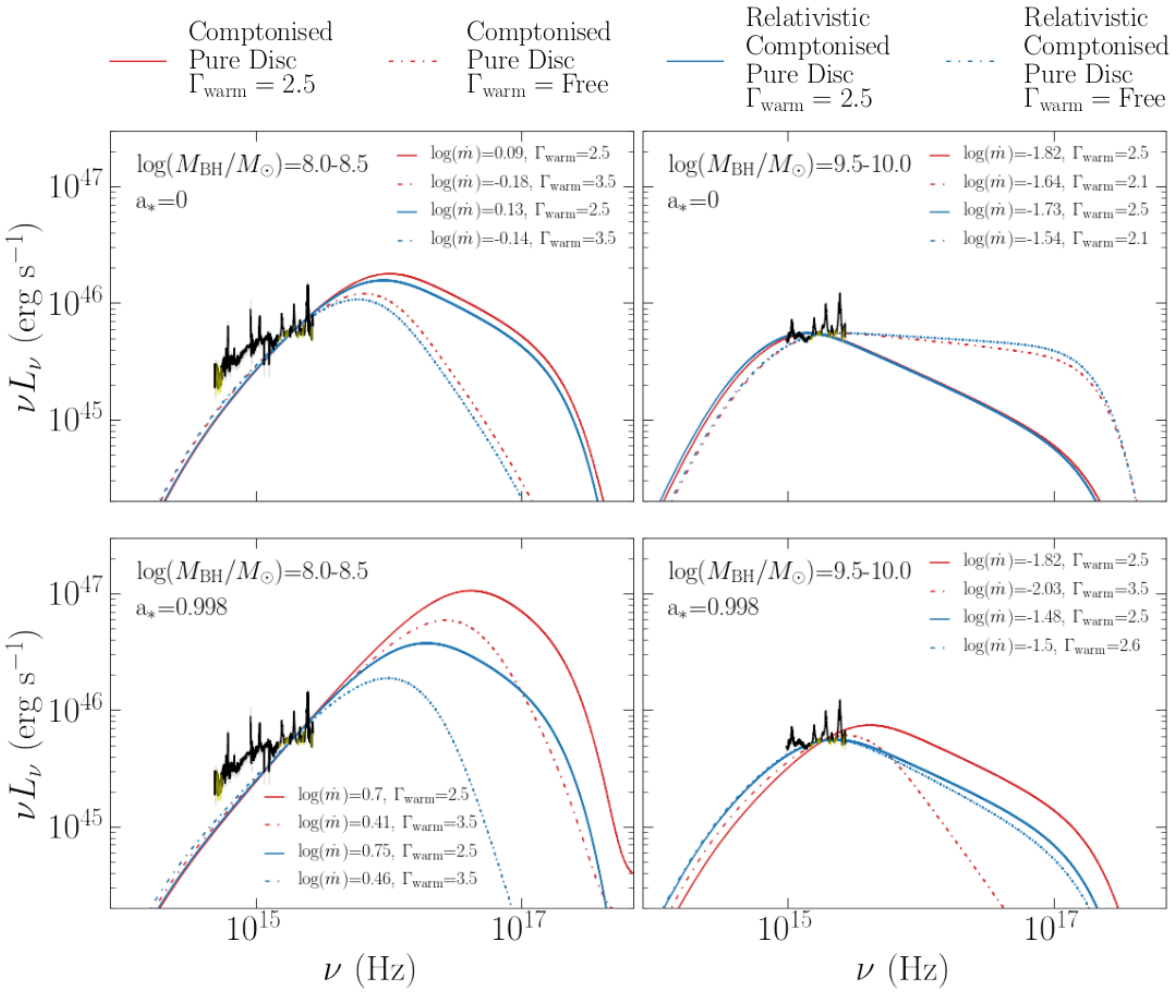

Figure 4.14: The same as Fig 4.13, but with Comptonised disc models. The solid red lines have $\Gamma_{\rm warm}$=2.5 (as in QSOSED), while the red dashed lines have $\Gamma_{\rm warm}$ as a free parameter. Again, models in blue include the self consistent ray tracing.

Comptonisation does not solve all the issues, there is clearly a need for $\Gamma_{\rm warm}$ to change as a function of $M_{\rm BH}$ or perhaps more likely physically with $\dot{m}$. However, it does at least allow the models to fit to the furthest UV emission observed in the data. The key issue is our (lack of) understanding of the physics which sets this emission component.

4.6 The structure of the accretion flow

The QSOSED model is a good zeroth order predictor of the intrinsic AGN spectrum (Figure 4). In particular, it is clear that the assumptions made about the hot Comptonisation region match fairly well to the X-ray luminosity and spectra seen in the 2-10 keV bandpass across most of the range of mass and $\dot{m}$ in our sample. This favours the underlying assumption that the hot corona dissipates up to the ADAF maximum of $L_{\rm diss} = 0.02 L_{\rm Edd}$. For accretion flows below this limit, the entire accretion power is dissipated in the hot corona rather than in a UV bright disc. This means there is no longer a strong UV ionising flux so no strong BLR, leading to an optical classification as a LINER or true type 2 Seyfert. This gives the observed mass dependence of the lowest luminosity spectra identified as QSOs in SDSS.

The outer disc extends inwards as the source luminosity increases above the ADAF limit, but unlike the stellar mass BHB at the same $L/L_{\rm Edd}$, the disc is still truncated so that there is substantial hard X-ray power with $L_{\rm diss} = 0.02 L_{\rm Edd}$. In the QSOSED model, this requirement sets the extent of truncation of the optically thick material, $R_{\rm hot}$. The inner disc only extends down close to the ISCO for $L \gtrsim 0.1 L_{\rm Edd}$, giving the changing fraction of disc to X-ray power seen in the data.

Similarly, the warm Comptonisation region in QSOSED is a fairly good match to the soft X-ray/UV data at zeroth order despite the assumptions underlying this component being completely phenomenological. In particular, the assumption that the radius of the warm Comptonisation region is $R_{\rm warm}$=$2R_{\rm hot}$ means that the extent of this component is limited at high $L/L_{\rm Edd}$ where $R_{\rm hot}$~$R_{\rm isco}$, whereas it covers most of the outer disc at $L/L_{\rm Edd}$~ 0.02. The condition $R_{\rm warm}$=$2R_{\rm hot}$ equivalent to assuming that the soft Comptonisation carries a more or less fixed fraction of the bolometric luminosity, with $L_{\rm warm} = 0.2 L_{\rm bol}$.

The QSOSED component which is the worst fit is the one which is best motivated physically, namely the standard disc emission assumed for the outer radii. For the highest masses, and lowest luminosities, the disc temperature should result in a peak below 1200Å even for an untruncated disc, let alone the quite strongly truncated disc assumed by QSOSED at low

$L/L_{\rm Edd}$. This mismatch is present in the literature but is disguised as individual objects seem fairly well fit using extreme black hole spin. This gives a higher temperature disc peak from the smaller $R_{\rm isco}$ size scales. However, these models only incorporated the effects of general relativity on the disc emissivity (Novikov-Thorne, as used here), but did not include its effect on ray tracing from the disc to the observer (Capellupo et al., 2016). There is gravitational and transverse red shift which depends on R, whereas the Doppler red and blue shifts from orbital motion depend on both R and inclination. These shifts are generally negligible for the optical/UV spectra of QSO with masses $\lesssim 10^9 M_\odot$, as the radii emitting these wavelengths are at $R \gtrsim 10R_{\rm g}$. However, for the most massive QSO the observed UV is produced close to $R_{\rm isco}$. Pure disc models with low spin do not have sufficiently high temperatures to match the observed UV. High spin models appear able to do this as the intrinsic disc temperature is higher when including the additional inner disc radii from R=(6-1.23)$R_{\rm g}$. However, this higher temperature emission is strongly gravitationally redshifted on its way out to the observer, which reduces the observed temperature back to something close to the low spin disc models.

Pure disc models are then not able to match the observed outer disc emission at the highest masses once the self consistent ray tracing is included. Either the masses are wrong, or the standard disc models are wrong (or both). We examine each of these in turn below.

4.6.1 Changes to the standard disc emission

We saw above that the models where the entire outer disc emitted a warm Comptonisation spectrum rather than (colour temperature corrected) blackbody enabled the accretion power to reach the highest UV energies observed. However, it requires a correlation of $\Gamma_{\rm warm}$ with $L/L_{\rm Edd}$, such that the lowest $L/L_{\rm Edd}$ AGN have hardest $\Gamma_{\rm warm}\sim$ 2 while the highest $L/L_{\rm Edd}$ have $\Gamma_{\rm warm}\sim$ 3.5. This is certainly supported from observational data, as multiple high $L/L_{\rm Edd}$ NLS1 show $\Gamma_{\rm warm}\sim$ 3.5 and this can be produced in the theoretical models when the underlying disc is not purely passive (Petrucci et al., 2018). However, this is completely dependent on the unknown structure of the warm Comptonisation region, and detailed models require much more understanding of this regime. Nonetheless, the warm Comptonisation models are the only disc based spectra which can match the data from the highest mass QSO at lower luminosity.

Alternatively, the accretion structure could be completely different to that expected from a disc. We based even our non-standard disc models on the standard Novikov-Thorne emissivity

Table 4.2: Three single epoch virial black hole mass estimates calculated using spectra taken from Boroson & Green (1992); Buttiglione et al. (2009); Torrealba et al. (2012), along with the measurement from Gravity Collaboration et al. (2018).

Source	Mass ($M_\odot$)
Boroson & Green (1992)	6.79×10^8
Buttiglione et al. (2009)	1.22×10^9
Torrealba et al. (2012)	2.51×10^8
Gravity Collaboration et al. (2018)	$(2.6 \pm 1.1) \times 10^8$

which assumes that the α viscosity remains constant across the disc. Yet it is starting to become clear that hydrodynamic convection couples to the magnetic dynamo, strengthening the heating in convectively turbulent regions (Jiang et al., 2016b; Jiang & Blaes, 2020). Hydrodynamic convection could be triggered by the strong bumps in opacity at 10^4 K from H, $10^{4.6-4.8}$ K from He and $10^{5.2}$ K from iron (a blackbody at 10^4 peaks at 3000Å). Massive stars have similarly bright UV emission, and these also show strong winds, powered by UV line driving, as well as turbulent convection triggered by the continuum UV opacity (Jiang et al., 2015; Ro, 2019). Alternatively, convection could be triggered by the radiation pressure instability. This becomes more dominant at highest $\log(M_{\rm BH}/M_\odot)$ though the expected dependence is rather slow, with $R_r/P_g = 7.3 \times 10^9 m_9^{1/4} \dot{m}^2 \alpha^{1/4} r^{-21/8}$.

Either warm Comptonisation of the outer disc, or a different accretion structure potentially can work by shifting the emission to higher temperatures, so it peaks in the unobservable EUV part of the spectrum. The observed UV/optical emission is instead produced by reprocessing of this EUV peak from dense material. The BLR must subtend a substantial solid angle, and dense clouds within the BLR can give a diffuse continuum (Korista & Goad, 2001, 2019; Cackett et al., 2018). There is also growing evidence for a dense clumpy wind on the inner edge of the BLR which again produces predominantly diffuse continuum (free-free and bound-free rather than lines) (Miniutti et al., 2014; Kaastra et al., 2014). A predominantly reprocessed origin for the optical/UV is the easiest way to explain its remarkably constant shape, as well as fitting in with the longer timescale lags seen in the continuum reverberation intensive monitoring campaigns.

4.6.2 Testing black hole masses

The unexpectedly constant shape of the optical-UV continuum that we observe could potentially indicate that the mass scaling relations are inaccurate at high $M_{\rm BH}$. We test the robustness of the single epoch virial mass estimates at high black hole masses using single

epoch spectra from from 3C273, comparing them to the mass derived from the independent method of spatially resolving the broad line region Gravity Collaboration et al. (2018). We obtained three independent optical spectra of 3C273 spanning 20 years from Boroson & Green (1992); Buttiglione et al. (2009); Torrealba et al. (2012). We use the PyQSOFit software package to measure the FWHM of the H β line profile, and the 5100Å continuum luminosity, and derive mass from the scaling relations of Mejía-Restrepo et al. (2016) as for the SOUX AGN sample. These results are shown in Table 4.2, and show a spread of 0.69 dex, with a mean mass which is higher than the Gravity result. Thus there could be a systematic overestimate of black hole mass, which could shift a source up between 1-2 of our 0.5 dex bins. This could give a systematic effect as there are many more lower mass black holes than high mass ones, so numbers scattered to higher masses are not compensated by number of higher mass objects scattered down (e.g. Davis et al. 2007).

However the constant shape of the optical-UV continuum is observed across 1.5-2 dex (4 bins) in $M_{\rm BH}$ in the SOUX sample (see Figure 4.11) and 2-2.5 dex (5 bins) in $M_{\rm BH}$ in the wider SDSS parameter space (see Figure B.2). This would require an extremely large systematic shift in the mass estimation that seems unlikely given the number of sources used in the creation of the composite spectra. Such a shift would have consequences for all black hole mass estimates.

Based on this preliminary study, it is possible that there is a systematic overestimation of $M_{\rm BH}$ at high masses, but this is not likely to be large enough to be the dominant factor causing the homogeneous optical-UV continuum shape that we observe. As the number of AGN with a spatially resolved broad line region increases, a more robust test of single epoch virial black hole mass estimates will hopefully become possible.

Aside from the possibility that the scaling relations are inaccurate, the $M_{\rm BH}$ values that we adopt from R20 could be contaminated, particularly where Mg II is utilised. This seems unlikely however, given that R20 perform a subtraction of the UV emission by fitting a velocity-broadened template, and that we perform a S/N cut. Even if a systematic overestimation in $M_{\rm BH}$ due to some kind of contamination existed in the masses quoted by R20, it would doubtful be large enough to cause the shape of the optical-UV continuum to remain constant over 2-2.5 dex.

For the SOUX sample, there is a $\approx 60\%$ overestimation in the Mg II masses with respect to H β masses in sources with both lines present. These masses were calculated in K22 using the scaling relations from Mejía-Restrepo et al. (2016). This only corresponds to 0.2 dex, an

offset of too small a magnitude to cause the effect we see over 2-2.5 dex in $M_{\rm BH}$.

4.7 Summary and Conclusions

We demonstrate the advantages of studying a sample large enough to investigate population statistics, but based on available high quality multi-wavelength data. This enables us to carry out detailed SED analysis over a wide parameter space rather than for individual objects.

We stack spectra from our sample on a 2D grid in $M_{\rm BH}$ and L_{2500}. All AGN in a single grid point should have the same black hole mass range and mass accretion rate (including uncertainties in black hole mass estimates, and inclination angle) so that their spectra can be averaged with confidence that the objects are similar. We compare the stacked SED in each grid point with the recent AGN SED model of KD18, QSOSED. This model consists of an inner hot flow with fixed X-ray heating of $L_{\rm x} = 0.02 L_{\rm Edd}$ (the maximum ADAF luminosity). This condition determines the inner radius of the optically thick disc, $R_{\rm hot}$, and thermalisation is assumed to be incomplete out to $R_{\rm warm}$=$2R_{\rm hot}$, making the soft X-ray excess connect to the UV downturn from warm Comptonisation, before thermalising to a blackbody from $R_{\rm warm}$ to $R_{\rm sg}$.

To a reasonable degree, the model in each grid box matches well the observed stacked SED across the entire range of mass accretion rates for intermediate masses $\log(M_{\rm BH}/M_{\odot})= 7.5-9.0$. This is despite it not being fit to the optical or X-ray data. This is quite strong evidence that the underlying assumptions in QSOSED are a fairly good description of the accretion flow, although to first order the X-ray flux is underpredicted by a factor 2 at the lowest $M_{\rm BH}$. However, at the highest masses, $\log(M_{\rm BH}/M_{\odot})$>9, there is a clear discrepancy in the shape of the optical-UV continuum predicted by the QSOSED models. This is very surprising, as this is the part of the spectrum dominated by the thermal disc, which is the part of QSOSED which has a solid theoretical basis. Instead, we remove the thermalised outer disc, so that all of the optically thick disc emission emerges as warm Comptonisation. This component is poorly understood, so we first use the same parameters as in QSOSED, ($\Gamma_{\rm warm}$=2.5, $kT_{\rm e,warm}$ = 0.2 keV) and fix $R_{\rm hot}$=10 (equivalent to $L_{\rm x}$=0.07 $L_{\rm bol}$). This gives a much better fit to the optical-UV continuum at high mass, but now does not match well the lower mass bins which were well fit with an outer standard disc. Allowing $R_{\rm hot}$ to vary does not improve the fits. There is a systematic shift between any disc model (warm Comptonised or thermal blackbody) and the optical-UV spectrum. Either the low $M_{\rm BH}$ can be fit with a thermal outer disc, which misses

the high $M_{\rm BH}$, or the high $M_{\rm BH}$ can be fit with a warm Comptonised disc, which misses the low $M_{\rm BH}$.

We examine this shift in more detail by constructing wide wavelength coverage composite SDSS spectra for each of the SOUX AGN sample grid points. We combine together optical (low redshift) quasar spectra, with mass from H β and UV (higher redshift) quasar spectra with mass from Mg II to produce a stacked spectrum in each grid point of $\log(M_{\rm BH}/M_{\odot})$ and L_{3000}. These optical-UV composites are of higher quality than the original SOUX SDSS-OM composites as the UV is now based on spectra rather than photometry. They show the same shift with respect to the thermal or Comptonised disc models (Figure 4.9). However, from these spectra we now clearly see that the SED peak shift is present in the models, but is not seen in the actual data. Furthermore, at a given luminosity the SDSS stacked spectra remain remarkably constant as a function of $M_{\rm BH}$ (Figure 4.11). All disc models are based on a size scale of the inner disc, irrespective of whether it is Comptonised or not, and this size scale increases as $M_{\rm BH}$ increases. This decreases the inner disc temperature until it peaks in the observable UV bandpass for the highest $M_{\rm BH}$ in our sample. (Figure 4.10). But the data show no sign of this predicted change (Figure 4.11). We demonstrate this by fitting the pure blackbody disc models to spectra at constant L_{3000} across 4 different mass bins (Figures 4.12 and 4.13). The only way to maintain the constant continuum shape is to systematically increase the spin so as to move the expected peak disc temperature above the UV bandpass. While there could conceivably be a physical connection between the most massive black holes and their spins (e.g. Volonteri et al. 2007, Fanidakis et al. 2011, Griffin et al. 2019, Huško et al. 2022) this requirement seems to be fine tuned, especially as we use stacked spectra rather than single objects. More fundamentally, such models are not self consistent as they only incorporate spin on the intrinsic disc emissivity, but do not include the general relativisitc effects of ray tracing from the origin in the disc to the observer. Increasing the spin can succeed in fitting the highest mass spectra as it reduces the size scale of the inner disc, thereby increasing the peak temperature. But this is significantly compensated by the increased gravitational redshift of the emission and so the high spin models cannot fit the UV data.

We explore whether this could be due to mass estimates being systematically biased so that $M_{\rm BH}$ is overestimated at high $M_{\rm BH}$. However, while this is possible, it seems unlikely to account for the magnitude of the effect we find. Therefore we consider that it is more likely the accretion structure is different than that shown by thin disc models, such that the 'disc' always peaks in the EUV bandpass, even at the highest black hole masses. Reprocessing of

this EUV component then gives the constant shape of the optical/UV spectrum (see also Lawrence 2018 for the same idea based on variability).

One possible way to do this is if the outer disc is completely covered by warm Comptonising material. But this also requires that the spectral index $\Gamma_{\rm warm}$ increases with $L/L_{\rm Edd}$, perhaps indicating a larger fraction of power is dissipated in the disc itself.

Instead, there could be a more fundamental change in the accretion flow if the dissipation always peaked in the EUV region, perhaps due to magneto-rotational instability coupling to the hydromagnetic turbulence generated by sharp changes in the opacity (Jiang et al., 2016a; Coleman et al., 2016, 2018; Jiang & Blaes, 2020). We will explore these ideas further in future work.

Chapter 5

Near-IR Spectroscopic Reverberation mapping: Mrk 509

5.1 Introduction

Active Galactic Nuclei (AGN) are powered by mass accretion onto a central Supermassive Black Hole (SMBH), and are the most luminous compact objects in the Universe. They exhibit a wide range of emission characteristics resulting in an AGN zoo of differing classifications. One of the fundamental distinctions between categories is that which separates AGN with both narrow and broad emission lines seen in their optical spectra (Type 1) from those which only display narrow lines (Type 2). Historically unified schemes proposed a basic AGN anatomy consisting of a central SMBH, a thin accretion disc, separate compact broad and extended narrow line regions and an outer dusty toroidal structure, with each of these components existing on different physical size scales. It has been argued, through these unified schemes, that orientation effects resulting in the obscuration of the Broad Line Region (BLR) for certain viewing angles is the physical mechanism separating AGN into type 1 and type 2 sources (Lawrence, 1987, Ant; Urry & Padovani, 1995; Netzer, 2015). Originally the dichotomy between type 1 and 2 AGN was solely explained by these orientation effects. However the predominant view now is more holistic, with differences in emission characteristics attributed to numerous factors, such as mass accretion rate and black hole mass in addition to orientation (Jin et al., 2012a; Kubota & Done, 2018).

Despite this slight paradigm shift, much evidence for the presence of a large dusty obscuring structure, still commonly referred to as the torus for ease, remains. Strong infrared continuum

emission at $\lambda > 1$ μm, thought to be reprocessed optical/UV emission from the accretion disc, and observations of broad emission lines in the polarised spectra of some type-2 AGN (Antonucci & Miller, 1985; Lawrence, Ant; Cohen et al., 1999; Lumsden et al., 2001; Tran, 2003) are prime examples of evidence for the torus. For a small subset of AGN, constraints on the spatial extent of the torus have been placed at (0.1-6) pc through near-IR, mid-IR (Tristram et al., 2009; Pott et al., 2010; Kishimoto et al., 2011; Burtscher et al., 2013; Kishimoto et al., 2013) and sub-mm interferometric observations (García-Burillo et al., 2016; Imanishi et al., 2018) along with high angular resolution mid-IR imaging (Packham et al., 2005; Ramos Almeida et al., 2011). These findings are roughly consistent with a power-law size-luminosity relation $r_{\rm dust} = L^{0.5}$ (Kishimoto et al., 2009, 2011). Modelling the torus has also provided a useful probe of this structure. Through the fitting of infrared SED's, using the CLUMPY torus model (Alonso-Herrero et al., 2011), and spectroscopy, tight constraints have been placed on a range of torus properties. This and other studies indicate that the torus is of a clumpy geometry (Ramos Almeida et al., 2011; Ichikawa et al., 2015). However for the general case the precise structure and geometry of the torus remains unknown, largely due to the size scales in question, rendering direct imaging impossible with current technology for the vast majority of sources.

Where interferomtery is not possible, the inner dusty structures of AGN can be probed with reverberation mapping techniques, which are predominantly used in mapping the spatial extent of the broad line region in AGN, but can also be used to measure the physical size of other components such as the accretion disc or hot dust (Clavel et al., 1989; Glass, 1992; Sitko et al., 1993; Nelson, 1996; Oknyanskij & Horne, 2001; Glass, 2004). This method assumes that the variable near-IR continuum arises from hot dust reprocessing of the variable optical/UV emission from the accretion disc and therefore the inner edge of the torus can be measured as the time lag between these two variable signals. This method is well tested, with 20 AGNs having been reverberation mapped in the near-IR through long term coordinated optical and near-IR monitoring campaigns (Clavel et al., 1989; Sitko et al., 1993; Nelson, 1996; Oknyanskij & Horne, 2001; Glass, 2004; Minezaki et al., 2004; Suganuma et al., 2006; Koshida et al., 2014; Vazquez et al., 2015). Dust radii measured in this way are systematically smaller than radii found with near-IR interferometry (Oknyanskij & Horne, 2001; Kishimoto et al., 2007; Nenkova et al., 2008). The same discrepancy is seen between reverberated radii and those determined by the dust sublimation temperature and AGN bolometric luminosity. Previous studies have attributed this offset to the dust geometry, advocating for a bowl-shaped rather than a spherical (doughnut-like) structure, caused by the anisotropy of the

accretion disc emission (Kawaguchi & Mori, 2010, 2011). This shape would also be seen if the dust formed part of a radiatively accelerated outflow from the disc, or a disc wind (Hönig & Kishimoto, 2017).

Reverberation mapping has most commonly been carried out with photometric data, largely due to the relative ease in obtaining high cadence lightcurves for multiple objects simultaneously. However photometric reverberation mapping has limitations that can be overcome by using spectroscopy instead. The lightcurves are not true reflections of the continuum flux variability as, due to fairly wide bandwidth filters, they are often contaminated by broad line emission, which superimpose their own intrinsic variability originating from the BLR, and show reverberation, but from different radii than the hot dust emission (Peterson et al., 2004; Koshida et al., 2014). With time-resolved spectroscopic data these line profiles can be reliably separated from the continuum emission, providing not only an uncontaminated continuum lightcurve, but also allowing for the assembly of simultaneous emission-line lightcurves. This allows us to investigate the connection between variations of emission-line flux and profile shape simultaneously with dust continuum variations. In the scheme that the BLR is limited by the dusty torus, we could even gain insights into the shape of the BLR, current evidence would suggest that the BLR, like the torus could be bowl shaped (Goad et al., 2012). Most BLR reverberation mapping is carried out on the optical lines, however, and so is most probably sampling a smaller size scale than the most extended part of the BLR. In this work we extend emission-line reverberation studies to infrared lines, namely Pa ϵ and Pa β. In addition to line measurements, spectroscopy also makes the simultaneous measurement of dust temperature possible, making astro-chemical analysis possible through sublimation arguments. This technique is now viable due to the advent of short-cross dispersed near-IR spectrographs mounted on medium sized telescopes, and has been utilised in Landt et al. (2019) which focused on NGC 5548. A similar version of this can be achieved with simultaneous multi-band photometry, however this method is less precise, presenting another advantage of opting for spectroscopy over photometry.

5.1.1 This Chapter

In this work we present our findings from a 9-month long spectroscopic near-IR reverberation campaign on the nearby, well studied AGN Mrk 509 conducted between 2019 February to November. In this Chapter, we focus mainly on the near-IR continuum emission, but also investigate the variability of Pa β and Pa ϵ.

Due to various observational constraints we are not able to report confident results, however we do present a detailed discussion on the limitations of performing this technique and discuss ways in which the difficulties inherent to such ground based near-IR spectroscopic campaigns can be overcome. We ultimately demonstrate the feasibility of measuring both host dust and Paschen lags with a better sampled data set and develop a re-scaling procedure that can be used in cases where there is no viable scaling feature present in the spectra.

We adopt here cosmological parameters $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_{\rm M} = 0.3$, and $\Omega_\Lambda = 0.7$, which give a luminosity distance to Mrk 509 of 150 Mpc and an angular scale at the source of 682 pc per arcsec.

5.2 The science target

This study focuses on the Type-1 AGN Mrk 509. We selected this target for several reasons: (i) it has a previously measured, short and variable dust response time of (126 ± 11) days (Koshida et al., 2014), therefore within our 6 month baseline of observations we should be able to recover a hot dust lag measurement; (ii) it has clearly discernible broad emission lines in the Paschen series allowing us to separate narrow- and broad-line components; (iii) multiple reverberation campaigns have determined its black hole mass. Peterson et al. (2004) reported the black hole mass to be $(1.43\pm0.12)\times10^8$ M$_\odot$. In addition Mrk 509 has a spatially resolved torus radius of (296 ± 31) lt-days from near-IR interferometry (Gravity Collaboration et al., 2020).

Mrk 509 (J2000 sky coordinates R.A. $20^h44^m09.8^s$ and Decl. $-10°43'24.7''$) can be observed from both the northern and southern hemispheres. Its relatively low redshift of $z = 0.0344$, enables us to sample the hot dust component and several near-IR coronal lines using a cross-dispersed near-IR spectrum. This AGN is relatively bright in the near-IR (2MASS $J = 11.6$ mag, $K = 10.0$ mag; Skrutskie et al. (2006)) ensuring good quality near-IR spectra are possible using reasonable exposure times on a 4 m class telescope and inhabits a compact host galaxy, perhaps of type S0 (Kriss et al., 2011).

5.3 The observations

Our aim is to measure a reverberation signal in the near-IR hot dust emission and the Paschen emission lines of Mrk 509, and to measure the dust temperature. We therefore monitor Mrk

509 using near-IR spectroscopy and also require well sampled optical photometry to trace the accretion disc variations.

5.3.1 The near-IR spectroscopy

Using the Spex spectrograph at the 3 m NASA Infrared Telescope Facitity (IRTF) on Manaukea, Hawaii, we observed the source Mrk 509 between 2019 May and November (semesters 2019A and 2019B) (Rayner et al., 2003). Having requested 36 nights of observing to obtain a cadence of ∼1 week, we were scheduled 24, 2-2.5 hour, observing windows between 2019 April and November. We obtained a total of 17 near-IR spectra as we lost seven of the scheduled observations due to weather and site inaccessibility. The journal of observations is listed in Table 5.1.

All observations were carried out using the (0.7-2.55) μm short cross-dispersed mode (SXD) equipped with the 0.3"x15" slit, oriented at the parallactic angle. Using this narrow slit allowed us to minimise the contamination from the host galaxy, whilst giving a sufficiently high spectral resolution, of R=2000 or full width at half maximum (FWHM) $\sim$ 150 km s^{-1}, to discern clearly any narrow and broad emission lines and study their profiles. We observed the source using a standard ABBA nodding pattern. As the source is not extended we nodded along the slit rather than on sky, therefore obtaining an image of the target in both the A and B frames from which the background subtraction could be carried out. The usual on-source exposure time was 32×120s. On seven of the seventeen observing runs we obtained fewer exposures than this due to either weather or technical issues. We also note that on 2019 September 10 we achieved 38 exposures due to spare time. This exposure time was chosen in order to guarantee high enough (S/N) to reliably measure not only the broad, but also the narrow emission line profiles, which can be key in performing photometric corrections to the spectra, as discussed in Section 5.3.3.

We observed a standard star to correct for telluric absorption and for flux calibration. Therefore after the science target we observed a nearby (in airmass and position) A0 V star, HD 203893 which has accurate optical magnitudes ($B = 6.89$ and $V = 6.84$). Where possible we took flats and arcs after the science target. Photometric corrections to the spectra in order to investigate intrinsic variability is carried out through a separate process detailed in Section 5.3.3.

The data were reduced using SPEXTOOL (version 4.1) a software package created for use by Spex users and based in Interactive Data Language (IDL) (Cushing et al., 2004). This

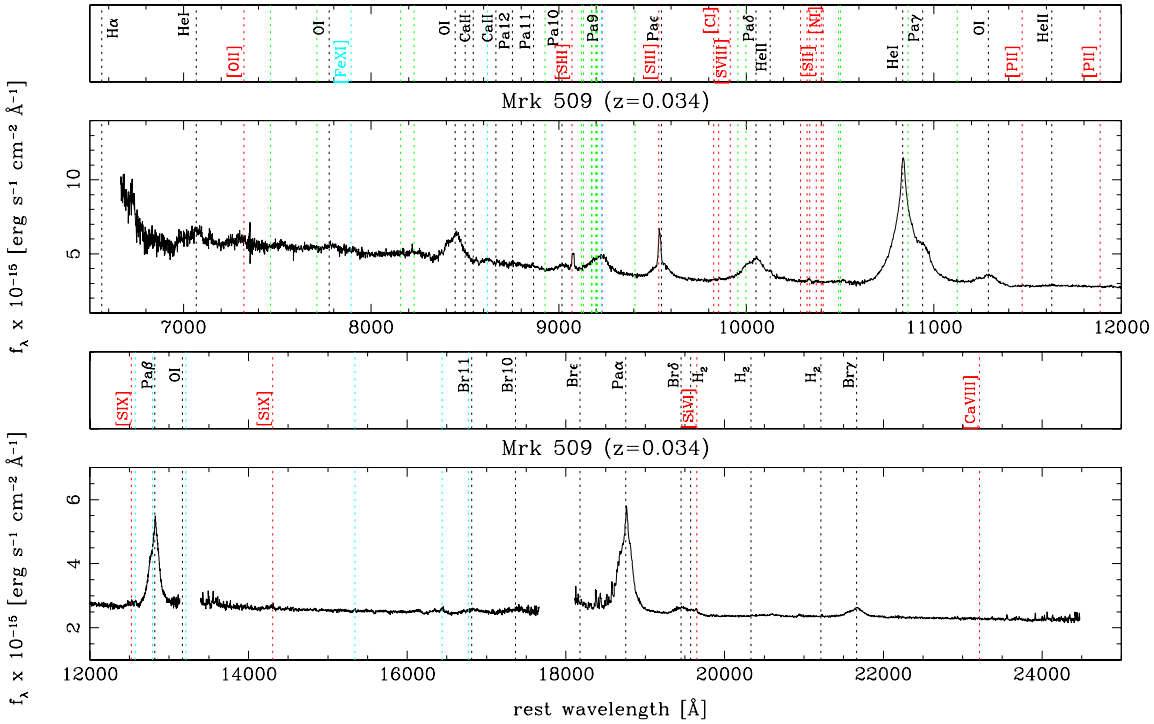

Figure 5.1: IRTF SpeX near-IR spectrum from 2019 May 9 shown as observed flux versus rest-frame wavelength. Emission lines listed in Table 4 of Landt et al. (2008) are marked by dotted lines and labeled; black: permitted transitions, green: permitted Fe II multiplets (not labeled), red: forbidden transitions and cyan: forbidden transitions of iron (those of [Fe II] not labeled).

package was used on both the science target and telluric standard. Spextool carries out all of the procedures required in the production of fully reduced spectra. This includes the preparation of calibration frames, the subtraction of flat fields, the fitting and subtraction of local background, interactive spectral extraction from science frames using an optimally weighted extraction scheme (Horne, 1986), telluric correction of the science spectra, merging of orders, and combination of the different frames into a single reduced science spectrum for each observing window. The final spectrum was corrected for Galactic extinction using a value of $A_V = 0.079$, which we derived from the Galactic hydrogen column densities published by Dickey & Lockman (1990). In Figure 5.1, we show the spectrum from 2019 May 9 as a representative example.

Table 5.1: IRTF Journal of observations

Observation Date	MJD	exposure (s)	airmass	aperture (arcsec^2)	PA (°)	continuum S/N			Standard Star airmass	seeing (arcsec)	cloud condition
						J	H	K			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
2019 May 9	58613	32×120	1.39	0.3×1.3	293	40	92	103	1.32	0.8	clear
2019 May 20	58624	32×120	1.24	0.3×1.4	106	48	82	117	1.23	0.8	cirrus
2019 May 30	58634	32×120	1.18	0.3×1.3	100	52	93	117	1.21	0.5	photom.
2019 Jun 8	58643	32×120	1.16	0.3×1.3	94	45	81	111	1.21	0.5	photom.
2019 Jun 14	58649	32×120	1.16	0.3×1.2	266	40	73	109	1.24	0.5	photom.
2019 Jun 18	58653	32×120	1.17	0.3×1.2	83	49	85	114	1.26	0.7	clear
2019 Jul 15	58679	32×120	1.26	0.3×1.2	108	46	85	103	1.24	0.4	photom.
2019 Aug 26	58721	24×120	1.39	0.3×1.1	113	52	84	106	1.38	0.5	cirrus
2019 Sep 2	58728	16×120	1.22	0.3×1.5	104	35	75	93	1.27	0.6	cirrus
2019 Sep 10	58736	38×120	1.17	0.3×1.2	97	56	84	110	1.21	0.7	cirrus
2019 Sep 19	58745	24×120	1.16	0.3×1.1	100	54	75	95	1.21	0.5	cirrus.
2019 Sep 26	58752	24×120	1.31	0.3×1.4	230	56	77	115	1.31	0.7	cirrus
2019 Oct 1	58757	26×120	1.30	0.3×1.3	109	59	75	103	1.27	0.6	clear
2019 Oct 21	58777	32×120	1.17	0.3×1.6	265	51	59	97	1.25	1.2	cloudy
2019 Oct 31	58787	30×120	1.19	0.3×1.4	79	44	76	119	1.32	0.8	clear
2019 Nov 4	58791	16×120	1.21	0.3×1.2	76	40	67	109	1.27	0.7	clear
2019 Nov 14	58801	24×120	1.23	0.3×1.4	255	37	61	95	1.12	0.5	clear

The columns are: (1) Universal Time (UT) date of observation; (2) Modified Julian Date (MJD); (3) exposure time; (4) mean airmass; (5) extraction aperture; (6) slit position angle, where PA= 0° corresponds to east-west orientation and is defined east through north; S/N in the continuum over $\sim$ 100Å measured at the central wavelength of the (7) J, (8) H, and (9) K-bands; for the telluric standard star; (10) mean airmass; (11) seeing in the K band; (12) cloud condition. All observations were taken contemporaneously with the telluric standard star HD 203893.

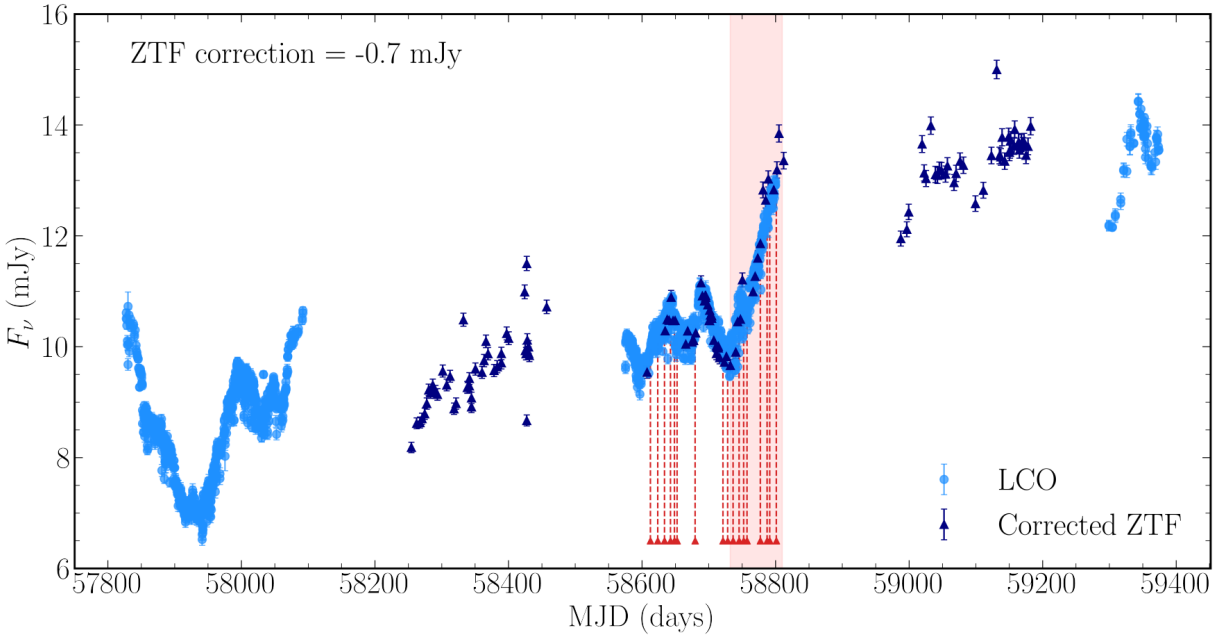

Figure 5.2: LCO g-band observations of Mrk 509 with complementary corrected ZTF g band data. The most significant variability feature measured contemporaneously with our near-IR spectra is shaded in red. Archival LCOGT data (MJD<58200) is shown to extend the time baseline. Spectral epochs are marked by red arrows.

5.3.2 Complementary Optical photometry

The LCOGT observations

Between 2019 April 2 and November 12 we monitored Mrk 509 in the optical with the 1 m and 2 m robotic telescopes of the Las Cumbres Observatory (LCOGT; Brown et al., 2013) almost daily in four bands (g, r, i and z_s). This baseline of observations begins ∼1 month before our near-IR spectra and ends 2 days before the end of our near-IR campaign. We use here mainly the lightcurves in the g and z_s filters, which have a central wavelength of 4770 Å and 8700 Å and a width of 1500 Å and 1040 Å respectively. The Sinistro cameras on the 1 m telescopes have a field-of-view (FOV) of $26.5' \times 26.5'$ and a plate scale of $0.389''$ per pixel. The Spectral cameras on the 2 m telescopes have a FOV of $10.5' \times 10.5'$ and a plate scale of $0.304''$ per pixel.

The frames were first processed by LCOGT's BANZAI pipeline (McCully et al., 2018) in the usual way (bias and dark subtraction, flat-fielding and image correction) and were subsequently analysed with our custom-made pipeline, which has been described in detail by Hernández Santisteban et al. (2020). In short, after performing aperture photometry with a diameter of $7''$ and subtracting a background model, we produced stable light curves by constructing a curve of growth using the standard stars on each individual frame and measuring

Table 5.2: Observed optical nuclear fluxes from GROND photometry

Observation Date	g' $\lambda_{\rm eff} = 4587$ Å (mJy)	r' $\lambda_{\rm eff} = 6220$ Å (mJy)	i' $\lambda_{\rm eff} = 7641$ Å (mJy)
2019 Aug 26	9.7	13.3	10.9

Table 5.3: Observed near-IR nuclear fluxes from GROND photometry

Observation Date	J $\lambda_{\rm eff} = 12399$ Å (mJy)	H $\lambda_{\rm eff} = 16468$ Å (mJy)	K $\lambda_{\rm eff} = 21706$ Å (mJy)
2019 Aug 26	17.4	22.7	40.3

the correction factors required to bring all different light curves to a common flux level. This approach mimicked extraction from an azimuthally averaged point-spread function (PSF), but without actually performing a PSF extraction. A colour correction and correction for atmospheric extinction were applied before the photometric calibration. Finally, we used comparison stars in each field to perform an image zero-point calibration at each epoch. In Figure 5.2, we show the g and in Figure 5.5 the z_s light-curve, with the latter overlapping in wavelength with the near-IR spectrum.

The ZTF data

The LCO g-band lightcurve has been supplemented with data taken from the Zwicky Transient Facility (ZTF) (Bellm, 2014). These data were corrected to match the flux scale of the LCO observations by interpolating the LCO flux lightcurve to the ZTF observation times, calculating a mean flux difference, and then subtracting this from the ZTF lightcurves. This correction value was $\sim$0.7 mJy. The resulting ZTF lightcurve is displayed in Figure 5.2 alongside the LCO observations. The ZTF data fills large gaps present in the LCO data-set, and therefore provides a helpful constraint on the behaviour of Mrk 509 during these time periods.

The GROND observations

We obtained simultaneous optical and near-IR photometry in seven bands with the Gamma-Ray Burst Optical and Near-Infrared Detector (GROND; Greiner et al., 2008) mounted on the MPG 2.2 m telescope, La Silla, Chile. The observations were taken on 2019 August 26, simultaneous with one of our near-IR spectra. The optical channels have a FOV of $5.4' \times 5.4'$ and a plate scale of $0.158''$ per pixel, whereas both the FOV and plate scale of the

near-IR channels is larger ($10' \times 10'$ and $0.6''$ per pixel, respectively). The filter passbands are relatively wide ($\sim 1000 - 3000$ Å). Images were reduced with the IRAF-based pipelines devoloped by R. Decarli and G. De Rosa (Morganson et al., 2012), modified and extended for our purpose. After bias and dark subtraction and flat-fielding of each individual image, a scaled median sky image constructed from separate sky exposures was subtracted from each frame. The dither step size was $18''$ for both the science and sky images, with five dither positions obtained for each.

In Table 5.3, we list the optical and near-IR nuclear fluxes. They were obtained with GALFIT (Peng et al., 2002), a software package that models the object's surface brightness profile with a point spread function (PSF), for the unresolved AGN, and a host galaxy component. We modelled the host galaxy with only a bulge due to the relatively low S/N of the images. In addition, GALFIT simultaneously models the PSF fluxes of the reference stars, which we used to flux-calibrate the optical and near-IR images based on the known stellar fluxes as listed in the SDSS (Abazajian et al., 2009) and 2MASS (Skrutskie et al., 2006). We note that we had to omit the z-band images due to lack of SDSS reference stars in this area and that the host galaxy parameters in the H and K bands had to be constrained (based on the other images) rather than left to vary.

5.3.3 Photometric corrections to the spectra

For the purposes of this work it is essential to distinguish between intrinsic AGN variability and extrinsic sources of variability such as changing night to night observing conditions. Therefore to obtain meaningful lightcurves from the spectra it was necessary to perform a photometric correction to the spectra. Failing to do this could result in a contamination of the final dust response lightcurves Bentz et al. (2016).

Several software packages employ different approaches to perform the photometric re-scaling of spectra, calibrating out time-dependent instrumental effects, e.g PREPSPEC (written by Keith Horne and discussed in Shen et al. (2016)) and GW92 (Van Groningen & Wanders, 1992). This work uses the open source PYTHON package MAPSPEC (Fausnaugh, 2017) which is an updated and improved version of the GW92 package. MAPSPEC , like GW92, aligns an assumed non variable line profile of a given observed spectrum to a reference spectrum that has been constructed, by applying a smoothing kernel, a flux re-scaling factor, and a wavelength shift.

In order to correct for time-dependent instrumental effects, a spectral feature which is intrinsically non variable over the baseline of observations is required. Traditionally, for optical

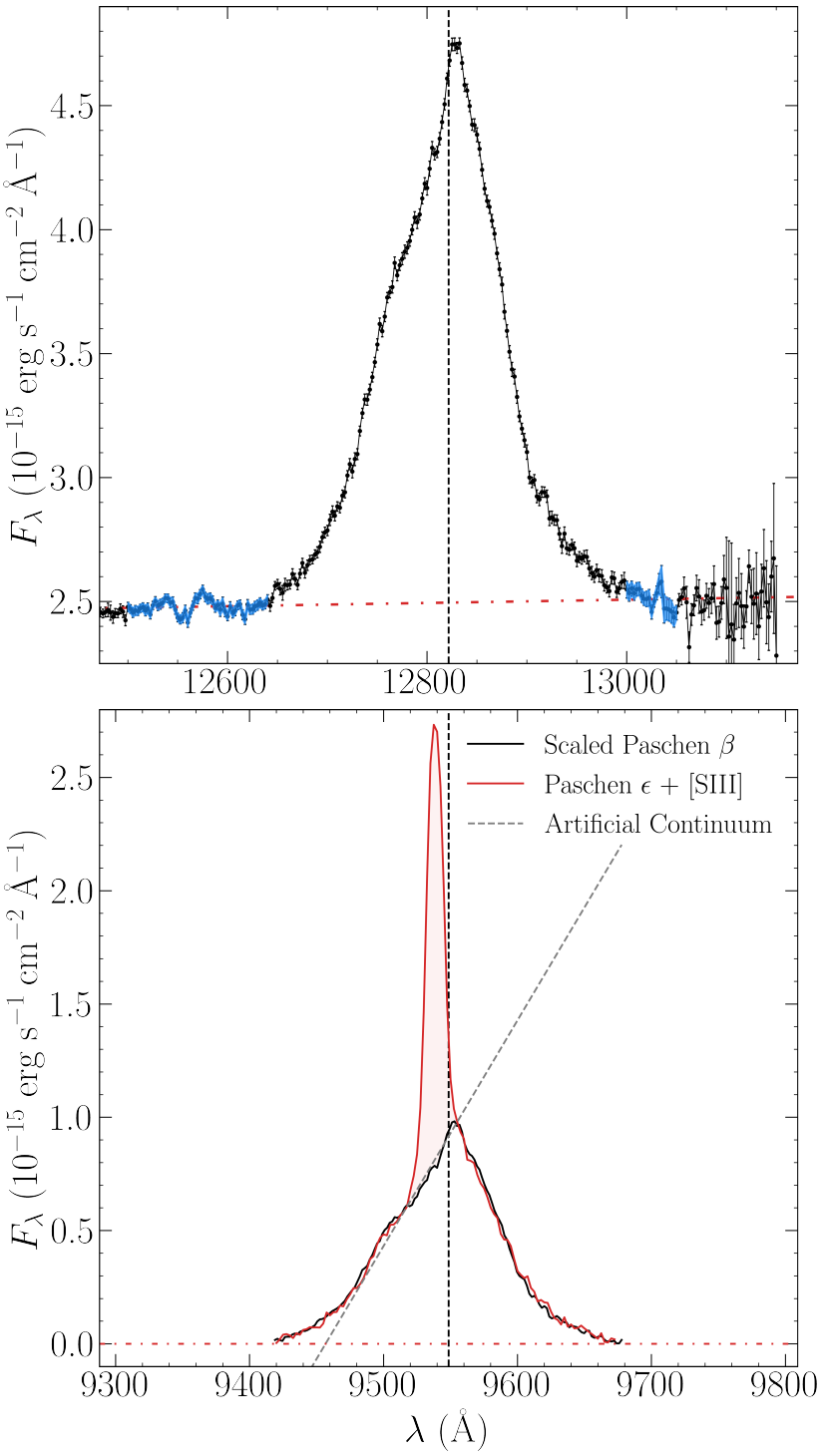

Figure 5.3: **Top Panel:** Pa β emission line from Mrk 509 spectrum taken on 26 August 2019. Red line represents a linearly fitted continuum, subtracted in order to calculate enclosed flux. The continuum has been fit using a blueward clean continuum window of (12500-12640) Å and the redward clean continuum window of (13000-13050) Å, these regions are highlighted in blue. The redward window is necessarily short as beyond this there is significant telluric absorption. **Bottom Panel:** Mean spectrum of Mrk 509 from all 4 nights with photometric conditions. Pa β profile (black) scaled to match Pa ϵ (red), in order to isolate the constant [S III] λ9531 component from any variable broad line emission. The slope of the Pa β profile in the [S III] window is shown as the artificial continuum in grey. Both panels centre on the rest frame wavelengths of 12821.6 Å and 9548.6 Å for Pa β and Pa ϵ respectively and have a matched velocity range.

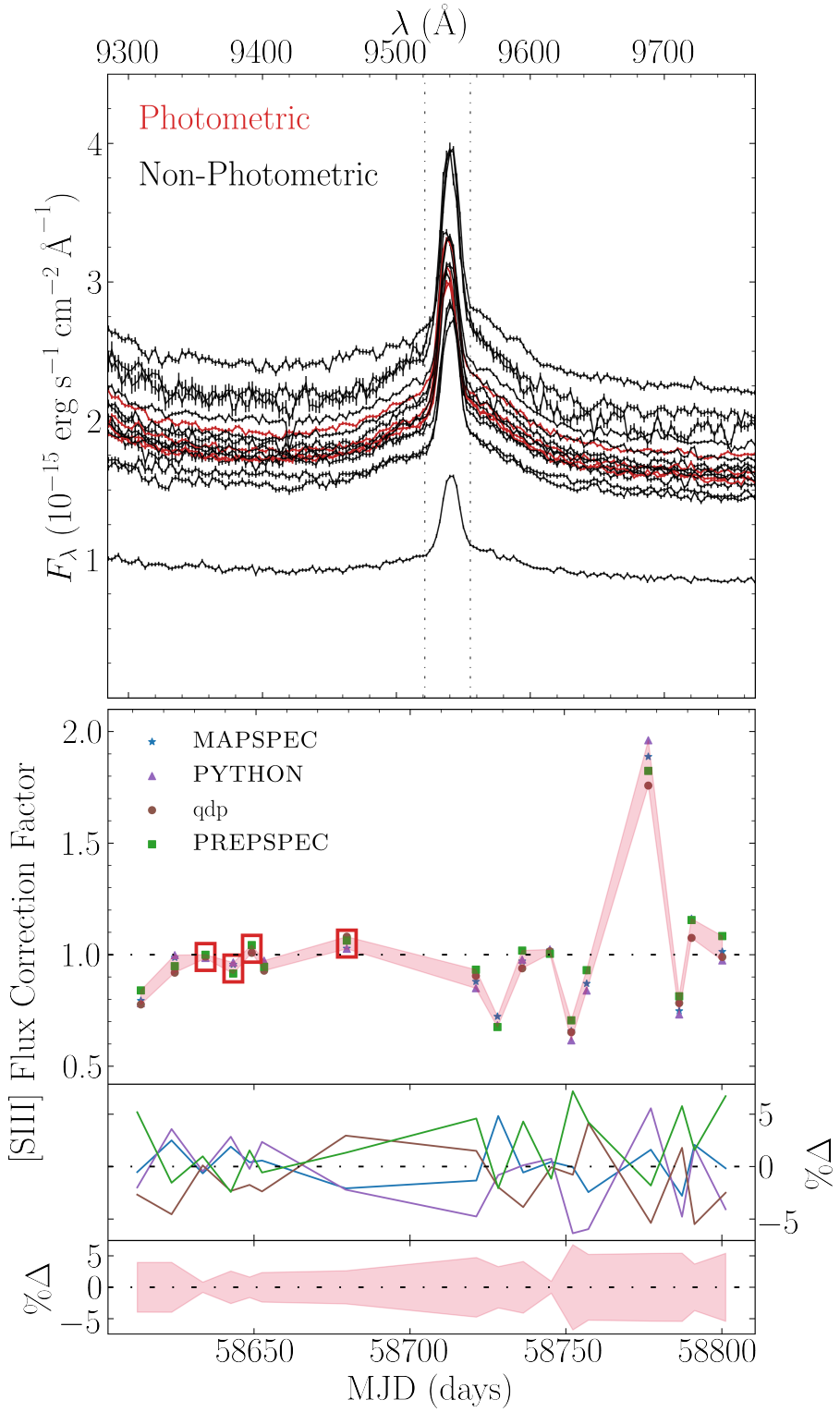

Figure 5.4: **Top Panel:** All 17 Mrk 509 spectra, showing the region around Pa ϵ and [S III]. The four photometric spectra are coloured red, the others black. **Second Panel:** [S III] flux correction factors calculated using four different methods for each night of observation, photometric nights are enclosed by red boxes. The four methods used are as follows; MAPSPEC using spectra with an appended artificial continuum (see Figure 5.3), interactive fitting routine qdp with a linear, narrow Gaussian and broad Gaussian component, numerically integrating under the [S III] line profile after the removal of the artificial continuum and PREPSPEC. **Third Panel:** Percentage of the mean correction factor for each value obtained via the four methods described above, coloured similarly to the panel above. **Bottom Panel:** Percentage range of the four methods around their mean.

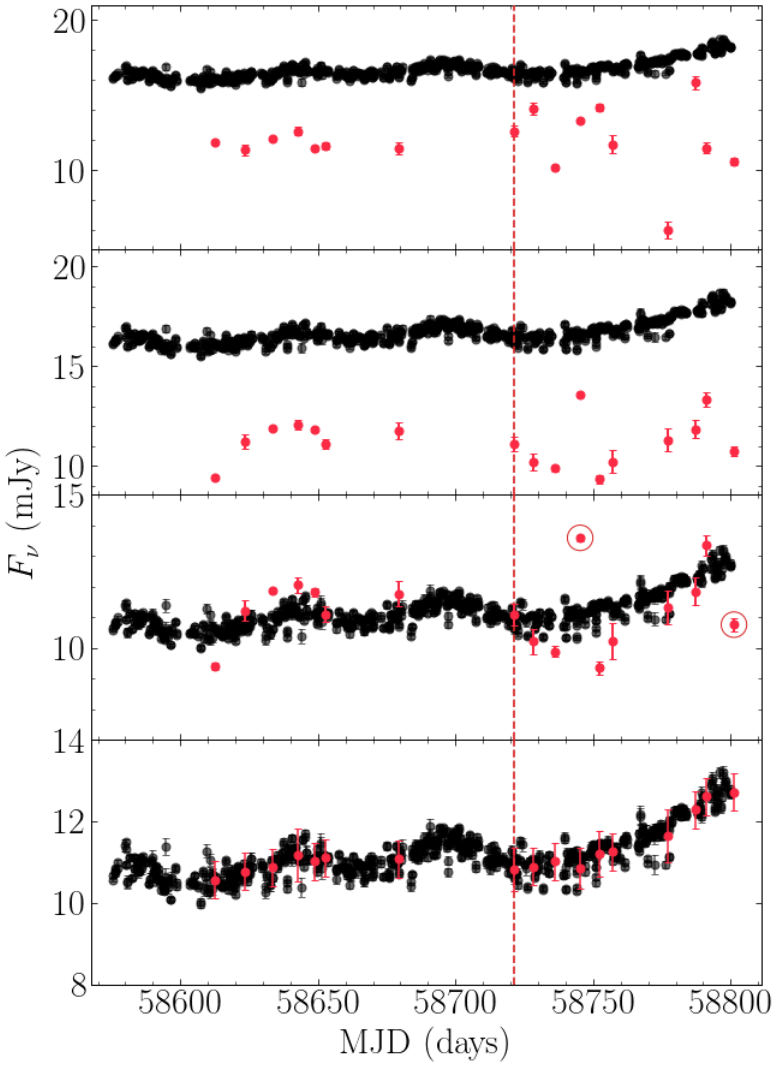

Figure 5.5: **Top Panel:** LCO z_s band photometry, uncorrected for host galaxy contribution in black, with all 17 spectral LCO z_s band (8180-9220 Å observed frame) uncorrected fluxes plotted in red. **Second Panel:** LCO z_s band photometry, uncorrected for host galaxy contribution in black, with all 17 spectral LCO z_s band ((8180-9220) Å observed frame) fluxes, corrected with MAPSPEC scaling factors are plotted in red.**Third Panel:** LCO z_s band photometry, corrected for a constant host galaxy contribution in black, with all 17 spectral LCO z_s band (8180-9220 Å observed frame) fluxes, corrected with the MAPSPEC scaling factors plotted in red. Spectral fluxes measured on MJD 58745.30 and 58801.20 are encircled in red as they do not match well with the LCO z_s band photometry. The night of 26$^{\text{th}}$ August 2019 is highlighted with a dashed red line. **Bottom Panel:** The MAPSPEC corrected lightcurve has been further corrected to the LCO z_s band photometry by scaling to the weighted mean of all photometric data points within three days of each spectral data point. The errors displayed are defined by the spread of the data to which the spectral points have been corrected, or the mean spread of the photometry within any given 3 day period of the spread if the correction was less than this value. These uncertainties have been propagated to our corrected lightcurves along with the MAPSPEC uncertainties displayed in the panel above.

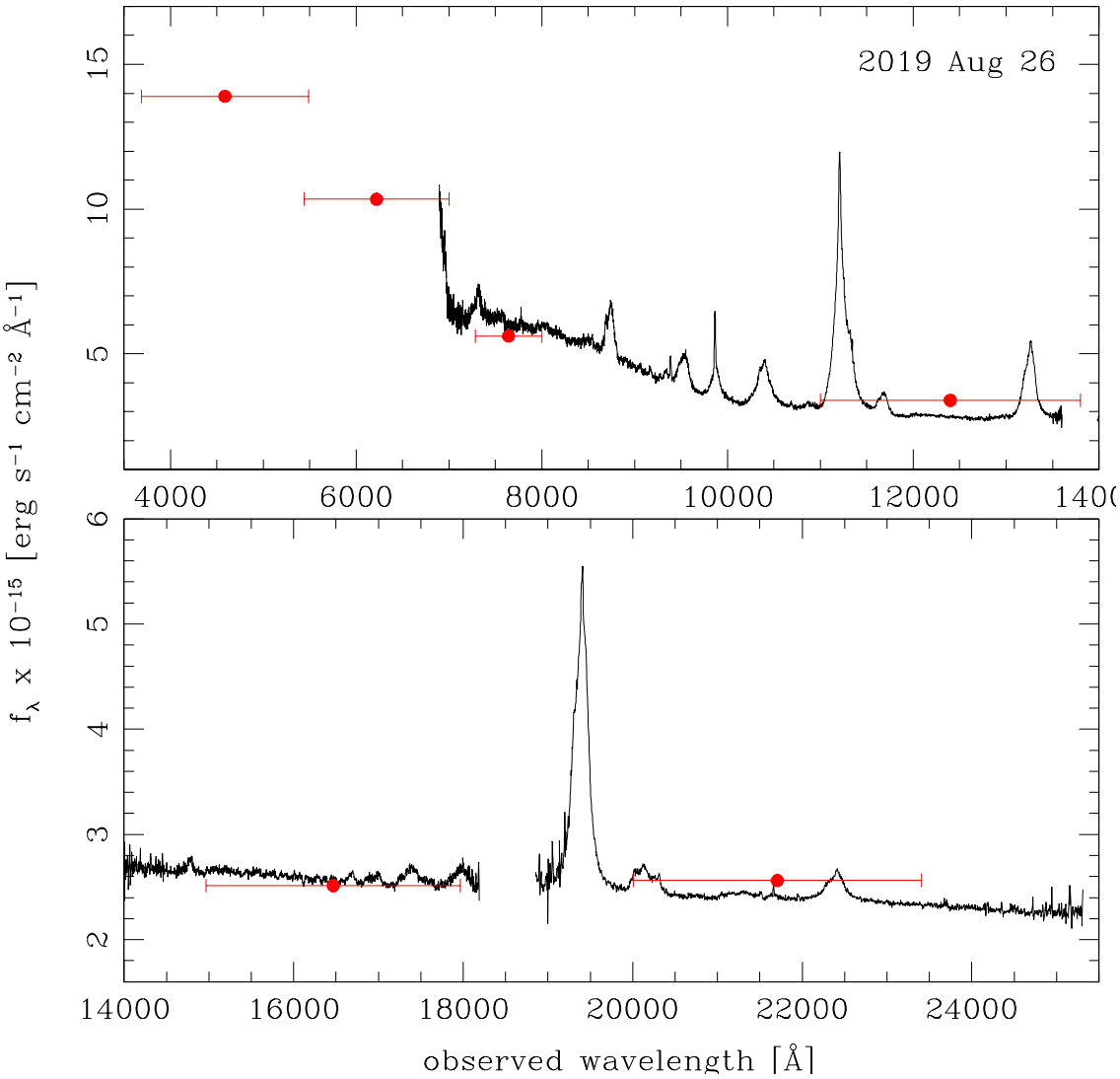

Figure 5.6: IRTF SpeX near-IR spectrum from 26$^{\text{th}}$ August 2019, highlighted in Figure 5.5 (black), overlaid with the optical and near-IR nuclear fluxes from the GROND photometry obtained on the same day (red filled circles).

campaigns, this is the [O III] λ5007 narrow emission line. Confirmed via numerous IFU observations [O III] λ5007 is separated from the ionising source by a pc-to-kpc scale region and therefore constant on timescales much larger than the baseline of observations typically used in reverberation studies such as this (Fausnaugh, 2017). However, this emission line is not visible in the near-IR, therefore we have selected the [S III] λ9531 narrow emission line, which is analogous to the [O III] λ5007 region due to its extent.

A complication in the use of the [S III] λ9531 narrow emission line is that it is blended with the variable broad Pa ϵ. It was therefore important to separate these two profiles to allow for accurate flux scaling. In order to do this a mean spectrum of our 4 photometric nights was constructed. Using this mean spectrum we isolated and fit the broad Pa β to the broad Pa ϵ profile in velocity space and then converted back to wavelength space. This fit is displayed in Figure 5.3.

A linear fit to the Pa β profile within the [S III] window of (9520 - 9560) Å was performed to calculate the slope of an artificial continuum which could be used to isolate the intrinsically non-variable [S III] narrow line from the variable broad Pa ϵ upon which it sits. Figure 5.4 top panel shows the Pa ϵ and [S III] emission region from all 17 Mrk 509 spectra with the real continuum and the broad Pa ϵ profile, with the bottom panel displaying the [S III] scale factors calculated in four different ways.

Using MAPSPEC, a reference spectrum was constructed from 4 photometric spectra using the artificial continua as displayed in Figure 5.3 as the continuum level, thereby isolating the [S III] profile. This reference spectrum was used alongside all 17 spectra with artificial continua in order to calculate re-scaling factors. MAPSPEC allows for three different smoothing methods, a Delta function, a Gaussian and Gauss-Hermite smoothing kernel. We have taken the final scaling values as the Gauss-Hermite smoothing value as this is the most stable of the three methods (Fausnaugh, 2017). We performed a test using two different continua, the broad line profile and the artificial continua shown in Figure 5.3, resulting in an average difference in scale factor of 3.5% between the two continuum options for the Gauss-Hermite method compared to an average difference in scale factor of 10.4% for the delta function, therefore confirming this as the most robust smoothing method.

To quantify the uncertainties from the scaling procedure we have calculated scaling factors using three alternative methods. Firstly by calculating the integrated flux in the [S III] line using the artificial continua and scaling these areas to the mean area of all four photometric nights. Secondly, using the interactive fitting package qdp, we modelled the [S III] region with

a linear continuum plus broad and narrow Gaussian components. The linear continuum and broad line components were subtracted and the flux enclosed by the narrow line was calculated and scaled to the mean of the photometric nights. Finally we used PREPSPEC to calculate a set of flux correction factors. In Figure 5.4, data points representing the photometric nights (enclosed in red) show very little spread in correction factor and all lie within a few percent of unity, as expected. Across all observations, the corrections factors for each method of calculation described above agree within $\lesssim 5\%$. Figure 5.4 shows that on average MAPSPEC is closest to the mean correction factor for each observation, but the distribution around the mean for the other three correction factors appears random, indicating that one method alone does not significantly increase the spread displayed in the bottom panel.

The two central panels of Figure 5.5 show the MAPSPEC corrected LCO z_s-band spectral fluxes compared with LCO z_s-band photometry. Whilst an exact agreement with the photometry would not be necessary to confirm the validity of the correction factors, it is clear that there are two measurements which are very discrepant with respect to the contemporaneous photometric measurements. These large discrepancies reduce the confidence in the scaling procedure and the assumptions upon which it is based, perhaps indicating that there is some extension in the [S III] region causing variability in this line between observations not due to atmospheric effects. We therefore scale to the LCO z_s-band photometry, as shown in the bottom panel of Figure 5.5. To do this, we bin the photometric data into 3-day windows around each spectral data point and scale to the weighted mean of this value, we then propagate the weighted standard error of this bin forward as the scaling uncertainty, and in any bin where this is less than the average spread of the data in any 3-day period we adopt this value as the uncertainty on the correction.

5.3.4 Variability in the broad Paschen emission lines

As Mrk 509 is a Type 1 Seyfert it exhibits both broad and narrow emission lines in its optical and near-IR spectra. The Paschen series are a group of hydrogen recombination transitions that fall within the spectral range of the IRTF Spex instrument, and therefore are detected in our spectra. Figure 5.1 shows these broad emission lines, the most prominent of which are Pa ϵ, Pa β and Paα. It is clear from their profile shapes that these emission lines are emitted from the BLR, which is observed to vary due to its illumination by the accretion disc (Peterson et al., 2004). Therefore it should be possible to measure variability in the Paschen lines from our spectra, and to search for a reverberation signal from our 'driving' optical LCO

lightcurve and thereby place constraints on the location of the Paschen BLR.

To do this we have selected the Pa ϵ and Pa β lines, opting to reject Paα due to the effect of telluric absorption on its blueward wing, which can be seen in Figure 5.1, but is far more prominent on nights of poorer quality. Pa ϵ is surrounded by a relatively clean continuum, however there is a slight complication as on top of its broad profile sits [S III], procedures for the removal of Pa ϵ from the [S III] are discussed in Section 5.3.3. This technique can be used to isolate the Pa ϵ flux, and is discussed in more detail in Section 5.4.2.

5.3.5 The spectral continuum components

We aim to compare response- and luminosity-weighted dust radii. For the dust time delay we assembled the lightcurves of both the irradiating flux and hot dust. For the calculation of the luminosity-weighted dust radius we measured the dust temperature and estimated the total accretion disc luminosity that heats the dust. The large wavelength coverage of our cross-dispersed near-IR spectra gives in Mrk 509 roughly half of the hot dust SED, and the improved wavelength coverage of SpeX in the blue samples a considerable part of the accretion disc spectrum, which is expected to dominate the total continuum flux up to a rest-frame wavelength of ~ 1 μm (Landt et al., 2011a,b). We note that, since we used a relatively small spectral aperture of 0.3" which corresponds to a physical size of ∼205 pc at the source, the contribution from the host galaxy to the total observed continuum flux is expected to be small in this luminous AGN.

Following Landt et al. (2019), we decomposed the spectral continuum into two components. We first approximated the rest-frame wavelength range of $\lesssim 1$ μm with an accretion disc spectrum, which we subsequently subtracted from the total spectrum. For its calculation we adopted a black hole mass of $M_{\rm BH} = 1.09 \times 10^8$ solar masses, which is the value derived by optical reverberation campaigns (Peterson et al., 2004; Bentz & Katz, 2015) using a geometrical scaling factor of $f = 4.3$ to convert the measured virial product (Grier et al., 2013). Furthermore, we assumed that the disc extends to $r_{\rm out} = 10^4 r_{\rm g}$, where $r_{\rm g} = GM_{\rm BH}/c^2$ is the gravitational radius, with G the gravitational constant and c the speed of light, the value of $\dot{m}$ was resultant from each fit and an efficiency of $\eta = 0.057$ was assumed. We then fitted the resultant hot dust spectrum at wavelengths > 1 μm with a blackbody, representing emission by large dust grains, and with two blackbodies modified by a power-law of the form $Q_\lambda(a) \propto \lambda^\beta$, approximating with $\beta = -1$ and $\beta = -2$ the emissivity of sub-micron silicate and carbon dust grains, respectively (Landt et al., 2019, see their Figure 8). Despite our fitting of only

Table 5.4: Physical parameters for the calculation of luminosity-weighted dust radii

Observation Date	accretion disc	blackbody ($\beta = 0$)			silicate dust ($\beta = -1$)			carbon dust ($\beta = -2$)		
	log $L_{\rm uv}$ (erg/s)	$T_{\rm d}$ (K)	log $L_{\rm d}$ (erg/s)	$R_{\rm d,lum}$ (lt-days)	$T_{\rm d}$ (K)	log $L_{\rm d}$ (erg/s)	$R_{\rm d,lum}$ (lt-days)	$T_{\rm d}$ (K)	log $L_{\rm d}$ (erg/s)	$R_{\rm d,lum}$ (lt-days)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
2019 May 09	45.28	1369±11	44.66	168	1170±8	43.78	1591	1022±6	44.52	1022
2019 May 20	45.33	1386±14	44.64	174	1182±10	43.72	1652	1031±8	43.94	1063
2019 May 30	45.36	1362±12	44.65	187	1164±9	43.68	1763	1018±6	44.05	1129
2019 Jun 08	45.35	1362±8	44.67	184	1167±6	43.69	1734	1022±5	44.22	1107
2019 Jun 14	45.36	1352±10	44.65	189	1158±8	43.68	1781	1013±6	44.44	1140
2019 Jun 18	45.33	1363±12	44.66	180	1164±9	43.68	1703	1017±6	43.97	1093
2019 Jul 15	45.36	1357±13	44.66	188	1160±9	43.70	1775	1014±7	43.87	1138
2019 Aug 26	45.35	1353±14	44.69	187	1157±10	43.68	1764	1011±8	44.04	1132
2019 Sep 02	45.25	1381±15	44.61	160	1180±10	43.70	1511	1030±8	44.91	971
2019 Sep 10	45.32	1367±15	44.70	177	1167±10	43.72	1675	1018±9	44.04	1078
2019 Sep 19	45.41	1376±12	44.72	194	1173±8	43.73	1839	1022±7	43.98	1187
2019 Sep 26	45.33	1396±16	44.69	172	1190±11	43.75	1629	1038±8	44.07	1049
2019 Oct 01	45.35	1365±14	44.69	184	1166±10	43.70	1737	1018±7	44.04	1116
2019 Oct 21	45.38	1374±14	44.68	188	1171±10	43.76	1782	1021±8	44.33	1149
2019 Oct 31	45.47	1354±11	44.69	214	1156±9	43.70	2029	1010±6	43.92	1302
2019 Nov 04	45.47	1316±14	44.72	227	1127±10	43.70	2134	986±8	44.02	1366
2019 Nov 14	45.41	1364±14	44.63	197	1165±11	43.69	1864	1018±7	44.49	1196

The columns are: (1) Universal Time (UT) date of observation; (2) total accretion disc luminosity; for a blackbody emissivity (3) dust temperature; (4) total dust luminosity and (5) dust radius; for an emissivity law appropriate for silicate dust with small grain sizes of $a \lesssim 0.1$ μm (6) dust temperature; (7) total dust luminosity and (8) dust radius; for an emissivity law appropriate for carbon dust with small grain sizes of $a \lesssim 0.1$ μm (9) dust temperature; (10) total dust luminosity and (11) dust radius.

two components, there is evidence to suggest that at 1μm diffuse BLR continuum emission could account for 25% of the total continuum emission with this increasing to 35% at 8000Å (Korista & Goad, 2019). This could have an effect on our spectral decomposition, and on our spectral lightcurves, particularly those extracted at wavelengths <1μm, see Section 5.4.1. Table 5.4 lists the relevant physical parameters extracted from the spectral decomposition. We obtain average temperatures of $\langle T \rangle = 1365 \pm 4$ K, 1166±3 K and 1018±3 K for emissivity laws with $\beta = 0$, -1 and -2, respectively.

As in Landt et al. (2019), we calculated luminosity-weighted dust radii, $R_{\rm d,lum}$, from the best-fit dust temperatures assuming radiative equilibrium between the luminosity of the irradiating source and the dust:

$$\frac{L_{\rm uv}}{4\pi R_{\rm d,lum}^2} = 4\sigma T^4 \langle Q^{\rm em} \rangle, \tag{5.3.1}$$

where σ is the Stefan-Boltzmann constant and $\langle Q^{\rm em} \rangle$ is the Planck-averaged value of $Q_\lambda(a)$. We approximated $L_{\rm uv}$ with the accretion disc luminosity and have used for the Planck-averaged emission efficiencies in the case of $\beta = -1$ a value of $\langle Q^{\rm em} \rangle = 0.0210$ appropriate for silicates of $T = 1259$ K and $a = 0.1$ μm (Laor & Draine, 1993) and in the case of $\beta = -2$ a value of $\langle Q^{\rm em} \rangle = 0.0875$ appropriate for graphite of $T = 1000$ K and $a = 0.1$ μm (Draine, 2016). The average luminosity-weighted dust radii are $\langle R_{\rm d,lum} \rangle = 186 \pm 4$ light-days, 1763±36 light-days and 1132±23 light-days in the case of a blackbody, and small-grain silicate and carbon dust, respectively.

5.3.6 The mean and variable ($F_{\rm var}$) spectrum

The Fractional mean variance ($F_{\rm var}$) is a measure of fractional "excess variance", and therefore is a measure of intrinsic variability above that due to flux measurement errors. $F_{\rm var}$ is calculated as,

$$F_{\rm var} = \frac{(\sigma^2 - \delta^2)^{1/2}}{\langle F \rangle}, \tag{5.3.2}$$

following Rodríguez-Pascual et al. (1997), where $\langle F \rangle$ is the sample mean of the $N = 17$ flux data F_i;

$$\langle F \rangle = \frac{1}{N} \sum_{i=1}^{N} F_i, \tag{5.3.3}$$

and σ is the square root of the sample variance, calculated as,

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(F_i - \langle F \rangle)^2. \tag{5.3.4}$$

Finally, δ^2 is the mean square uncertainty of the fluxes, calculated as,

$$\delta^2 = \frac{1}{N}\sum_{i=1}^{N}\delta_i^2, \tag{5.3.5}$$

where δ_i is the uncertainty on each flux measurement F_i.

Figure 5.7 shows the mean variance (numerator of $F_{\rm var}$ in Equation 5.3.2) calculated for each spectral element for the 17 observations of our campaign. These values have been plotted against wavelength to create an intrinsic mean variance spectrum. Both the mean (black) and mean variance (red and blue) spectra were normalised to a clean continuum window just blueward of 1μm, and are shown with their 1σ uncertainties derived using Monte-Carlo methods. In blue are all of the the mean variance values that are more than a factor of 3 times their MC derived uncertainties from zero, and therefore defined as significantly variable to the 3σ level. It is clear that most of the variability is found at wavelengths $< 1\mu$m, and arises from the accretion disc. With increasing wavelength the accretion disc contribution decreases and therefore the variability decreases. Almost all data points (aside from those affected by a few telluric absorption features) at wavelengths >1 μm do not show significant variability to the 3σ level. The data gaps seen in the mean variance spectrum, again predominantly at wavelengths >1 μm are due to a lack of any real variability, characterised by the mean square uncertainty of the flux being larger than the variance of the flux itself.

The baseline and cadence of our data are insufficient to recover any 3σ significant variability in the hot dust emission in Mrk 509, and recovers relatively little 1σ variability, however we do detect 3σ variability in the accretion disc. The mean variance spectrum also reveals that there is little significant variability in the broad emission lines, indicating a response from the broad line region that our baseline and cadence make us unable to sample with any real confidence. This is evidenced by the dropping out of the mean variance spectrum over the broad emission line regions. This is further discussed in Sections 5.3.4 and 5.5.2.

The shape of the mean variance spectrum reveals that we see a very slight increase in the level of variability in the K-band when compared to the H-band, this is also reflected in the values listed in Table 5.6, and discussed in Section 5.4.1. The K-band variability shown

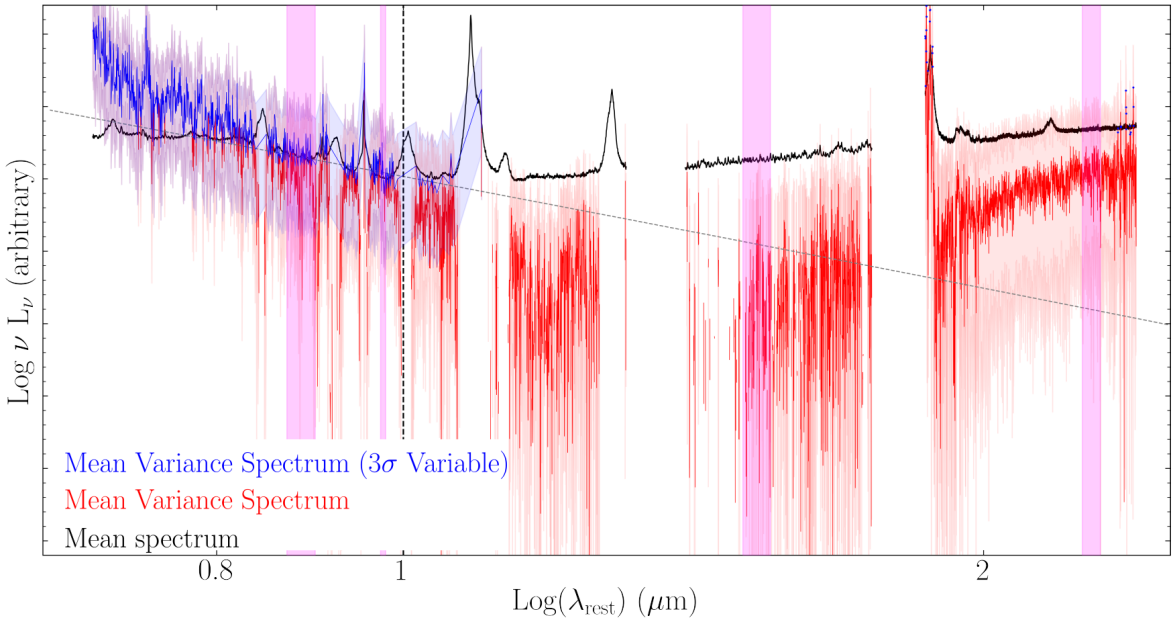

Figure 5.7: The mean (black) and mean variance spectrum (red) for our campaign normalised at rest-frame ~ 1 μm (vertical dashed line). The spectrum of a standard accretion disc (black dotted line) is observed out to ~ 1.2 μm in the variable component compared to out to ~ 1 μm in the mean spectrum. The 3σ variable mean variance values are shown in blue. Extraction windows for (0.870 - 0.900)μm, (0.973 - 0.979)μm, H - (1.50-1.55)μm and K - (2.25-2.30)μm are highlighted in fuchsia.

here is barely significant, therefore this increase in statistical variability does not necessarily reflect physical reality.

Figure 5.8 shows the $F_{\rm var}$ spectra (calculated as detailed in Equation 5.3.2) for the Pa ϵ and Pa β regions. For each line we show two plots, firstly we plot the δ and σ values, defined in Equations 5.3.5 and 5.3.4 respectively, as a function of wavelength. Secondly the $F_{\rm var}$ values are plotted along with their 1σ uncertainties. We compute and plot an upper limit for any spectral element for which the δ value exceeds the σ value. It is clear that the broad line profiles for both Pa ϵ and Pa β display a lower level of variability than the surrounding continuum regions. In the case of Pa ϵ the [S III] narrow line profile shows an $F_{\rm var}$ amplitude of $\sim 6\%$. This indicates that when scaled to the photometry, the [S III] profile is significantly variable. Whilst the narrow-line region has been found to vary over large timescales, it is unlikely that we are sensitive to this over the 183 days of our observations. The narrow-line region has been found to be extended over large physical size scales, meaning that we could see variation in the [S III] region due to our chosen slit width and observational effects such as seeing (Peterson et al., 2013). For reasons discussed in Section 5.3.3 we believe scaling to the photometry to be more reliable than assuming a non-variable [S III] component.

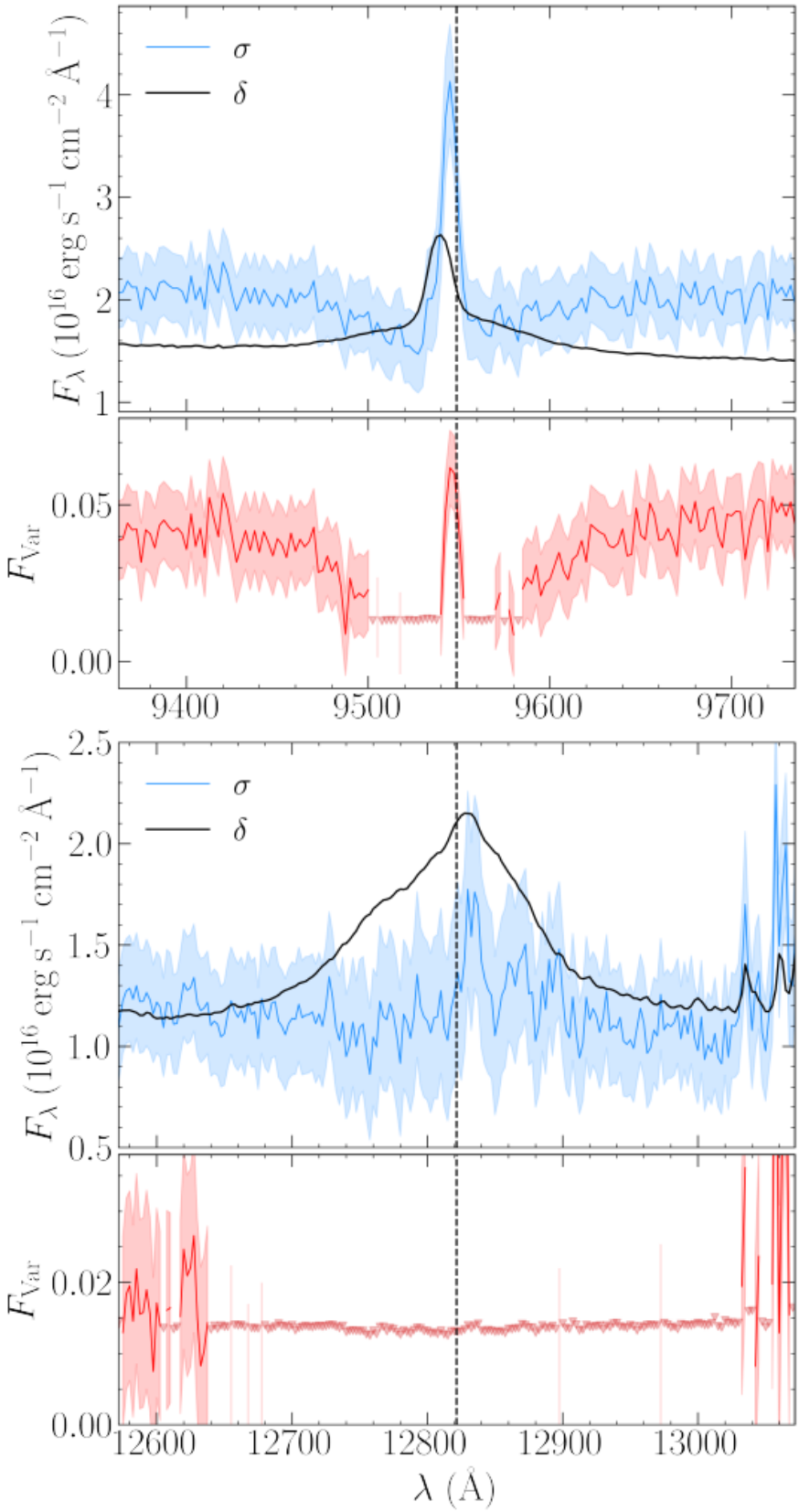

Figure 5.8: δ, σ (top panels) and $F_{\rm var}$ spectrum (bottom panels) for Pa ϵ (top plot) and Pa β (bottom plot) calculated through the process described in Section 5.3.6. $F_{\rm var}$ values are shown in red with their 1σ uncertainties calculated via Monte Carlo methods, represented by the red shaded region, upper limits on $F_{\rm var}$ are shown as red triangles. σ values are displayed as the blue lines along with their 1σ uncertainties calculated via Monte Carlo methods. δ values shown as the black solid line.

Table 5.5: Fractional mean variance ($F_{\rm var}$) in each relevant waveband, unweighted mean flux ($\langle F \rangle$), sample variance (σ) and the RMS uncertainty (δ) calculated following Rodríguez-Pascual et al. (1997).

Waveband	$F_{\rm var}$	$\langle F \rangle$ (mJy)	σ (mJy)	δ (mJy)
g	0.1584±0.0003	9.766	1.55	0.09
LCO - $z_{\rm s}$	0.0851±0.0002	15.88	1.36	0.11
(0.870 - 0.900)μm	0.037±0.012	10.61	0.61	0.46
(0.973 - 0.979)μm	0.042±0.012	9.92	0.60	0.43
H - (1.50-1.55)μm	0.007±0.014	18.08	0.81	0.80
K - (2.25-2.30)μm	0.029±0.013	34.89	1.83	1.54

5.4 Spectral and emission line lightcurves

5.4.1 The observed spectral continuum light-curves

Lightcurves for four spectral continuum windows are constructed from fluxes extracted from near-IR spectra and displayed in Figure 5.9. In addition, we have the g-band photometric light-curve for the accretion disc, shown in Figure 5.2. Ideally, we would like to measure the accretion disc flux at the shortest wavelengths, where it is least contaminated by the hot dust, however, the S/N ratio of the near-IR spectra decreases significantly towards short wavelengths. Therefore, we measured the accretion disc flux in the 60 Å wide rest-frame wavelength region of $\lambda = 9730 - 9790$ Å, which lies between the two broad hydrogen emission lines Pa ϵ and Pa δ and is known to be line-free, and also in the wider (300 Å) rest-frame wavelength region of $\lambda = 8700 - 9000$ Å, which is not contaminated by significant emission-line flux in Mrk 509 (see Figure 5.1).

We measured the flux in two line-free, 500 Å wide rest-frame wavelength regions, namely, $\lambda = 2.25 - 2.30$ μm, which is towards the red end of the K band and thus close to the peak of the blackbody spectrum, and $\lambda = 1.50 - 1.55$ μm, which is close to the middle of the H band. These bands were selected due to the absence of any emission line contamination and the lack of any telluric effects even in the worst quality spectra. These two wavebands have a significantly lower contribution from the accretion disc than the two shorter wavelengths selected. In Figure 5.9 the 30% rise in g-band flux between MJD ∼57830 and ∼58800 has no clear counterpart in the H and K bands as opposed to the shorter-wavelength bands which possess this distinctive rise also seen in our g-band lightcurve, thus demonstrating the reduction in accretion disc contribution well. The reduction in accretion disc contribution at longer wavelengths is well documented in Mrk 509 and similar sources (Hernán-Caballero et al., 2015, 2017; Kuraszkiewicz et al., 2003). Hernán-Caballero et al. (2017) show that at H

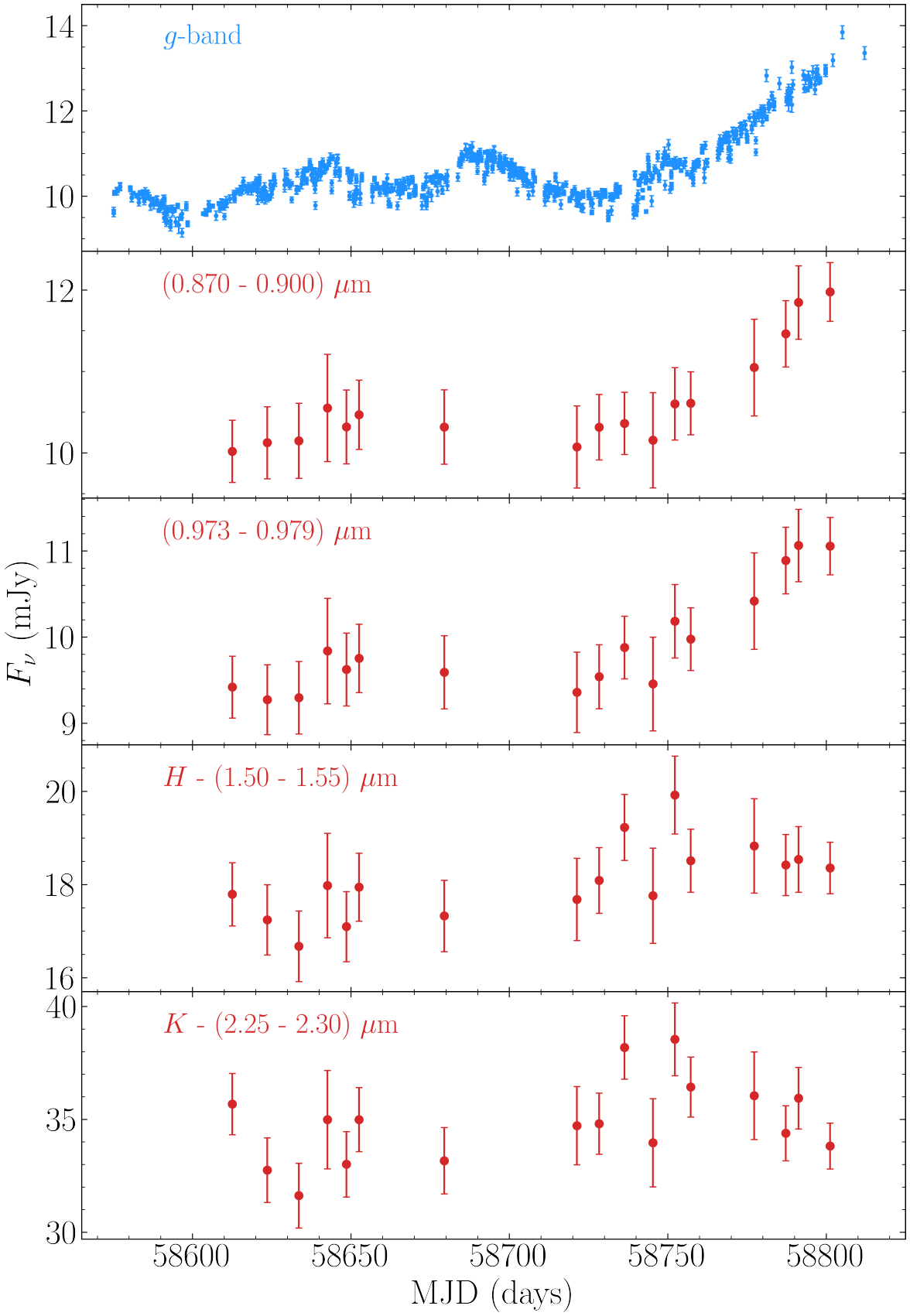

Figure 5.9: Top-bottom, LCO g-band, (0.870-0.900) μm, (0.973-0.979) μm, H-band (1.50-1.55) μm and K-band (2.25-2.30) μm spectral lightcurves. Fluxes have been calibrated onto an absolute scale using MAPSPEC scale factors and then a secondary photometric correction, therefore any variability is considered to be intrinsic.

the contributions from the accretion disc and hot dust are roughly equal, whereas at K the hot dust emission significantly dominates over that of the accretion disc. It is worth noting that the shorter wavelength variability that we sample could contain a significant contribution from diffuse BLR continuum emission, or a wind on the inner edge of the BLR, rather than simply being from the disc (Netzer, 2021; Kara et al., 2021).

$F_{\rm var}$ has been calculated for all four lightcurves. These values are presented in Table 5.5, and show that the variability generally decreases with increasing wavelength and therefore with decreasing contribution from the accretion disc.

From the (126±11) lt-day lag measured by Koshida et al. (2014) and the interferometrically determined radius of (296±31) lt-days (Gravity Collaboration et al., 2017), along with our luminosity weighted dust radius of (186±4) lt-days, it appears that we should expect to measure a lag of (100-200) days. A time lag of this size would necessarily require a smoothing timescale of a similar order meaning that the hot dust emission originating from the torus would not display a high level of variability on short timescales. This could explain why, with our limited baseline and number of observations, 17 data points across 183 days, we detect little variability in the hot dust emission. We would only expect the already low level of variability seen in the H and K-band lightcurves to decrease if the accretion disc component were removed, and therefore we would almost certainly not be sensitive to variability in the hot dust emission alone.

5.4.2 Emission line light-curves

To construct lightcurves, we extract the Pa ϵ and Pa β line fluxes. For Pa ϵ we first fit a linear continuum to redward and blueward clean continuum regions, shown as the red dot-dashed line in Figure 5.3. We subtract this linear continuum, and then integrate under the data using Simpson's method. We must subtract the [S III] component from each observation in order to obtain the flux enclosed by Pa ϵ. As detailed in Section 5.3.3 we opt to scale to the photometry due to major discrepancies in the scaling procedures relying on the [S III] being constant. Therefore the subtracted [S III] values vary by $\sim 6\%$ in line with our calibration procedure. The flux extraction process was slightly more straightforward for the Pa β profile: we fit a linear continuum to relatively clean blueward and redward continuum regions, again displayed by the dashed line in Figure 5.3. The selected redward continuum window is necessarily short in this case as this region is affected by telluric absorption. We then subtract this continuum and simply integrate under the line, using Simpson's method, to give the fluxes displayed

Table 5.6: Fractional mean variance ($F_{\rm var}$) in each relevant waveband, unweighted mean flux ($\langle F \rangle$), sample variance (σ) and the RMS uncertainty (δ) calculated following Rodríguez-Pascual et al. (1997).

Line	$F_{\rm var}$	$\langle F \rangle$ (erg s^{-1}cm^{-2})	σ (erg s^{-1}cm^{-2})	δ (erg s^{-1}cm^{-2})
Pa ϵ	0.060±0.012	0.86	0.06	0.04
Pa β	0.030±0.007	2.860	0.11	0.07

in Figure 5.10. The Paschen line flux uncertainties shown in Figure 5.10 were calculated using the Monte-Carlo method, with 1000 iterations, selecting the 1σ bounds of the resultant Gaussian distribution and then numerically propagating this with a 2% systematic from the continuum selection.

Figure 5.10 shows the lightcurve Pa ϵ for Pa β, it is clear that there is not a high level of variability. This is confirmed by the calculated $F_{\rm var}$ values listed in Table 5.6.

5.5 Response-weighted dust and emission line radii

5.5.1 Continuum and hot dust

JAVELIN is an open source reverberation mapping package (Zu et al., 2011). Through the use of a Bayesian framework alongside an MCMC sampler, JAVELIN creates a statistical model of a 'driving' lightcurve and then assumes 'response' lightcurves can be modelled as smoothed, scaled and shifted versions of this. The main assumption upon which JAVELIN relies is that AGN variability is stochastic, and therefore can be accurately modelled as a Damped Random Walk (DRW).

Utilising the LCO and ZTF g-band data as a 'driving' lightcurve, and the four spectroscopic near-IR lightcurves shown in Figure 5.9 as 'response' lightcurves we run JAVELIN with 10^5 iterations and a (0-100) day limit on the time lag search range in the form of strong preferential priors. This lag range was selected in order to prevent a significant number of spectral data points being shifted into a seasonal gap in the 'driving' lightcurve. The results are displayed in Figure 5.11.

Figure 5.11 shows that the (0.870-0.900) μm and (0.973-0.979) μm bands clearly reverberate on timescales shorter than ∼25 lt-days, whereas the H and K-band lightcurves tend towards longer lags. This is to be expected given that the two shorter-wavelength bands are dominated by the accretion disc, which would reverberate on much shorter time scales than any hot dust emission which would dominate to a larger degree at H and even more so at K.

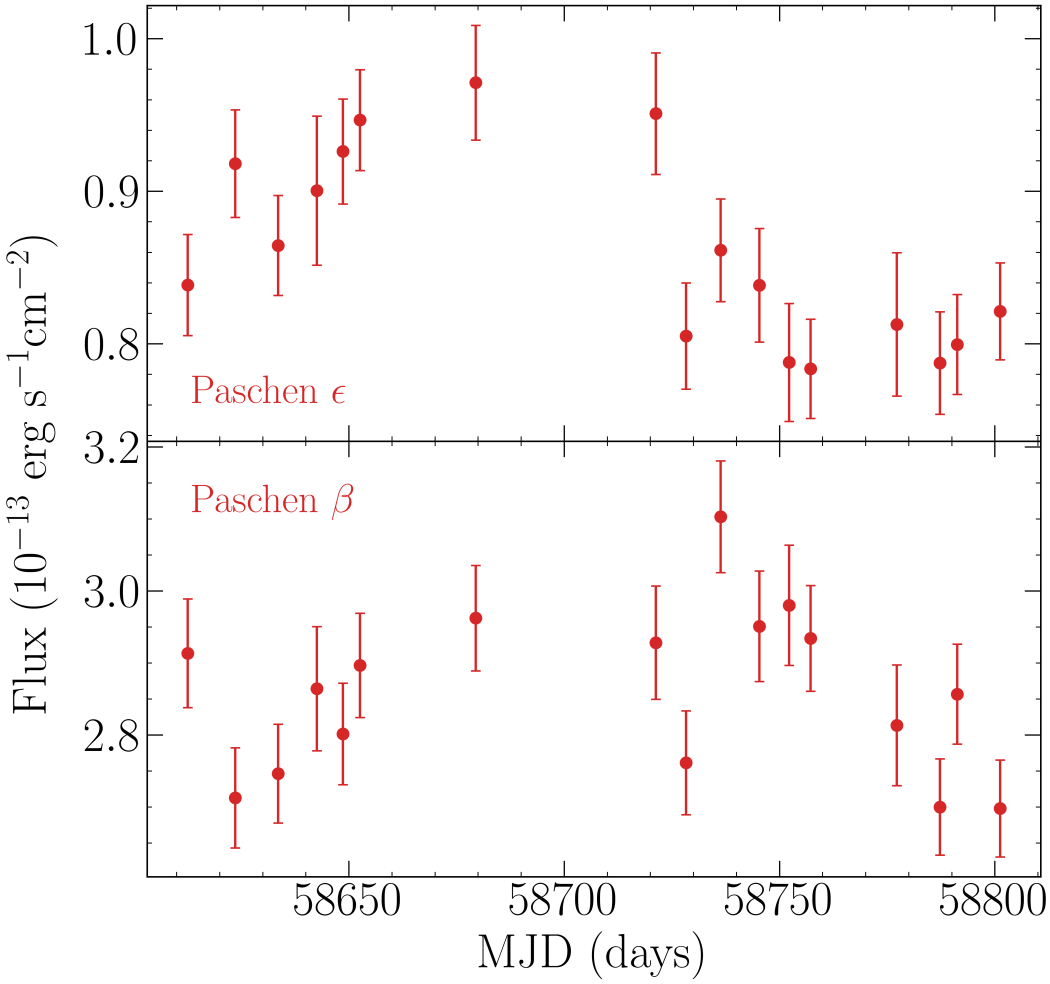

Figure 5.10: Pa ϵ and Pa β emission line lightcurve, constructed using the procedure described in Section 5.4.2.

Due to the limited sampling of our data and ∼150 day seasonal gaps in the 'driving' g-band lightcurve, there are strong aliasing problems, and we cannot therefore report any confident lag. However we can state that the H and K-band lightcurves do not reverberate at <40 days. This is consistent with previous findings. Koshida et al. (2014) measure a reverberated dust lag in the K band of (126.8±11) lt-days, and Gravity Collaboration et al. (2020) measure the spatially resolved torus at a radius of (296 ± 31) lt-days using near-IR interferometry.

We are not only limited by a lack of significant variability features over the majority of our campaign, but also the data gaps present in both the spectral dust, and photometric continuum lightcurves. The large seasonal data gap shown clearly in Figure 5.2 between MJD ∼58450 and ∼58550 limits our ability to measure lags from the hot dust. The K-band lag measured by Koshida et al. (2014) of (126.8±11) lt-days would be extremely difficult to measure with our data-set as a lag of this length would place much of the spectral data into a seasonal data gap, and the length of the lag would smooth out what little variability we do sample in the hot dust response. In fact any lag of more than ∼65 days would place a significant proportion of our already limited data sample into this seasonal gap, which greatly limits the ability to measure a credible hot dust response with our current sampling.

In Appendix: C we present a wider range of positive and negative lags to demonstrate the prevalence of alias lags in all of the spectral lightcurves. We also utilise PyCCF, a python package which performs a cross correlation analysis on two unevenly sampled lightcurves using linear interpolation allowing the measurement of a peak and centroid of the cross correlation function (CCF) (Sun et al., 2018). The PyCCF analysis is presented in Appendix C in Figure C.4. Figure C.4 clearly shows that there are two relatively strong lag peaks in the (0.870-0.900) μm, (0.973-0.979) μm lightcurves close to zero, the uncertainty distribution is strongly skewed towards the negative for both of these wavebands, indicating as we found with JAVELIN that any lag is not likely to be large. The lag distributions for H and K show much more noise with only a strong negative lag, the error distribution here is much wider, yet is skewed to positive values indicating no significant lag, but a bias towards larger values, as found with JAVELIN.

5.5.2 The broad emission line region

Literature values for the H β lag in Mrk 509 place the Balmer BLR at $82.3^{+6.3}_{-5.6}$ days and 79^{+6}_{-7} days for $\tau_{\rm cent}$ and $\tau_{\rm peak}$ respectively, using Cross-Correlation Function (CCF) analysis (Peterson et al., 2004). Unless Mrk 509 has changed state significantly since these measurements

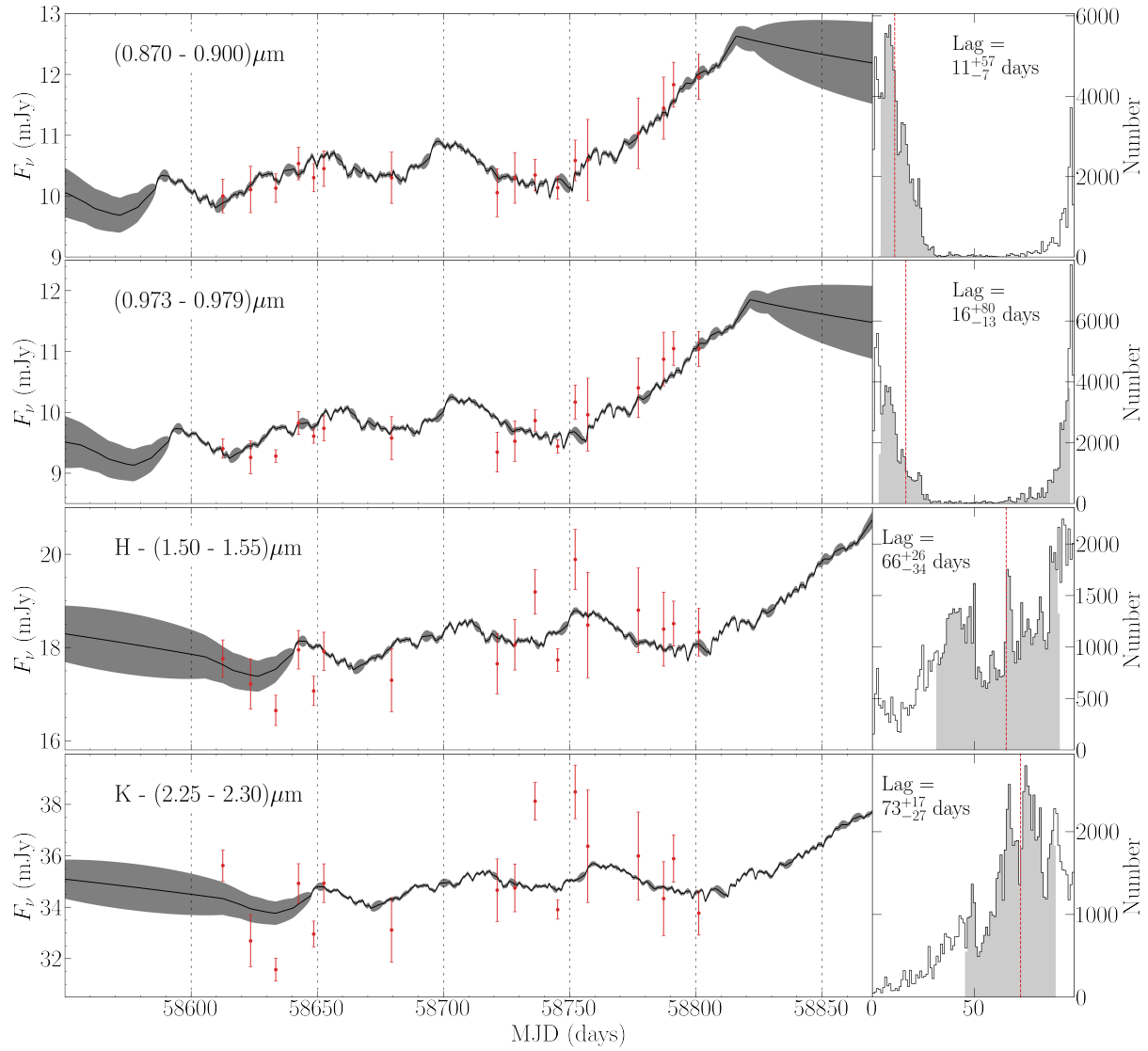

Figure 5.11: JAVELIN results for the four spectroscopic bands, (0.870-0.900) μm, (0.973-0.979) μm, (1.50-1.55) μm, (2.25-2.30) μm, displayed from top to bottom. Each run consists of 10^5 iterations, with the determined lags displayed on the right hand of each panel. The 17 red data points represent spectral data points, and the continuous grey curve is the smoothed, shifted and scaled JAVELIN model of the 'driving' g-band lightcurve displayed in Figure 5.2. Lag limits placed at (0-100) days.

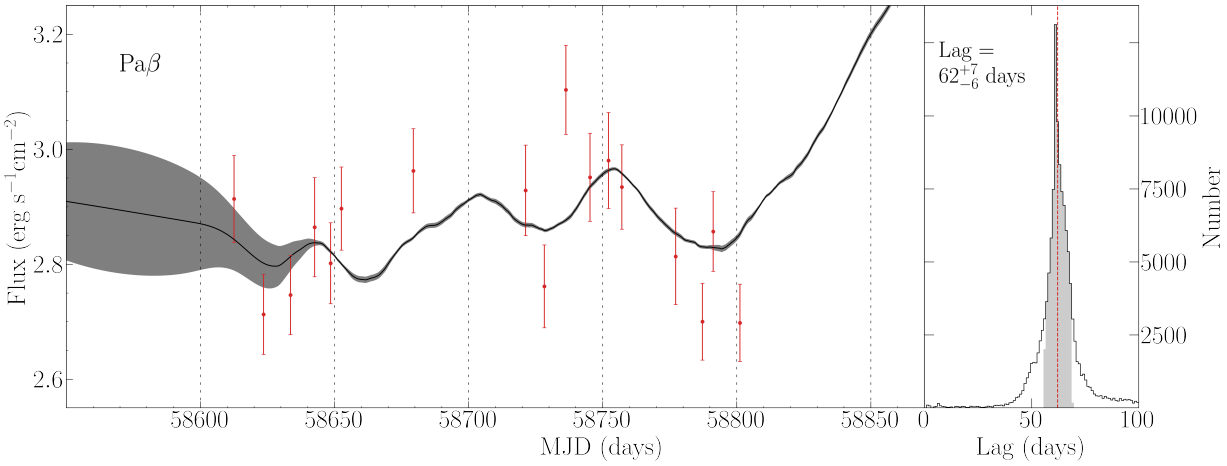

Figure 5.12: JAVELIN results for the Pa β broad emission line. Each run consists of 10^5 iterations, with the determined lags displayed on the right hand of each panel. The 17 red data points represent enclosed emission line fluxes, and the continuous grey curve is the smoothed, shifted and scaled JAVELIN model of the 'driving' g-band lightcurve displayed in Figure 5.2. Lag limits placed at (0-100) days.

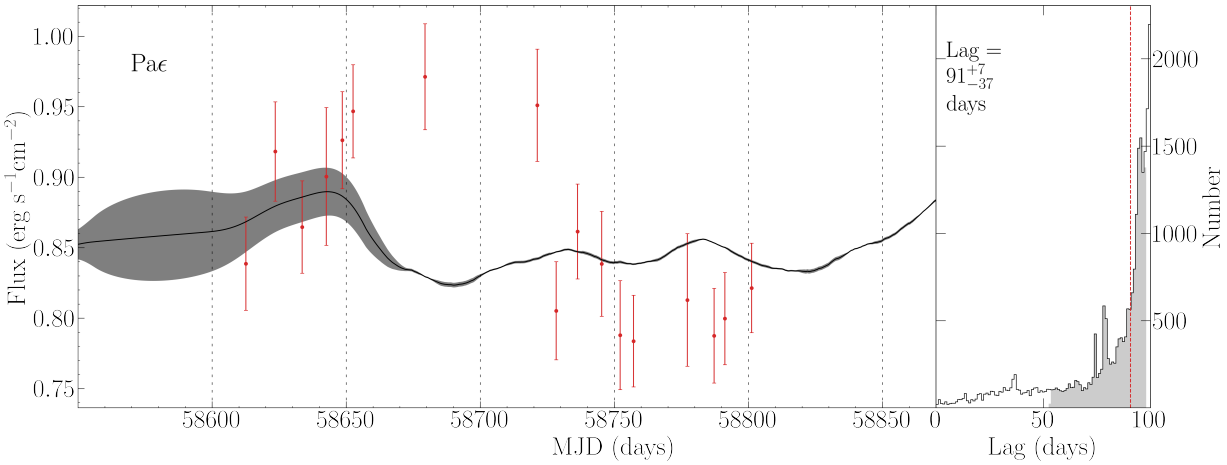

Figure 5.13: JAVELIN results for the Pa ϵ broad emission line. Each run consists of 10^5 iterations, with the determined lags displayed on the right hand of each panel. The 17 red data points represent enclosed emission line fluxes, and the continuous grey curve is the smoothed, shifted and scaled JAVELIN model of the 'driving' g-band lightcurve displayed in Figure 5.2. Lag limits placed at (0-100) days.

were taken (1989-1996), we would expect to see the Paschen line lags to at least be this large. As with measuring reverberation signals in the hot dust of Mrk 509, this presents a problem due to the sparse sampling, short baseline of observations and data gaps present in our data.

As in the case of the hot dust lightcurves, we used JAVELIN to search for Paschen line reverberation signals, the results of which are displayed in Figure 5.12.

JAVELIN reports a lag of 62^{+7}_{-6} days in Pa β. This result is consistent (within 2σ) with previous measurements of the H β lag (Peterson et al., 2004). Our reverberation result of 62^{+7}_{-6} days, does show some smoothing, with a value of 21^{+29}_{-19} days, however this is not as large as we would expect for a lag of this size. In addition, Figure 5.12 shows that with a lag of 62 days, 3 out of 17 data points are shifted into a data gap and do not align with any optical data points. Upon visual inspection, we see that the remaining 10 data points do not show a good agreement with the pattern of variability displayed by the optical lightcurve. Therefore we do not report this result with a high level of confidence. However, as with the H and K-band spectral lightcurves, we are able to place a lower limit on the Pa β lag of ∼40 days as we see no counterpart in the lightcurve of the distinctive rise in flux in the last 80 days of the campaign, marked in red in Figure 5.2.

We can place this same lower limit on Pa ϵ which shows the same distinct absence of the major rise in flux present in our 'driving' lightcurve. We measure the Pa ϵ at 91^{+7}_{-37} days, shown in Figure 5.13. This lag places 6 of our 17 data points into a data gap and therefore is not reported with a high level of confidence despite being consistent with previous measurements of the H β lag (Peterson et al., 2004).

JAVELIN results for Pa ϵ and Pa β are displayed in separate plots as the analysis was carried out separately. Running JAVELIN separately resulted in a fit with higher level of smoothing in both cases, and therefore in some senses a more physical result.

In Appendix: C we present a wider range of positive and negative lags to demonstrate the prevalence of alias lags in both Pa ϵ and Pa β with both JAVELIN and PyCCF as with the four spectral lightcurves. The lag distributions obtained with PyCCF, presented in Figures C.6 and C.5 are very similar for both Pa ϵ and Pa β. There are very few lag peaks with an extremely large uncertainty rage on the centroid lags in both cases. This result indicates that there is strong aliasing and no significant lag to report.

5.6 Summary and conclusions

We monitored the AGN Mrk 509 to document its infrared (hot dust) and optical (accretion disc) variations. Our IRTF spectra span 0.7 to 2.4 microns at 17 epochs over 200 days from 2019 May to Nov. The g-band monitoring spans 5 seasons with sub-day cadence during the IR campaign.

Using spectral decomposition we measure luminosity weighted dust radii for different astrochemical regimes to be $\langle R_{\rm d,lum} \rangle = 186 \pm 4$ light-days, 1763±36 light-days and 1132±23 light-days in the case of a blackbody, small-grain silicate and carbon dust, respectively. Comparison with the previously measured photometric dust response time of (126±11) days (Koshida et al., 2014) and an interferometrically determined radius of (296 ± 31) lt-days (Gravity Collaboration et al., 2020) suggests that the dust emits like a black body and therefore is most likely to be composed of large carbonaceous grains. This is consistent with results for the Seyfert NGC 5548 (Landt et al., 2019).

Our g-band lightcurve shows a distinctive rise by ∼30% in the last 80 days of the 2019 season, with no clear counterpart in the H and K lightcurves, which have intrinsic variations less than ∼4%. The little variability we do detect in H and K does not reverberate on timescales less than 40 days and therefore we find a lower limit on the dust response in Mrk 509 which is consistent with previously measured reverberation measurements and also interferometric measurements of the hot dust radius (Koshida et al., 2014; Gravity Collaboration et al., 2020).

We are also limited in our ability to recover a dust lag due to seasonal gaps in both the driving and response lightcurves. This combined with the poor cadence and low levels of variability in the dust response means that application of a large smoothing width would remove any significant variability, which is already low.

We measure lags of 62^{+7}_{-6} days and 91^{+7}_{-37} days for Pa β and Pa ϵ respectively. The Pa β lag is considered a more reliable measurement than that determined for Pa ϵ due to a more concentrated peak, fewer data points placed in seasonal gaps of the 'driving' lightcurve and a more reasonable smoothing result. The Pa ϵ result is consistent with literature values within 1σ and the Pa β result consistent within 2σ (Peterson et al., 2004). However due to the seasonal gaps in our 'driving' lightcurve and poor sampling of the emission line lightcurves, we do not report either of these results with a high level of confidence. As with the H and K-bands, we can place a lower limit on the lag of ∼40 days as we do not detect any lags

below this value and do not recover the distinctive variability feature seen in our 'driving' lightcurve during the last ∼80 days of our spectral campaign.

We construct mean, mean variance and $F_{\rm var}$ spectra, these confirm that over the time frame of our observations the hot dust shows relatively little significant variability compared with the significant variability detected in the accretion disc at wavelengths $<1\mu$m.

A longer baseline of observation, and a higher cadence is needed to recover a reverberated signal from the hot dust and place constraints on the physical size scales of this emission region. The main result of this work is to present an honest discussion of the difficulties and limitations inherent to performing ground based near-IR spectroscopic reverberation mapping and to develop techniques with which to scale spectra using a challenging dataset. These techniques have been applied to a much more robust dataset in Chapter 6.

Chapter 6

STORM2 - Near-IR spectroscopic reverberation mapping on Mrk 817

6.1 Introduction

This work presented in this Chapter is part of the STORM2 campaign, a large collaboration taking simultaneous multi-wavelength photometry and spectroscopy focusing on the type-1 AGN Mrk 817 (Kara et al., 2021).

STORM2 aims to provide detailed insights into the multi-wavelength behaviour of Mrk 817 over the course of ∼2 years of observations. The basis for this campaign was the success of the AGN Space Telescope and Optical Reverberation Mapping program (AGN STORM) which focused on NGC 5548 (De Rosa et al., 2015; Kriss et al., 2019). There were many interesting and unexpected results from AGN STORM, but the main insight was that intensive multi-wavelength monitoring can be an incredibly powerful tool in determining the relationships between different spectral properties. STORM2 should be an even more fruitful endeavour due to its wider range in wavelength and longer duration.

One notable difference is the addition of near-IR spectroscopy allowing for reverberation mapping of the Paschen lines simultaneously with the Balmer lines from optical spectroscopy. Understanding the connection between the near-IR and optical BLR is of great importance, especially given the spatial resolution of the Paschen BLR of 3C 273 by Gravity Collaboration et al. (2018) and the potential for this technique to be applied to other sources.

With near-IR spectroscopy, is it also possible to perform reverberation mapping on the hot dust continuum as shown in Chapter 5 and in Landt et al. (2019).

6.1.1 This Chapter

Here we build upon the work presented in Chapter 5 with a much larger dataset. Using the photometric scaling techniques presented in Chapter 5 we attempt to constrain the spatial extent of the inner torus and the Paschen BLR. Using two different methods we obtain consistent Pa β lags which alongside contemporaneous literature values of the H β lag in this source, provides evidence that the Paschen and Balmer Broad line regions are co-spatial.

6.2 The science target

As part of the STORM2 campaign, this work focuses on the type-1 AGN Mrk 817 (Kara et al., 2021). This target is a suitable subject for a near-IR spectroscopic reverberation mapping campaign such as this for several reasons: (i) it has a previously measured, short and variable dust response time of (89.8 ± 8.6) days (Koshida et al., 2014), therefore within our 494 day baseline of observations we should theoretically be able to recover multiple lags and investigate any variations; (ii) it has clearly discernible broad emission lines in the Paschen series allowing us to separate narrow- and broad-line components; (iii) multiple reverberation campaigns have determined its black hole mass. Peterson et al. (2004) reported the black hole mass to be $(4.94 \pm 0.77) \times 10^7$ $M_\odot$.

As part of the STORM2 campaign Mrk 817 has undergone extensive multi-wavelength monitoring, preliminary continuum and broad emission line lags across this range have been published in Kara et al. (2021). This will allow for useful comparison with our reverberation results for the hot dust emission and broad Paschen lines. This work shall therefore contribute to the formation of a relatively comprehensive picture as to the physical geometry of Mrk 817.

Mrk 817 (J2000 sky coordinates R.A. $14^h36^m22.068^s$ and so Decl. $+58°47'39.38''$) can be observed from the northern hemisphere. It is at a relatively low redshift of $z = 0.0313$, thus enabling us to sample the hot dust component and several near-IR coronal lines using a cross-dispersed near-IR spectrum. This AGN is relatively bright in the near-IR (2MASS $J = 12.9$ mag, $K = 10.9$ mag; Skrutskie et al. (2006)) ensuring good quality near-IR spectra are possible using reasonable exposure times on a 4 m class telescope.

We adopt here cosmological parameters $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$, which give a luminosity distance to Mrk 817 of 137 Mpc and an angular scale at the source of 626 pc per arcsec.

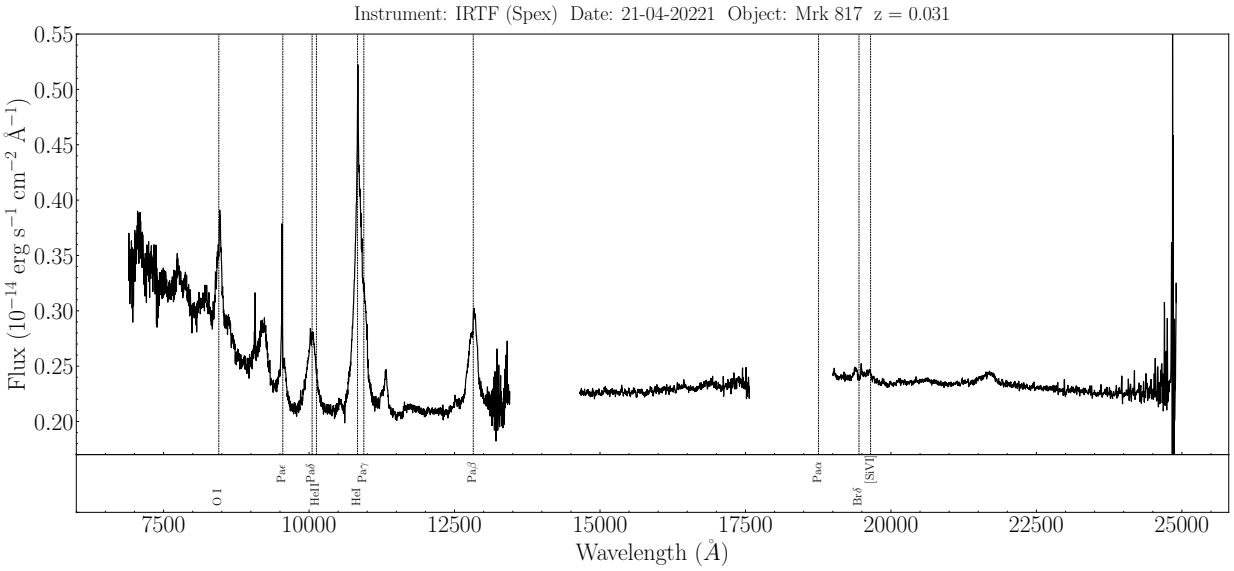

Figure 6.1: IRTF SpeX near-IR spectrum from 2021 April 21 shown as observed flux versus rest-frame wavelength. Several emission lines, including the Paschen Hydrogen series are marked with dashed lines.

6.3 The observations

The aim of this work is to measure a reverberated signal in the near-IR hot dust emission and the Paschen emission lines of Mrk 817, simultaneously with taking measurements of the dust temperature. We therefore monitor Mrk 817 using near-IR spectroscopy and also require well sampled optical photometry in order to trace the accretion disc.

6.3.1 The near-IR spectroscopy

Using the Spex spectrograph at the 3 m NASA Infrared Telescope Facitity (IRTF) on Manaukea, Hawaii, we observed the source Mrk 817 between February 2021 and June 2022 (semesters 2021A and 2021B and 2022A) (Rayner et al., 2003). Having requested 52 nights of observing in order to obtain a cadence of ∼1 week, we were scheduled 50, 2-2.5 hour, observing windows between February 2021 and June 2022. We obtained a total of 40 near-IR spectra as we lost thirteen of the scheduled observations due to weather and technical issues and were awarded an extra night of observations on 22nd July 2022. The journal of observations is listed in Table 5.1.

All observations were carried out using the (0.7-2.55) μm short cross-dispersed mode (SXD) equipped with the 0.3"x15" slit, oriented at the parallactic angle. Using this narrow slit allowed us to minimise the contamination from the host galaxy, whilst giving us a sufficiently

high spectral resolution, of R=2000 or full width at half maximum (FWHM) $\sim$ 150 km s^{-1}, in order to discern clearly any narrow and broad emission lines and study their profiles. We observed the source using a standard ABBA nodding pattern. As the source is not extended we nodded along the slit rather than on sky, therefore obtaining an image of the target in both the A and B frames from which the background subtraction could be carried out. The usual on-source exposure time was 32×120s. On 28 of the 40 successful observing runs we obtained more exposures than this, on 6 we obtained fewer exposures, and 6 we obtained the expected 32 exposures. This exposure time was chosen in order to guarantee high enough (S/N) to reliably measure not only the broad, but also the narrow emission line profiles, which were key in scaling the spectra onto an absolute spectral flux scale.

It was necessary to observe a standard star in order to correct our science spectrum for telluric absorption and for flux calibration. Therefore after the science target we observed a nearby (in airmass and position) A0 V star, HD 121409 which has accurate optical magnitudes (B = 5.66 and V = 5.70). Where possible we took flats and arcs after the science target. Flux calibration onto an absolute spectral flux scale in order to investigate intrinsic variability is carried out through a separate process which is detailed in section 5.3.3.

The data were reduced using SPEXTOOL (version 4.1) a software package created for use by Spex users and based in Interactive Data Language (IDL) (Cushing et al., 2004). This package was used on both the science target and telluric standard. Spextool carries out all of the procedures required in the production of fully reduced spectra. This includes the preparation of calibration frames, the subtraction of flat fields, the fitting and subtraction of local background, interactive spectral extraction from science frames using an optimally weighted extraction scheme (Horne, 1986), telluric correction of the science spectra, merging of orders, and combination of the different frames into a single reduced science spectrum for each observing window. The final spectrum was corrected for Galactic extinction using a value of $A_{\rm V}$ = 0.0181, which was obtained from the IRSA/IPAC Galactic Dust Reddening and Extinction service using values derived from Schlafly & Finkbeiner (2011). Figure 5.1, we show the spectrum from 2021 April 21 as a representative example.

The LCOGT observations

Mrk 817 was monitored between 2020 November 24 and 2022 May 16 with a mean cadence of $\sim$ 0.5 days, using the following facilities; Calar Alto 2.2m (CAFOS - V), LCO 1m (Sinistro - BV, u', g', r', i', z_s) (LCOGT; Brown et al., 2013), Liverpool Telescope 2m (IO:O -

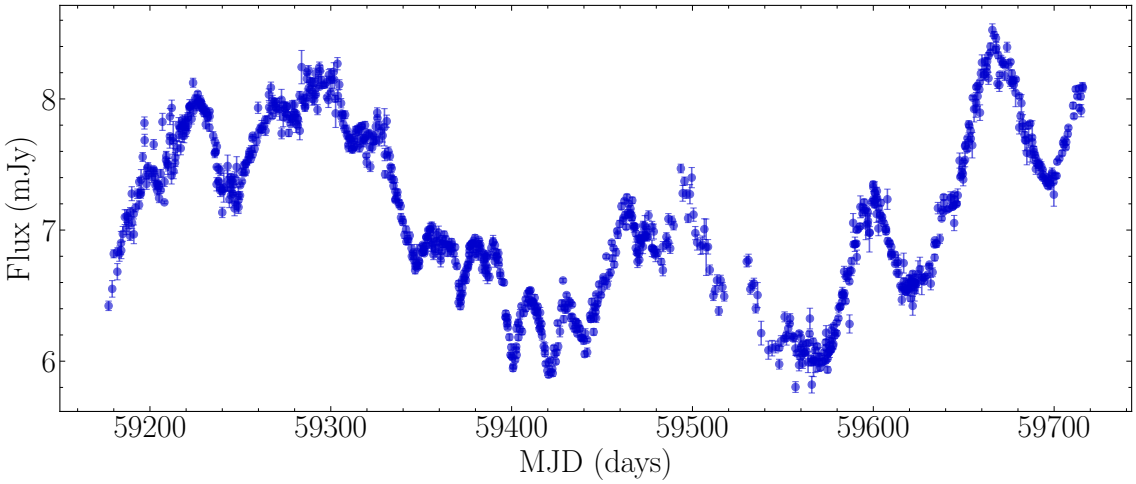

Figure 6.2: g-band lightcurve for Mrk 817.

u', g', r', i', z') (Steele, 2004), Wise Observatory 18-inch (QSI683 - g', r', i', z') (Brosch et al., 2008), Yunnan Observatory 2.4m (YFOSC - V) and Zowada Observatory 20-inch (g', r', i', z_s). Preliminary versions of all of these lightcurves, covering the time between 2020 November 24 and 2021 February are presented and discussed in Kara et al. (2021). The photometric lightcurves we utilise are the full g' and z_s filters, which have a central wavelength of 4770 Å and 8700 Å and a width of 1500 Å and 1040 Å respectively. The baseline of these observations begins ∼2 months before the start of and ends ∼2 months before the end of our near-IR campaign.

The standard image processing pipelines for each facility were used for the first steps of data reduction. Photometry was then carried out using an automated procedure added to the ASTROPY PHOTUTILS package (Astropy Collaboration et al., 2018) and described in detail in Bradley et al. (2020), for all of the facilities except Calar Alto and Yunnan. However we do not utilise the V-band data in this work.

Given that some facilities observed the source more than once per night, observations from the same facility in the same night were combined using a weighted average. The g'-band and z_s-band optical lightcurves we present use the data from multiple facilities, which was inter-calibrated using PyCALI (Li et al., 2014). This calibration was necessary due to the differences in wavelength-dependent throughput for each telescope. PyCALI uses a Bayesian approach, utilising the diffuse nested sampling algorithm of Brewer et al. (2009), in order to calculate additive and multiplicative factors used to scale the photometry onto a common flux scale. The uncertainties are also increased systematically by adding an extra term in quadrature with the original photometric errors.

Due to calibration issues the WISE data has been excluded from the z-band lightcurve, which is therefore a combination of the LCO z_s, Liverpool Telescope z', and Zowada Observatory z_s-band lightcurves. There is no significant difference in the Pan-STARRS z_s and SDSS z' lightcurves in terms of shape and reverberation lag, therefore these lightcurves could be easily combined (Kara et al., 2021).

Throughout this work we shall refer to the relevant composite lightcurves as simply the g-band and z-band.

6.4 Photometric Corrections to the Spectra

For the purposes of reverberation mapping, as with the work carried out on Mrk 509, it is essential to distinguish between intrinsic AGN variability and extrinsic sources of variability such as changing night to night observing conditions. We therefore perform a photometric correction to the spectra to obtain meaningful lightcurves. Failing to carry out this important calibration could result in a contamination of the final dust response lightcurves (Bentz et al., 2016).

One approach to correcting for time-dependent instrumental effects is to select a spectral feature which is intrinsically non variable over the timescale of a campaign. For optical campaigns, this is usually the [O III] λ5007 narrow emission line (Fausnaugh, 2017). However, this emission line is not visible in the near-IR, therefore we have selected the [S III] λ9531 narrow emission line, which is analogous to the [O III] λ5007 region due to its extent.

Several software packages employ different approaches to perform the photometric re-scaling of spectra, calibrating out time-dependent instrumental effects, e.g PREPSPEC (written by Keith Horne and discussed in Shen et al. 2016) and GW92 (Van Groningen & Wanders, 1992). This work uses the open source PYTHON package MAPSPEC (Fausnaugh, 2017) which is an updated and improved version of the GW92 package. MAPSPEC , like GW92, aligns an assumed non variable line profile of a given observed spectrum to a reference spectrum that has been constructed, by applying a smoothing kernel, a flux re-scaling factor, and a wavelength shift.

Using MAPSPEC, a reference spectrum was constructed from 7 photometric spectra. Here we opt to use the Pa ϵ broad line profile as the continuum level. This reference spectrum was used alongside all 40 spectra in order to calculate re-scaling factors. MAPSPEC allows for three different smoothing methods, a Delta function, a Gaussian and Gauss-Hermite smoothing

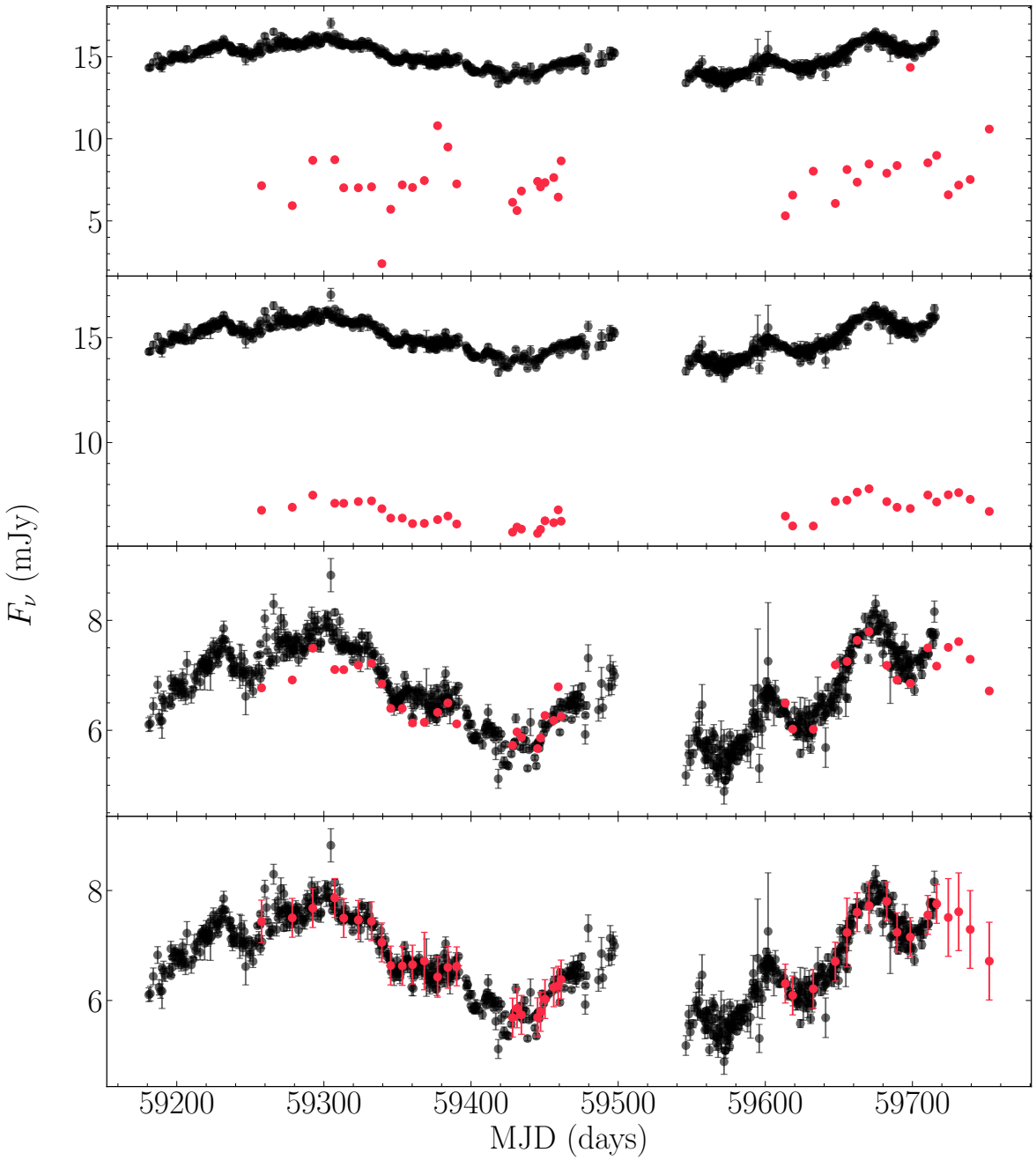

Figure 6.3: **Top Panel:** z band photometry, uncorrected for host galaxy contribution in black, with all 40 spectral z band (8180-9220 Å observed frame) uncorrected fluxes plotted in red. **Second Panel:** z band photometry, uncorrected for host galaxy contribution in black, with all 40 spectral z band (8180-9220 Å observed frame) fluxes, corrected with MAPSPEC scaling factors and plotted in red.**Third Panel:** z band photometry, corrected for a constant host galaxy contribution in black, with all 40 spectral z-band (8180-9220 Å observed frame) fluxes, corrected with the MAPSPEC scaling factors plotted in red. **Bottom Panel:** The MAPSPEC corrected lightcurve has been further corrected to the z band photometry, by shifted to the weighted mean of all photometric data points within three days of each spectral data point. The errors displayed are defined by the spread of the data to which the spectral points have been corrected, or the mean spread of the photometry within any given 3 day period of the spread of the correction was less than this value. These uncertainties have been propagated to our corrected lightcurves along with the MAPSPEC uncertainties displayed in the panel above.

kernel. We have taken the final scaling values as the Gauss-Hermite smoothing value as this is the most stable of the three methods (Fausnaugh, 2017).

For Mrk 509, we develop a technique of photometric correction that uses MAPSPEC as a first order correction, followed by a small second order correction factor to align z-band flux points extracted from each spectrum over the z_s bandpass (henceforth referred to as spectral z-band points) with the z-band photometry (see Chapter 5). Figure 6.3 shows the results of this procedure in the case of Mrk 817. We correct for a constant host galaxy contribution of $\sim$10 mJy. The initial MAPSPEC correction is much more accurate that in the case of Mrk 509, with no obviously discrepant points (see Figure 6.3, third panel). This indicates that the [S III] emission is not significantly extended beyond the slit width, or at least not as extended as we believe was the case for Mrk 509. However, the agreement with the photometry is of variable accuracy over the course of the observations, showing a much better agreement in the latter half of the campaign. This, of course, would be reversed by using a very slightly different host galaxy correction, however it is clear that the shape of the spectral lightcurve is different to that of the photometry. We therefore perform the secondary correction to the photometry.

We calculate the uncertainty on the correction by using the spread of the photometry. To do this, we bin the photometric data into 3-day windows around each spectral data point and scale to the weighted mean of this value, we then propagate the weighted standard error of this bin forward as the scaling uncertainty, and in any bin where this is less than the average spread of the data in any 3-day period we adopt this value as the uncertainty on the correction and therefore propagate this error forward, adding in quadrature with the measurement error on each spectral data point.

6.5 The mean and Variable $F_{\rm var}$ spectra

The Fractional mean variance ($F_{\rm var}$) is a measure of fractional "excess variance", and therefore is a measure of intrinsic variability above that due to flux measurement errors. $F_{\rm var}$ is calculated as,

$$F_{\rm var} = \frac{(\sigma^2 - \delta^2)^{1/2}}{\langle F \rangle}, \tag{6.5.1}$$

following Rodríguez-Pascual et al. (1997), where $\langle F \rangle$ is the sample mean of the $N = 40$, $\langle F \rangle$ is the unweighted mean, σ is the square root of the sample variance and δ^2 is the mean square uncertainty of the fluxes. All are fully defined in and Chapter 5.3.6.

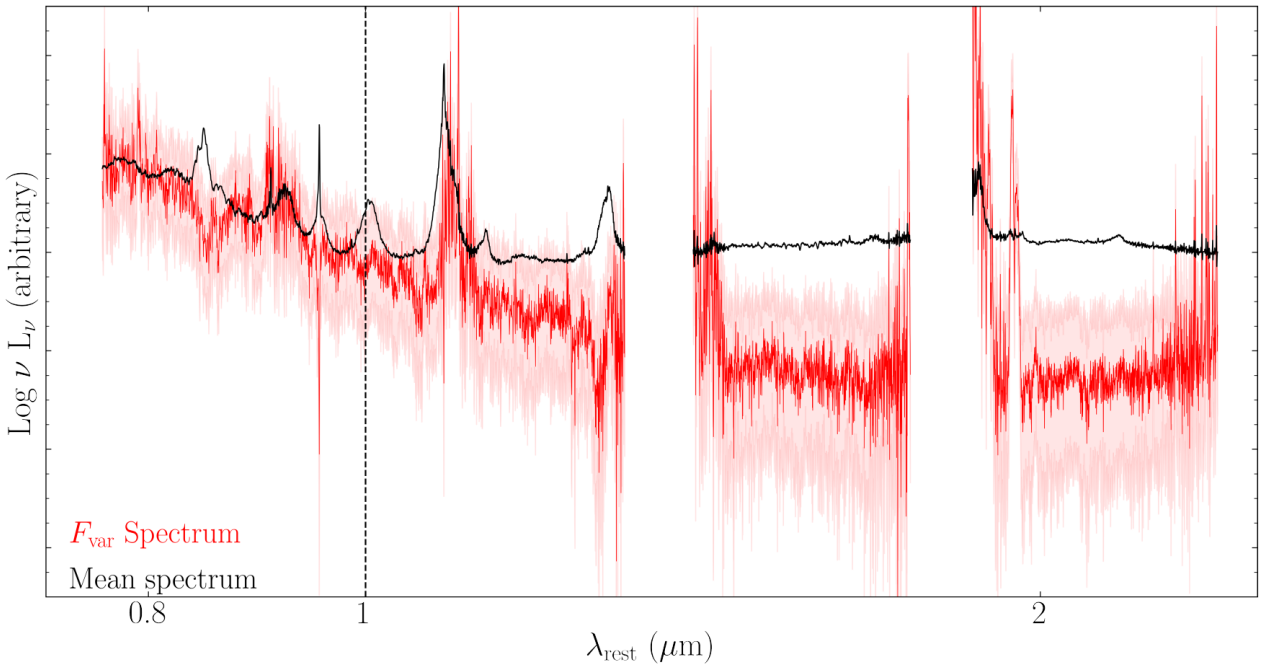

Figure 6.4: The mean (black) and variable ($F_{\rm var}$) spectrum (red) composed of the 40 spectra taken during our campaign, normalised at rest-frame $\sim$ 1 μm (vertical dashed line). The faded red region represents 1σ errors on our $F_{\rm var}$ spectrum calculated through MC methods. Below $\sim$ 1 μm the spectrum is dominated largely by the accretion disc. Beyond this value the accretion disc has a diminishing effect and towards $\sim$ 2 μm, the spectrum is dominated by hot dust emission.

Figure 6.4 shows the mean variance (numerator of $F_{\rm var}$ in Equation 6.5.1) calculated for each spectral element for the 40 observations of our campaign. These values have been plotted against wavelength to create an intrinsic mean variance spectrum. Both the mean (black) and mean variance (red) spectra were normalised to a clean continuum window just blueward of 1μm, and are shown with their 1σ uncertainties derived using Monte-Carlo methods. It is clear that for Mrk 817 we are sensitive to much more variability than in the case of Mrk 509 (see Chapter 5.3.6 and Figure 5.7). This does not necessarily indicate that Mrk 817 is more intrinsically variable than Mrk 509, but is most likely due to the longer baseline and cadence of the Mrk 817 campaign. As with Mrk 509, most of the variability exists at wavelengths $<$ 1μm where the mean variance spectrum matches the shape of the mean spectrum. Beyond this threshold, the mean variance spectrum falls away from the shape of the mean spectrum indicating that the level of variability is lessened. As we show for Mrk 509, this is likely due to the reduction in the contribution of the accretion disc at higher wavelengths (Hernán-Caballero et al., 2017). Hernán-Caballero et al. (2017) show that for Mrk 817, the contribution from the accretion disc and the hot dust about equal in the H-band, and the hot dust is far more dominant in the K-band. They also show that the relative contributions from the hot

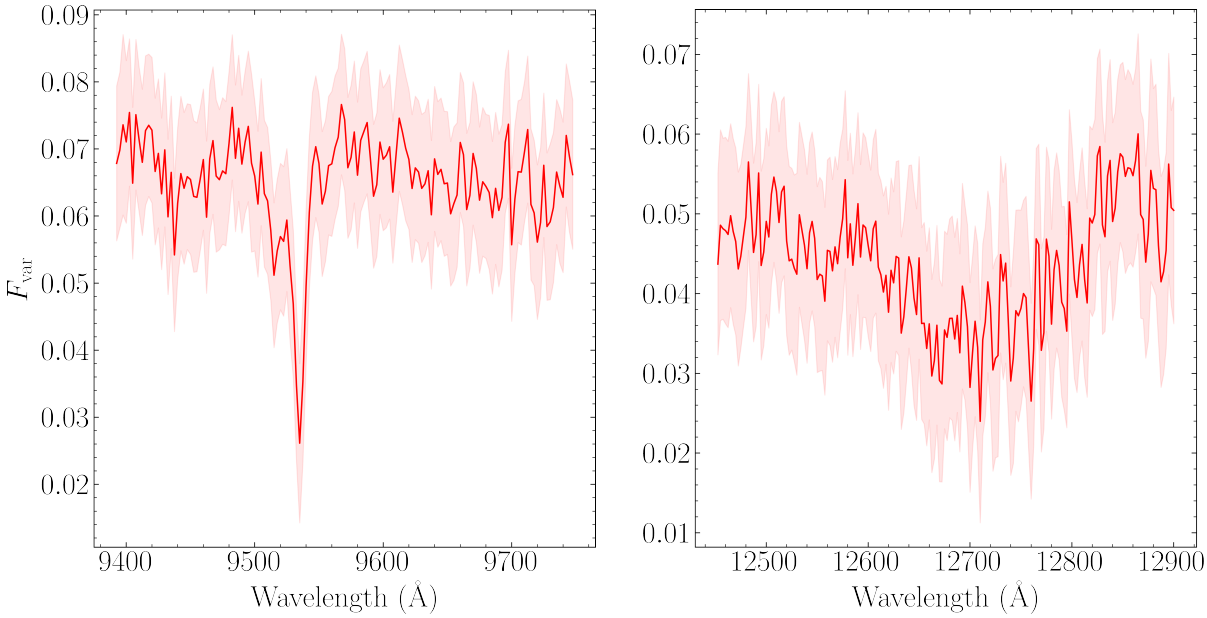

Figure 6.5: $F_{\rm var}$ (red) with 1σ errors (red faded) covering the Paschen ϵ (left) and Paschen β (right) spectral regions.

dust in both the H and K-bands is larger in the case of Mrk 817 than for Mrk 509. Due to the dominance of the hot dust response over that of the accretion disc, particularly in the K-band, we do not perform a full spectral decomposition.

Figure 6.5 shows the $F_{\rm var}$ spectra (calculated as detailed in Equation 5.3.2) for the Pa ϵ and Pa β regions. The $F_{\rm var}$ values are plotted along with their 1σ uncertainties. We compute and plot an upper limit for any spectral element for which the δ value exceeds the σ value. It is clear that the broad line profiles for both Pa ϵ and Pa β display a lower level of variability than the surrounding continuum regions. In the case of Pa ϵ the [S III] narrow line profile shows a much lower level of variability than the Pa ϵ broad line profile. This is the opposite trend than is seen in Mrk 509, providing further evidence that in Mrk 817 the [S III] emission region is not significantly extended beyond the slit.

6.6 Spectral and emission line lightcurves

6.6.1 The observed spectral continuum lightcurves

Lightcurves for four spectral continuum windows are constructed from fluxes extracted from near-IR spectra and displayed in Figure 5.9. In addition, we have the g-band photometric light-curve for the accretion disc, shown in Figure 6.2. Ideally, we would like to measure the accretion disc flux at the shortest wavelengths, where it is least contaminated by the

hot dust, however, the S/N ratio of the near-IR spectra decreases significantly towards short wavelengths. Therefore, we measured the accretion disc flux in the 60 Å wide rest-frame wavelength region of $\lambda = 9730 - 9790$ Å, which lies between the two broad hydrogen emission lines Pa ϵ and Paδ and is known to be line-free, and also in the wider (300 Å) rest-frame wavelength region of $\lambda = 8700 - 9000$ Å, which is not contaminated by significant emission-line flux in Mrk 817 (see Figure 6.1).

As in Mrk 509, we measured the flux in two line-free, 500 Å wide rest-frame wavelength regions, namely, $\lambda = 2.25 - 2.30$ μm, which is towards the red end of the K band and thus close to the peak of the blackbody spectrum, and $\lambda = 1.50 - 1.55$ μm, which is close to the middle of the H band. These bands were selected due to the absence of any emission line contamination and the lack of any telluric effects even in the worst quality spectra. Hernán-Caballero et al. (2017) show that at H the contributions from the accretion disc and hot dust are roughly equal, whereas at K the hot dust emission significantly dominates over that of the accretion disc. It is worth noting that the shorter wavelength variability that we sample could contain a significant contribution from diffuse BLR continuum emission, or a wind on the inner edge of the BLR, rather than simply being from the disc (Netzer, 2021; Kara et al., 2021).

$F_{\rm var}$ has been calculated for all four lightcurves. These values are presented in Table 6.1, and show that the variability generally decreases with increasing wavelength and therefore with decreasing contribution from the accretion disc. The trend of decreasing variability with increasing wavelength is clear to see from Figure 6.6. Most notably the $\sim$25% increase in the g-band flux between MJD 59560 and MJD 59660 is clearly visible in the (0.870-0.900) μm and (0.973-0.979) μm response. However there is a much less prominent increase in flux in the H and even more so the K-band, which takes place over a longer timescale. The same is true of the decrease in flux seen at the start of the campaign, with the response to the g-band being much less pronounced with increasing wavelength.

Even in the g-band, we calculate the low value $F_{\rm var}$=0.0893$\pm$0.0002, meaning that Mrk 817 does not display high levels of intrinsic variability across our campaign. Therefore the dust response in the H and K-bands shows even less intrinsic variability. This relatively low level of variability makes reverberation challenging.

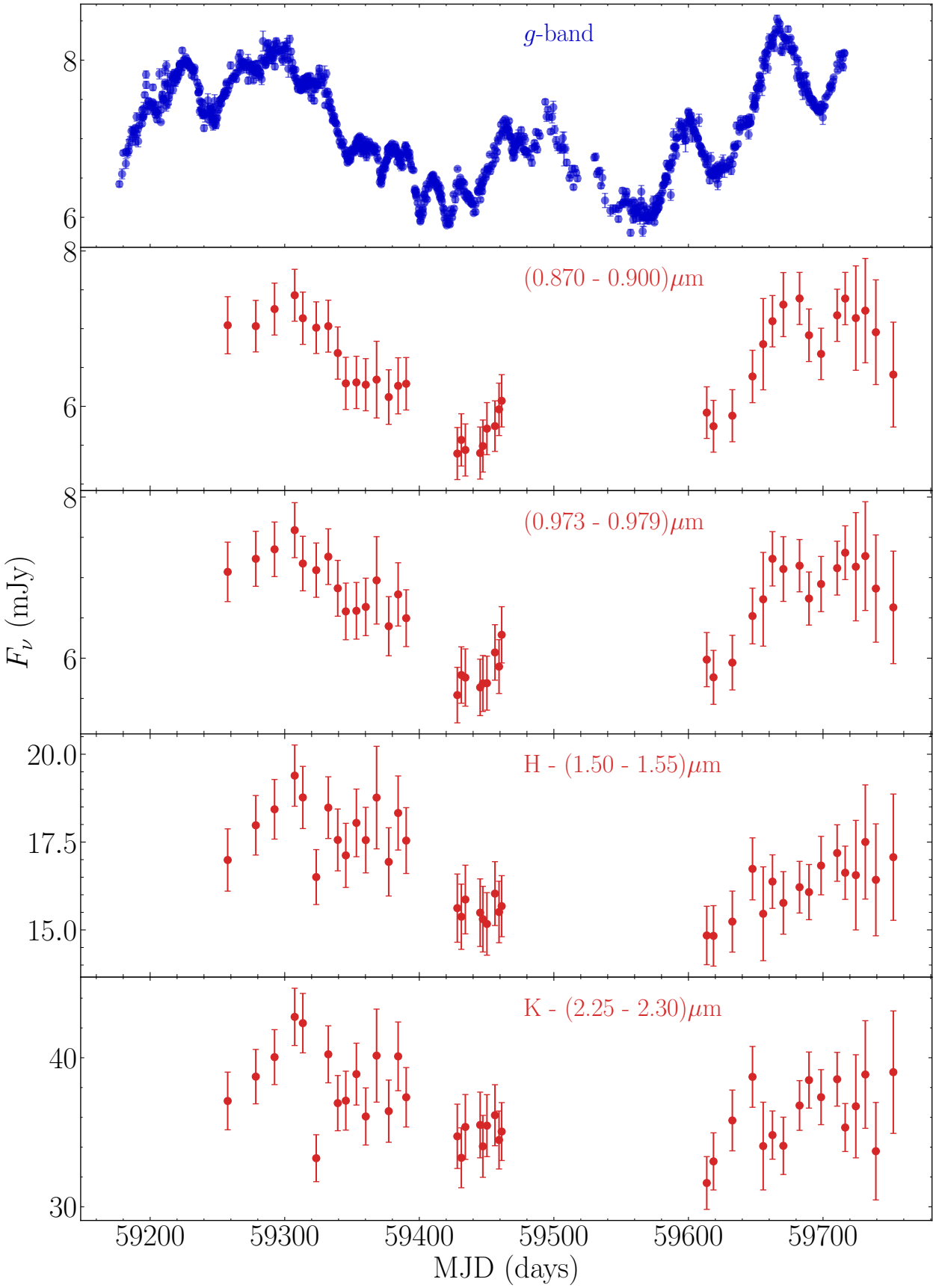

Figure 6.6: Top-bottom, g-band, (0.870-0.900) μm, (0.973-0.979) μm, H-band (1.50-1.55) μm and K-band (2.25-2.30) μm spectral lightcurves. Fluxes have been calibrated onto an absolute scale using MAPSPEC scale factors and then a secondary photometric correction.

Table 6.1: Fractional mean variance ($F_{\rm var}$) in each relevant waveband, unweighted mean flux ($\langle F \rangle$), squared variance (σ^2) and the mean square uncertainty (δ^2) calculated following Rodríguez-Pascual et al. (1997).

Waveband	$F_{\rm var}$	$\langle F \rangle$ (mJy)	σ^2 (mJy)	δ^2 (mJy)
g-band	0.0893±0.0002	7.118	0.405	0.001
z-band	0.0485±0.0003	14.934	0.544	0.018
(0.870 - 0.900)μm	0.07±0.01	6.49	0.42	0.21
(0.973 - 0.979)μm	0.06±0.01	6.62	0.35	0.18
H	0.03±0.01	16.70	1.44	1.17
K	0.03±0.01	36.71	6.72	5.39
Emission Line		(erg s^{-1}cm^{-2})	(erg s^{-1}cm^{-2})	(erg s^{-1}cm^{-2})
Pa β	0.099±0.007	1.195	0.017	0.003

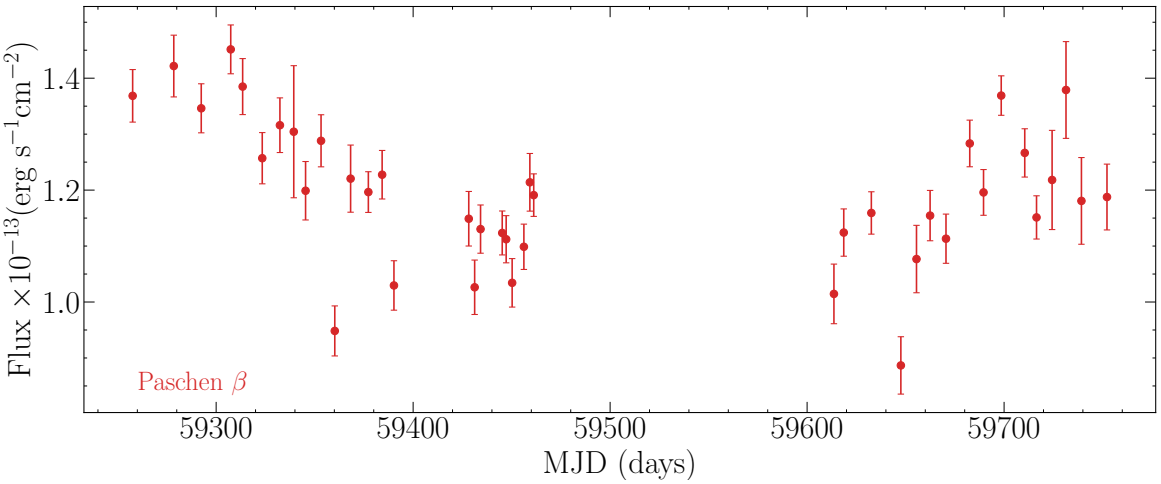

Figure 6.7: Pa β emission line lightcurve, constructed using the procedure described in Section 6.7.

6.7 Pa β Lightcurve

To construct lightcurves, we extract the enclosed Pa β fluxes. We fit a linear continuum to clean continuum windows redward and blueward of the Pa β broad line profile using the LMFIT module. We subtract this continuum level and integrate under the line profile. We determine the uncertainties on each flux measurement using Monte-Carlo methods, with a 2% systematic added in quadrature to account for uncertainties in the continuum selection.

The resultant lightcurve is shown in Figure 6.7, and the $F_{\rm var}$ value given in Table 6.1. We are sensitive to significant variability in the Pa β profile. There is a clear $\sim$ 30% decline in flux over the first half of our campaign, and a distinctive rise of similar magnitude in the second half.

6.8 Reverberation Results

6.8.1 Continuum and Hot Dust

We use JAVELIN, an open source reverberation mapping package (Zu et al., 2011). JAVELIN assumes that AGN variability is stochastic and can be modelled as a Damped Random Walk. JAVELIN uses a Bayesian framework alongside an MCMC sampler to create a statistical model of a 'driving' lightcurve. This 'driving' lightcurve is then smoothed, scaled and shifted when searching for lags in a 'response' lightcurve.

Here we use the g-band data as a 'driving' lightcurve, and the four spectroscopic near-IR lightcurves shown in Figure 6.6 as 'response' lightcurves we run JAVELIN with 10^5 iterations and a (0-100) day limit on the time lag search range in the form of strong preferential priors. This lag range was selected to prevent the lag distribution being skewed towards larger values for which the spectral data is not matched with data points in the 'driving' lightcurve. The results are displayed in Figure 6.8.

Figure 6.8 shows that the (0.870-0.900) μm and (0.973-0.979) μm bands clearly reverberate on timescales shorter than ∼10 lt-days showing clear lag peaks at $5.4^{+3.2}_{-2.6}$ and $6.4^{+21.4}_{-3.1}$ days respectively. Kara et al. (2021) calculate the g-band (4770Å) and z-band (8700Å) lag with respect to the *Swift/UVW2* (1928Å) lightcurve of $5.5^{+1.1}_{-1.2}$ days and $2.3^{+0.5}_{-0.5}$ days respectively. Therefore the lags we recover for (0.870-0.900) μm and (0.973-0.979) μm with respect to the g-band are consistent with the g-band to z-band lag presented in Kara et al. (2021). This gives added confidence in the accuracy of the scaling procedure. There is a small secondary peak at $\sim$ 27 days for (0.973-0.979) μm, however there is clearly no significantly reverberated signal at larger lags than this.

The longer wavelength H and K-bands both show much wider lag peaks, centered on higher lag values of $58.1^{+15.4}_{-10.8}$ and $45.4^{+30.2}_{-14.6}$ days. Koshida et al. (2014) recover a photometric K-band reverberated signal of $93.0^{+9.4}_{-8.5}$ days with respect to a photometric V-band lightcurve. We therefore underpredict the hot dust response with respect to the Koshida et al. (2014) result. This could be due to the fact that Koshida et al. (2014) perform a subtraction of the accretion disc, whereas we do not. The presence of some contribution of the accretion disc in the spectral lightcurves that we present could bias our lag measurements towards lower values. However given the low levels of variability in the H and K-bands before any spectral decomposition, we would not likely be able to detect any significantly reverberated signal after decomposition. It is also possible that a simple AD subtraction would not be suitable

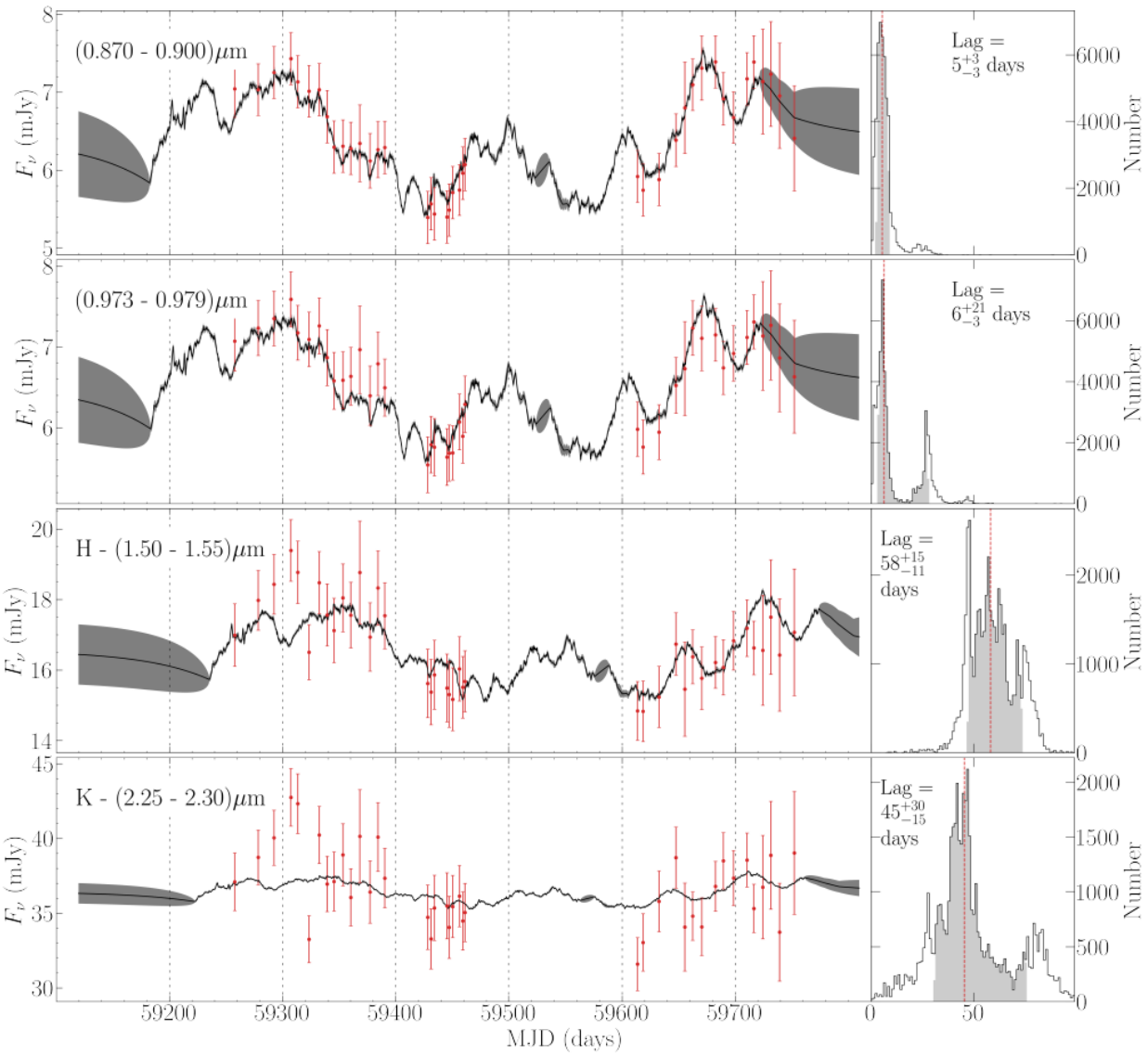

Figure 6.8: JAVELIN results for the four spectroscopic bands, (0.870-0.900) μm, (0.973-0.979) μm, (1.50-1.55) μm, (2.25-2.30) μm, displayed from top to bottom. Each run consists of 10^5 iterations, with the determined lags displayed on the right hand of each panel. The 40 red data points represent spectral data points, and the continuous grey curve is the smoothed, shifted and scaled JAVELIN model of the 'driving' g-band lightcurve displayed in Figure 5.2. Lag limits placed at (0-100) days.

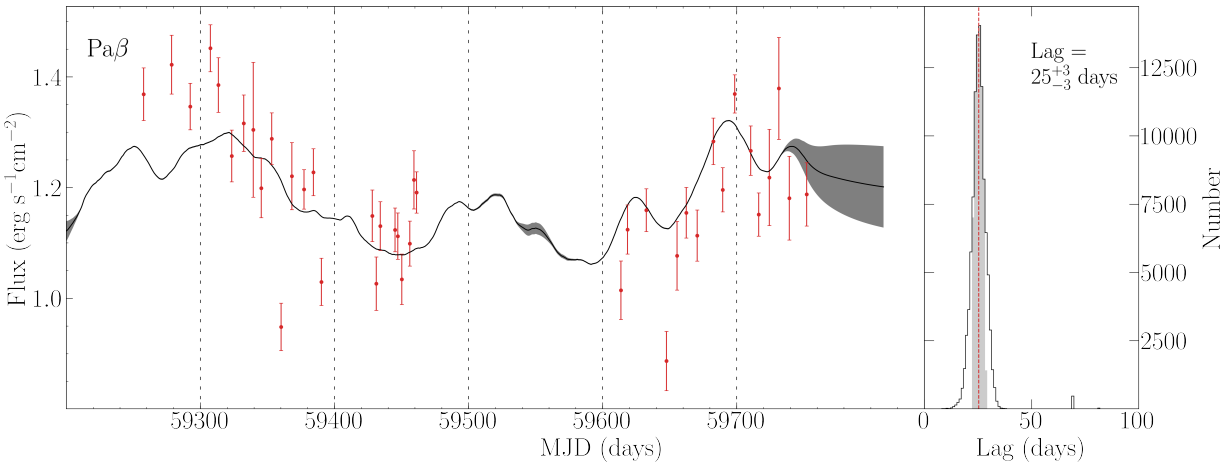

Figure 6.9: JAVELIN results for the Pa β broad emission line. Each run consists of 10^5 iterations, with the determined lags displayed on the right hand of each panel. The 40 red data points represent enclosed emission line fluxes, and the continuous grey curve is the smoothed, shifted and scaled JAVELIN model of the 'driving' g-band lightcurve displayed in Figure 6.2. Lag limits placed at (0-100) days.

as there could be some contribution from a diffuse BLR (Kara et al., 2021).

We perform a secondary reverberation analysis using PyCCF, the results of which are presented in Appendix D in Figure D.1. PyCCF gives consistent lag results with the JAVELIN analysis presented above for all four wavebands, giving particularly clear lag peaks for the (0.870-0.900) μm and (0.973-0.979) μm bands. This gives further confidence in the JAVELIN results that we present.

6.8.2 Pa β

We use JAVELIN to search for a reverberated signal between the g-band and Pa β lightcurves. We perform 10^5 iterations and search over a lag range of (0-100)days, as with the spectral lightcurves. These results are presented in Figure 6.9.

There is a very clear lag peak at $25.4^{+3.1}_{-3.1}$ days. The modelled 'response' lightcurves matches the broad emission trends of the spectral data well, and shows a fairly convincing agreement. Not only is the 'response' lightcurve shifted, but crucially it is smoothed by a value of 19.2 days. The posterior smoothing and lag distributions are plotted in Figure 6.10, the central smoothing width with uncertainties is $19.2^{+16.3}_{-16.4}$ days. For a lag result to be considered physically meaningful it is important for the smoothing width to be of a similar magnitude as the lag value, regardless of the geometry of the torus (Almeyda et al., 2017). Figure 6.10, shows that the resultant smoothing width is of similar magnitude to the lag peak, thus increasing the confidence in this result. In addition we search for a reverberated signal

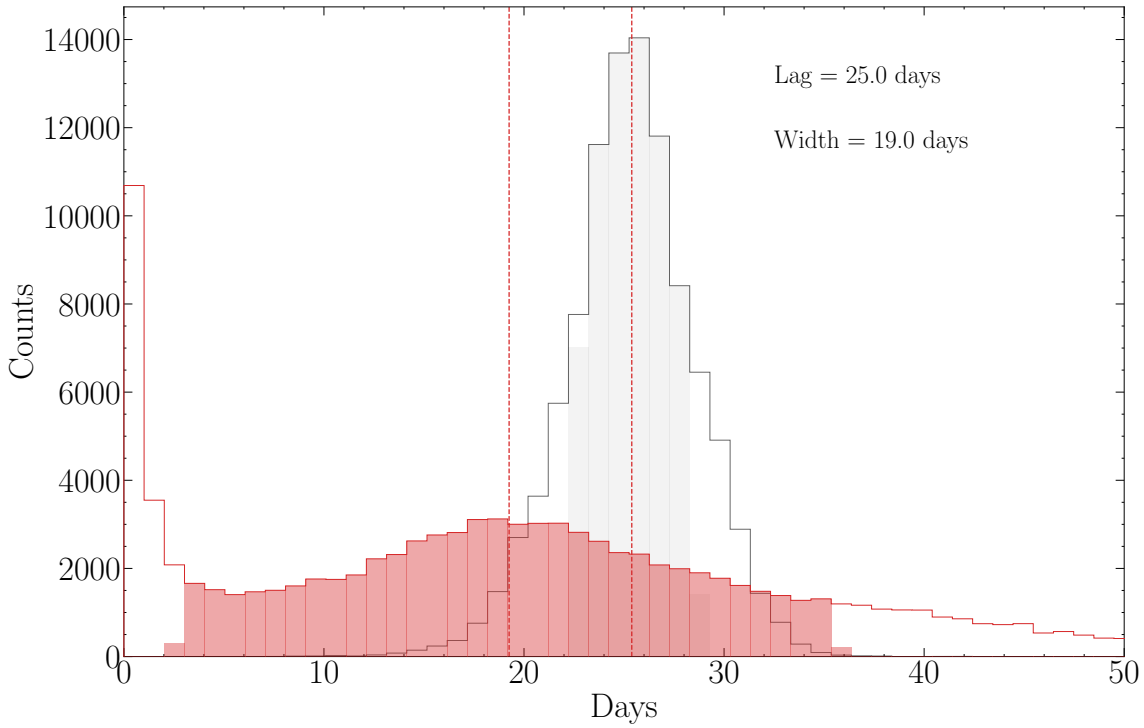

Figure 6.10: The posterior lag (grey) and width (red) distributions for the JAVELIN run on Pa β displayed in Figure 6.9. Uncertainties are shown in block grey and red for the lag and width distributions respectively.

over a wide baseline of (-300 - +300) days using PyCCF, the results of which are presented in Appendix D in Figure D.2. There is only one clear lag peak present, at $24.8^{9.4}_{9.3}$ days, a resulat that is consistent with the JAVELIN derived lag we present, therefore providing further confidence in this result.

Understanding the connection between the optical (Balmer) and the near-IR (Paschen) BLR is essential for building a general picture of the physics and geometry of the BLR. We find that the Pa β lag of $25.4^{+3.1}_{-3.1}$ days that we recover, is consistent with both the CCF and JAVELIN result for Hβ, (23.2 ± 1.6) days and $22.0^{+2.0}_{-1.4}$ days respectively, as reported in Kara et al. (2021). The Hβ lag reported in Kara et al. (2021) is measured against the 5100Å lightcurve which is slightly longer in wavelength than the g'-band lightcurve, central wavelength 4754Å against which our Pa β lag is measured. Therefore, as is the case, we would expect the Pa β lag to be slightly larger than the lag for Hβ. Our findings therefore indicate that the Balmer and Paschen BLRs are co-spatial.

6.9 Conclusion

We monitored the AGN Mrk 817 to document its infrared (hot dust) and optical (accretion disc) variations. Our IRTF spectra span 0.7 to 2.4 microns at 40 epochs between February 2021 and June 2022. The g'-band monitoring spans 2020 November 24 to 2022 May 16 with sub-day cadence during the IR campaign.

Our g-band lightcurve shows two main variability features, firstly a $\sim 25\%$ decrease beginning $\sim$59200 and a $\sim 25\%$ increase beginning at 59550. These features are plainly present in the (0.870-0.900)μm and (0.973-0.979)μm lightcurves which give clear reverberation results of $5.4^{+3.2}_{-2.6}$ and $6.4^{+21.4}_{-3.1}$ days respectively. These results are consistent with the continuum lags reported Kara et al. (2021). We report H and K-band lags of $58.1^{+15.4}_{-10.8}$ and $45.4^{+30.2}_{-14.6}$ days respectively. These results under-predict the photometric K-band lag of $93.0^{+9.4}_{-8.5}$ days reported by Koshida et al. (2014). This is most likely due to the remaining contribution from the accretion disc and diffuse BLR. However it is clear that we already do not sample high levels of variability in the H and K-bands and therefore do not perform a full spectral decomposition which would only increase the flux uncertainty and reduce the intrinsic variability further.

We construct mean, mean variance and $F_{\rm var}$ spectra, these confirm, as in the case of Mrk 509, that over the time frame of our observations the hot dust shows relatively little significant variability compared with the significant variability detected in the accretion disc at wavelengths $<1\mu$m.

The most confident and impotant result presented in this Chapter is the lag of $25.4^{+3.1}_{-3.1}$ days for Pa β with a $19.2^{+16.3}_{-16.4}$ days smoothing width, a result consistent with the Hβ lag measured for Mrk 817 by Kara et al. (2021). This implies that the Balmer and Paschen BLR are co-spatial. This is a novel result which could prove important as it is imperative that we are able to better understand the link between the Pashcen and Balmer broad line regions if we are to understand the BLR itself and to fully utilise the results from Gravity Collaboration et al. (2020) where the Paschen BLR of 3C 273 was spatially resolved.

Chapter 7

Summary and Future Works

There has been immense progress in Astronomy and Astrophysics over the last century. Less than a hundred years ago, there was still an active debate as to whether other galaxies even existed. Since then we have discovered millions of extra-galactic sources and began to unravel many of the profound mysteries pondered by humanity ever since the first person thought to look up and wonder. This story of progress is no more applicable than in the case of AGN which were discovered only ∼60 years prior to the writing of these words, and now are the focus of intensive research in hundreds of astronomy departments worldwide. For good cause too as there are still many questions that remain unanswered. I have highlighted some of these in this thesis and hope that the work presented here has contributed towards our ever increasing progress.

In this thesis, I focus on two aspects of AGN study:

- Near-IR spectroscopic RM aiming to constrain the spatial extent of the dusty torus and Paschen BLR.
- Broadband SED modelling and mass estimation of a medium sized sample of AGN spanning a wide range in mass and mass accretion rate, aiming to shed light on the nature of the disc.

7.1 Synopsis of Findings

7.1.1 Near-IR RM and the Torus

I present the results from two near-IR spectroscopic reverberation mapping campaigns firstly Mrk 509 and secondly Mrk 817 as part of the STORM2 campaign.

Due to an unfortunate observing run, we obtained only 17 spectra of Mrk 509 making the data-set a challenging one. Compounding this, we detect evidence of extension beyond the slit in the narrow emission line, leading to issue with the scaling procedure. We therefore develop a bootstrap scaling method, using MAPSPEC and the narrow [S III] as a first order correction before applying a secondary factor by scaling to the z_s-band photometry. Using this technique we are able to construct spectral lightcurves in four wavebands and emission line lightcurves for Pa ϵ and Pa β. Using a well sampled sub-day cadence g-band 'driving' lightcurve spanning 5 seasons, we search for reverberation signals with JAVELIN in the four spectral bands and both Pa ϵ and Pa β. We are able to place a lower limit on the H and K-band lags at > 40 days and place an upper limit of < 20 days in the (0.870-0.900) μm, (0.973-0.979) μm bands. These are consistent with literature values. We are able to perform a spectral decomposition and estimate the dust temperature of each spectrum. With these measurements and through sublimation temperature arguments we gain insights into the astro-chemistry of the inner edge of the dusty torus and conclude that the composition of the dust is likely to be large grains with a luminosity weighted radius $\langle R_{\rm d,lum} \rangle = (186 \pm 4)$ lt-days.

We repeat this approach for Mrk 817, with a more robust dataset, as part of the STORM2 campaign. We again scale to the photometry due to some inconsistencies between the MAPSPEC scaled spectral points and the z-band photometry. With this scaling method we construct four spectral lightcurves in the (0.870-0.900) μm, (0.973-0.979) μm, H and K-bands. Using JAVELIN we recover lags of 5 and 6 days in the (0.870-0.900) μm, (0.973-0.979) μm -bands respectively, a result that is consistent with contemporaneous photometric RM results in these bands. In the H and the K-bands we recover lags which underestimate K-band the hot dust response reported by Koshida et al. (2014), this could be due to some contribution from the AD biasing the lightcurve towards shorter lags. We do not perform a full spectral decomposition as we already sample relatively low levels of variability in both bands and this would only decrease with any AD contribution removed. We recover a Pa β lag of $25^{+3.1}_{-3.1}$ days with a smoothing width of $19.2^{+16.3}_{-16.4}$ days which is consistent with the contemporaneous Hβ lag reported by Kara et al. (2021). This result indicates that the Paschen and Balmer BLRs in Mrk 817 are co-spatial. It is of growing importance that we understand the connection between the Paschen and Balmer BLRs as the Paschen BLR is now resolvable with GRAVITY data in a few objects, so can be used to test single epoch mass estimates.

7.1.2 Broadband SED modelling and the nature of the disc

We assemble a sample of 696 type 1 AGN up to $z = 2.5$, all of which have an SDSS spectrum containing at least one broad emission line (H α, H β or Mg II) for mass estimation and an *XMM-Newton* X-ray spectrum containing at least 250 counts in addition to simultaneous optical/ultraviolet photometry from the *XMM* Optical Monitor. We name this sample the SDSS Optical/UV/X-ray or SOUX sample. The SOUX sample spans a wide range in luminosity, black hole mass and accretion rate. We compare single epoch mass estimates from the different lines and find broadly consistent mass estimates. Taking advantage of the wide range in mass and mass accretion rate spanned by our sample, we search for trends in $\alpha_{\rm ox}$ across this parameter space which we are then able to compare with the $\alpha_{\rm ox}$ predicted by the QSOSED model. We find that there are clear differences in $\alpha_{\rm ox}$ between both narrow/broad-line and radio-loud/quiet subsets of AGN. This indicates that the disc-corona coupling and kinetic output of an AGN are dependent on both the mass and mass accretion rate. We find that the narrow-line type 1 AGN in our sample are offset to lower masses and higher mass accretion rates than the typical broad-line quasars. Most of the AGN rest between $0.02 < L/L_{\rm Edd} < 2$ as defined by the QSOSED model.

This medium-sized sample enables us to stack spectra into bins of $M_{\rm BH}$ (0.5 dex) and L_{2500} (0.25 dex) to form average SEDs at a given mass and mass accretion rate. We compare these with the predictions of a new AGN SED model, QSOSED, which includes prescriptions for both a hot Comptonisation region replacing the inner disc, and a warm Compton region above some of the remaining disc to make the soft X-ray excess and UV downturn, and then an outer standard disc. This predicts the overall SED fairly well at intermediate masses, with $7.5 < \log(M_{\rm BH}/M_{\odot}) < 9.0$, but at higher masses the outer disc spectra in the model are far too cool to match the data. This is very surprising as the outer disc is the most physically justified component in the QSOSED model. We create better SDSS composite spectra for each grid point of the SOUX sample by stacking all AGN not just those with good X-ray data and use these to show that the mismatch is due to there being no significant change in spectral shape of the optical-UV continuum across several decades of $M_{\rm BH}$ at constant luminosity. We show for the first time that this cannot be matched by the standard approach for individual of simply allowing high black hole spin. The disc models apparently fit, but are not self-consistent as they do not include the General Relativistic effects for the emission to reach the observer. At such high spin, increased gravitational redshift compensates for almost all of the higher temperature emission from the smaller inner radii. We conclude that the observed

spectral shape of AGN across a wide range in $M_{\rm BH}$ and L_{3000} is remarkably constant, which does not match the predictions made by any current accretion flow models. Either the disc is completely covered by a warm Comptonisation layer whose properties change systematically with $L/L_{\rm Edd}$, or the accretion flow structure is fundamentally different to that of the standard disc models.

7.2 Future Works

7.2.1 Near-IR Spectroscopic Reverberation Mapping

The technique of near-IR spectroscopic reverberation mapping is a promising one, providing the existence of a robust well sampled data set and a highly variable source. It can allow for simultaneous RM along with temperature measurements providing insights into the spatial extent of the torus and its astro-chemistry. Applying this technique with a higher cadence on a highly variable source that has a well understood SED would potentially yield some very interesting results. This work could be carried out over a long baseline of observations, allowing for an investigation into the change in spatial extent of the torus during variability events. However the most important aspect of any future applications of this technique is the assurance of a reliable scaling technique. Detailed IFU imaging of the narrow [S III] region to ensure lack of extension beyond the slit width would be a good indicator.

7.2.2 The SOUX sample

We demonstrate that the shape of the Optical-UV continuum is static across a wide range in $M_{\rm BH}$, a result which could constitute part of an AGN crisis. Either the scaling relations used to estimate $M_{\rm BH}$ systematically overestimate $M_{\rm BH}$ at the largest masses, or AGN accretion does not occur via a disc. Whilst we favour the notion that AGN accretion is different than previously assumed, it is be important to test both of these hypothesis rigorously in future works.

The expanding work in spatially resolving the broad line regions of AGN is be extremely useful in testing the current $M_{\rm BH}$ scaling relations. This approach is only applicable to a very small number of targets currently, so complementary studies are also necessary. A systematic approach to checking the $M_{\rm BH}$ values in large catalogues such as (Rakshit et al., 2020), by including statistical tests on the effect of orientation among other factors would be extremely useful.

Assuming that the black hole masses are correct, it is imperative to reconcile the theoretical modelling of AGN SEDs with the trends exhibited by the data. This requires a paradigm shift away from the modelling of standard accretion discs, perhaps towards some form of Comptonisation or even an emission mechanism independent of $M_{\rm BH}$. Expanding the work on the SOUX sample to better understand its composition shall be an important step in this process. Obtaining high quality data on a much larger sample (thousands) of AGN across a larger range in mass and luminosity would allow a huge sample of SED composites to be constructed. This would enable detailed modelling experiments with rigorous atomic physics and full relativistic considerations.

7.2.3 The wider context

As the next generation of Space and ground based telescopes see first light and begin to impact the field of astronomy, the findings and conclusions of works such as this thesis can be built upon and tested ever more robustly.

The Large Synoptic Survey Telescope, now known as the Vera C. Rubin Observatory, shall one day map the entire sky visible from its perch on the Cerro Pachon ridge in north-central Chile every few nights over the course of ten years. This shall result in the largest and most detailed astronomical catalogue ever produced. The output of this telescope alone shall revolutionise the study of AGN, being particularly important to research focused on AGN variability. The utility of this facility does not only pertain to AGN research but extends to the mapping of our solar system and galaxy and even probing the nature of dark matter and dark energy.

Another ground based facility set to shake up the world of astronomical research is the Square Kilometer Array (SKA). The SKA is set to be the largest, most sophisticated radio telescope on the planet and shall be able to provide unprecedented resolution and therefore detailed radio images of AGN allowing for the probing of AGN radio morphologies in a more detailed fashion than ever before. This shall be potentially paradigm shifting and help to shed light on the origin of radio emission in AGN which are canonically termed 'radio-quiet' as well as potentially revealing the true morphology of many sources which, due to insufficient resolution, are wrongly classified as compact. As with the LSST the scope of this telescope is much larger than just AGN, the SKA is expected to obtain images of the very first stars and galaxies and to probe the large scale structure of the visible universe.

These exciting new facilities are not limited to our planet, but extend all the way out to

Earths Lagrange point 2 in fact. This is the home of the newly operational James Webb Space Telescope (JWST). This IR space telescope is set to revolutionise astronomy in almost every field from cosmology to the study of AGN. With unprecedented IR resolution JWST will be able to probe the IR emission regions of AGN like never before, perhaps even shedding light on the geometry and location of the torus and shall prove invaluable in studying reddened AGN with none of the observational constraints involved in observing across the IR bandpass from the Earth, some of which have been discussed in previous chapters of this thesis.

At the other end of the electromagnetic spectrum, observing in the X-rays, space telescopes such as ATHENA, IXPE and the proposed Lynx, will enable larger and more detailed studies than ever before on the X-ray emission from BHB systems and AGN. This shall enable the progression of studies such as that presented in Chapter 4, with detailed UV/X-ray SED modelling possible for ever growing samples across an ever widening parameter space. Better X-ray coverage of the AGN population shall be key in the expansion of the $\alpha_{\rm ox}$ work presented in Chapter 3, and shall enable the continuation of the work of Lusso et al. (2018) in the use of quasars as cosmological probes. (Lusso et al., 2018) already tentatively report a deviation from ΛCDM at high redshifts, potentially indicating an evolution in the behaviour of Dark Energy over time, it is data from these next generation X-ray telescopes that shall enable such findings to be robustly tested.

AGN research and astronomy and astrophysics in general is ever growing and moving in different an exciting directions. The era of big data is dawning and along with it shall come interesting and unexpected paradigm shifts in the way we view the Universe.

Appendix A

The SOUX AGN sample: Sample Definition

A.1 UV luminosities

In Section 3.4.1, for all possible AGN we estimate the UV luminosity at rest-frame 2500 from either the optical spectrum of a source, or its *XMM-Newton* OM or SDSS photometry. In Section 3.6 the UV luminosities were then used to calculate $\alpha_{\rm ox}$. This mixture of L_{2500} measures from three different sources is not ideal so here we assess the possible scale of the discrepancies that may be introduced, and their effect on $\alpha_{\rm ox}$ across the sample.

Mixing the spectroscopic and photometric measures of L_{2500} raises the following issues:

- Variability: only the OM photometry was recorded simultaneously with the X-ray data; the SDSS spectra and photometry were not.
- Aperture effects: the *XMM* OM photometric aperture (12–35″ diameter) is larger than the SDSS spectral fiber diameter (3″ SDSS spectrograph; 2″ BOSS spectrograph).
- Flux sampling: L_{2500} is arguably best determined from spectra, from which the flux is measured from a narrow (100 wide) window centred on rest-frame 2500 . In contrast the OM bands sample flux from a larger wavelength window: the effective widths of the U and UVW1 bands are $\approx$ 700 Å and therefore they possibly contain some broad line emission. Whilst we took photometric fluxes from a band containing 2500 the bands will not generally be centred on this wavelength. The Galactic dust dereddening we have applied to the UV fluxes will be more accurate for the spectroscopic measures

than the photometric ones (although Galactic extinction is very low in the SDSS survey area).

For 387 AGN in our sample we can determine the UV luminosity from both the SDSS spectrum and *XMM* OM photometry[1]; a comparison of the two measures is shown in Fig. A.1b. We have calculated the difference $\Delta \log(L_{2500}) = \log(L_{2500})^{\rm spec.} - \log(L_{2500})^{\rm phot.}$ and show this as a function of redshift, $\log(L_{2500})^{\rm spec.}$ and $\log(L_{2500})^{\rm phot.}$ in Figs.A.1d, f and h respectively. Fig. A.1b and Table A.1 show that on the whole there is no systematic difference between $\log(L_{2500})^{\rm spec.}$ and $\log(L_{2500})^{\rm phot.}$; the relationship is approximately one-to-one across 5 orders of magnitude in luminosity. There is substantial scatter, of ≈ 0.19 dex (i.e. around a 50 per cent difference in the luminosities). The scatter introduced will likely increase uncertainties on the relationships with L_{2500} , but since there is no strong systematic effect evident it appears unlikely that these differences would substantially alter the relationships. In our comparisons of $\alpha_{\rm ox}$ determined from the data and the QSOSED model (Section 3.6, Figs. 3.16, 3.18 and 3.19) we have in any case calculated averages in 0.5–1 dex luminosity bins. We do not see any evidence that $\Delta \log(L_{2500})$ changes with redshift, although as Fig A.1d shows, we do not have both spectroscopic and photometric L_{2500} UV measurements for AGN at $z \lesssim 0.5$. Aperture effects will be greatest for low-redshift AGN with resolved host galaxies.

In a sample of 51 $z < 0.4$ AGN, Jin et al. (2012a) noticed a discrepancy between flux levels of the SDSS spectra and *XMM* photometry for 17 sources, all with $z \lesssim 0.24$. The photometric points were above the extrapolation from the optical spectra, so it was proposed that the discrepancies were due to aperture effects, as described above. To obtain better agreement, Jin et al. (2012a) reanalysed the optical-UV images for affected sources and remeasured the photometry using a smaller aperture, excluding more of the extended host galaxy flux. The optical-UV photometric fluxes were generally reduced by ~ 50 per cent when using a smaller aperture (see Fig.1 of Jin et al. 2012a). A bespoke reanalysis of the OM photometry for AGN in the SOUX sample is not feasible. Instead we note that a 50 per cent change in L_{2500} corresponds to a ~ 0.1 change in $\alpha_{\rm ox}$ (see Fig. A.2). The SOUX sample consists of AGN drawn from NLS1 and quasar catalogues, so our AGN are generally UV-bright and not strongly contaminated by their host galaxy flux in the UV. Therefore, we do not anticipate that determining L_{2500} from photometry, or mixing photometric and spectroscopic measures, will have a very substantial effect on our results.

[1] We do not assess the SDSS photometry, since we used this to determine L_{2500} for only 26 sources in the sample.

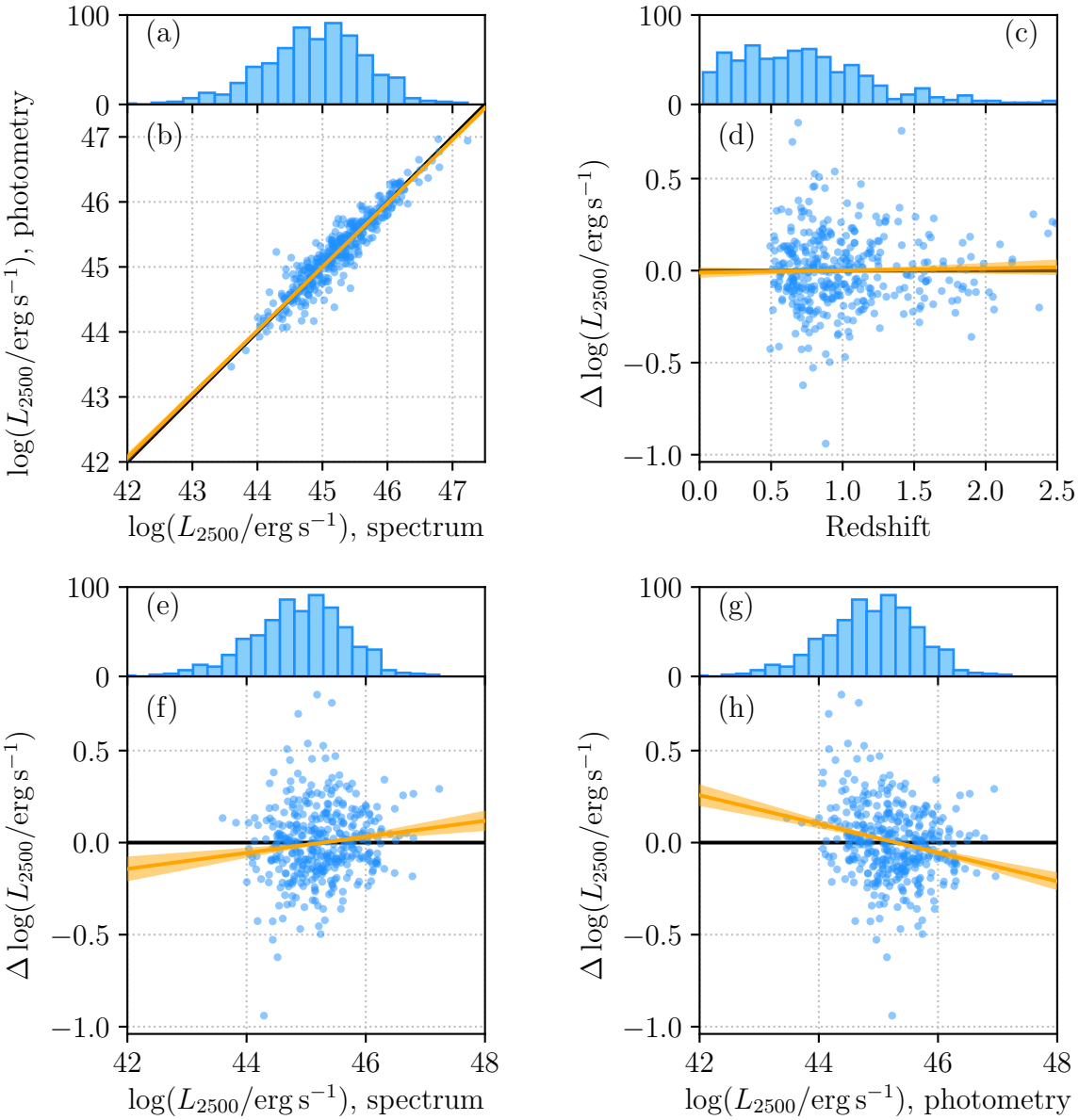

Figure A.1: A comparison of the UV luminosity at rest frame 2500 (L_{2500}) measured using either the SDSS optical spectrum or *XMM* OM photometry. Panels (b, d, f and h) show points for 387 AGN with both L_{2500} measurements. Panel (b) shows a direct comparison of the luminosities, with the solid black line designating the one-to-one relation. We define $\Delta \log(L_{2500}) = \log(L_{2500})^{\text{spec.}} - \log(L_{2500})^{\text{phot.}}$; panels (d, f and h) show $\Delta \log(L_{2500})$ as a function of redshift, $\log(L_{2500})^{\text{spec.}}$ and $\log(L_{2500})^{\text{phot.}}$, respectively. In panels (b, d, f and h) the best-fitting linear regression is indicated as a solid orange line and the 1σ confidence region is shaded pale orange. Panels (a, e and g) show the distribution of all 673 L_{2500} measurements used to determine α_{ox} (from either spectra or photometry, including sources which have L_{2500} determined from only one of these). Panel (c) shows the redshift distribution of all 696 AGN.

Table A.1: Linear regression parameters for the comparisons of L_{2500} measurements from spectra and photometry

Indep. var.	Dep. var.	Slope	Intercept	σ
$\log(L_{2500}^{\rm spec.})$	$\log\left(L_{2500}^{\rm phot.}\right)$	0.98 ± 0.02	1.0 ± 0.8	0.19
$\Delta\log(L_{2500})$	Redshift	0.01 ± 0.03	-0.01 ± 0.3	0.19
$\Delta\log(L_{2500})$	$\log(L_{2500}^{\rm spec.})$	0.04 ± 0.02	-2.0 ± 0.8	0.19
$\Delta\log(L_{2500})$	$\log\left(L_{2500}^{\rm phot.}\right)$	-0.08 ± 0.02	3.5 ± 0.8	0.18

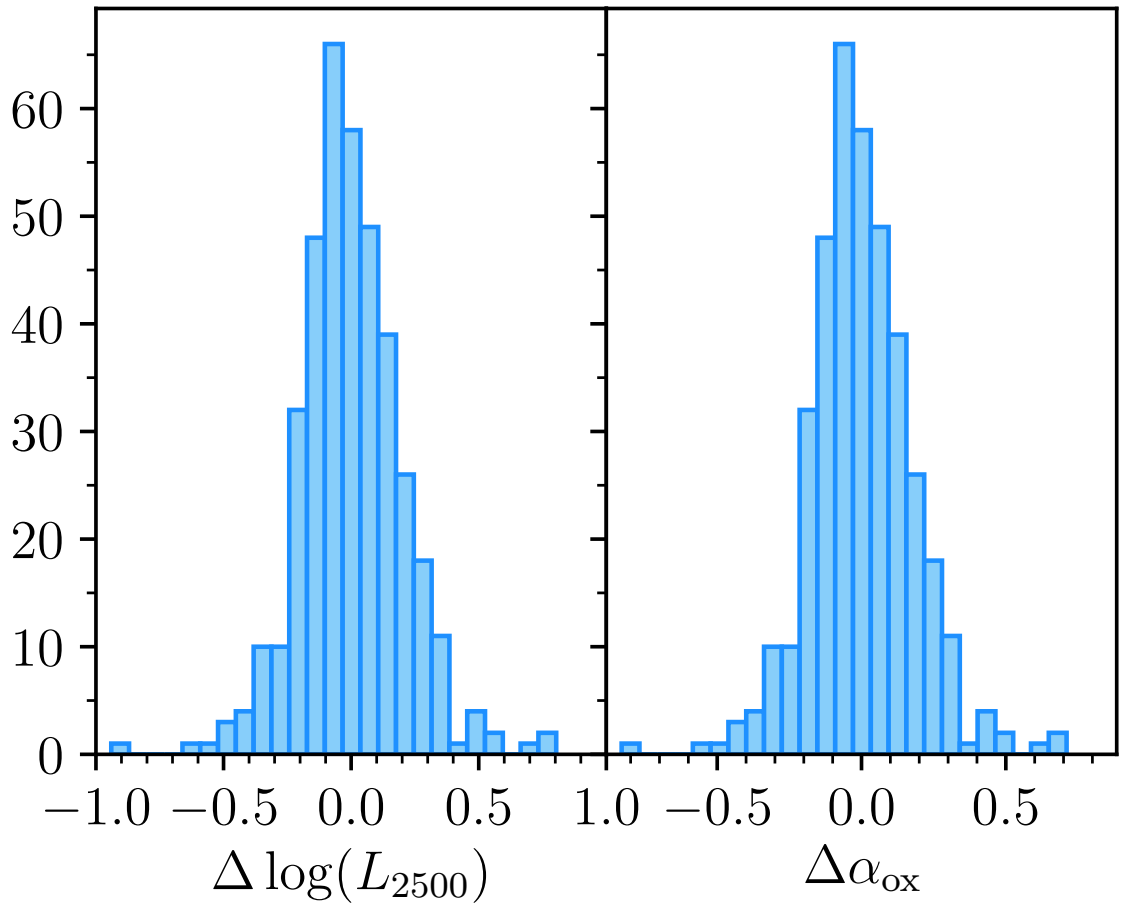

Figure A.2: *Left*: The difference in $\log(L_{2500})$ measurements made either from the optical spectrum or *XMM* photometry for 387 AGN with both measurements. We define $\Delta\log(L_{2500}) = \log(L_{2500})^{\rm spec.} - \log(L_{2500})^{\rm phot.}$. *Right*: The resultant change in $\alpha_{\rm ox}$ the change in $\log(L_{2500})$.

Table A.2: Quality flags for the fitted continuum

Bit	Condition
-	no quality flags raised
0	the luminosity or its uncertainty are zero, NaN or infinite
1	fractional uncertainty of the luminosity > 1.5
2	the continuum slope or its uncertainty are zero, NaN or infinite
3	the continuum slope has reached the lower or upper limit in the fit
4	the uncertainty on the continuum slope > 0.3
5	the continuum fit $\chi^2_\nu > 50$

A.2 AGN with multiple optical spectra

Whilst in this study we consider only one optical spectrum per AGN, many sources have been observed multiple times. Here we show quantitatively the number of spectra for the AGN in the SOUX sample. We use ASTROQUERY (Ginsburg et al. 2019) to search for all optical spectra of our 696 AGN in SDSS DR14. The query performs a $2''$-radius search around the input coordinates (taken from the Pâris et al. 2018 and R17 catalogues) for sources with spectroscopy. We checked that all spectra returned by the query relate to the same object (i.e. they have the same `ObjID`) and that the specific spectrum analysed in this work is one of those. No results were returned by ASTROQUERY for seven AGN; for these we manually checked for all spectra in the SDSS Science Archive Server[2]. We show the distribution of the number of spectra in Fig. A.3. Whilst 83 per cent of the AGN in our sample have only one optical spectrum, 17 have many (> 49). The wealth of optical data for these AGN may be explored further in future studies.

A.3 Quality flags

We assess the quality of measurements determined from the spectral fits by assigning a quality flag. If no issues are raised the quality flag value is 0 and the measurement may be considered to be reliable. The quality flags are integer numbers which are the sum $2^{\text{Bit }0}+2^{\text{Bit }1}...+2^{\text{Bit }n}$ for set bits. The bit numbers and the conditions for setting them are listed in Tables A.2 and A.3. For example, if the fractional uncertainty of the FWHM of a measured broad H β line is greater than 2 and the uncertainty on the velocity offset of the line is greater than 1000 km s^{-1} then bits 5 and 8 are set and the H β quality flag has the value $2^5 + 2^8 = 32 + 256 = 288$.

[2] `https://dr14.sdss.org/optical/spectrum/search`

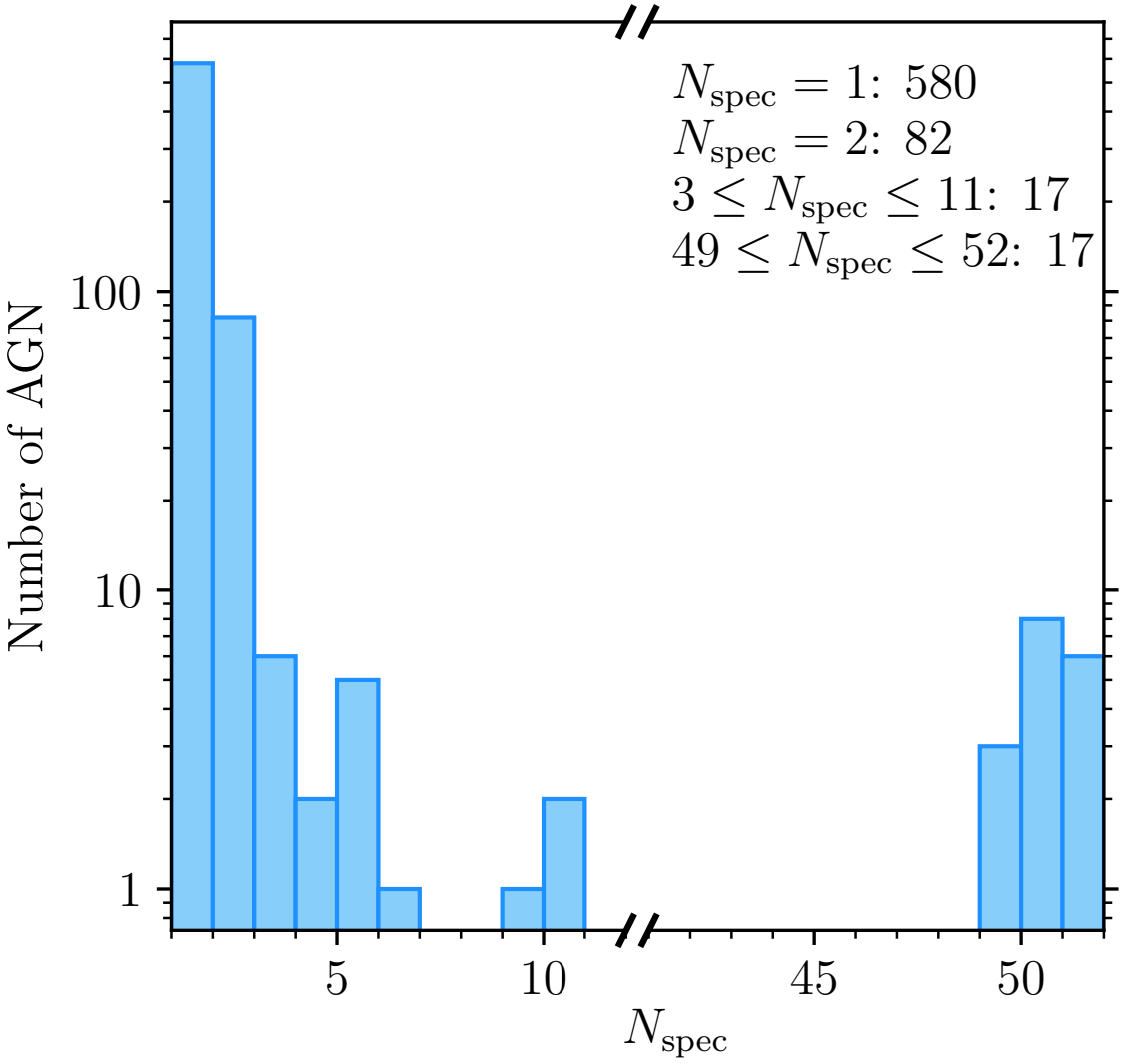

Figure A.3: The number of SDSS optical spectra ($N_{\rm spec}$) available for the 696 AGN in our sample, up to and including DR14. The majority of our sample (83 per cent) have only one spectrum, but a small number have ≈ 50 spectra.

Table A.3: Quality flags for the fitted emission lines

Bit	Condition
-	no quality flags raised
0	S/N in the continuum around the line < 3
1	the luminosity or its uncertainty are zero, NaN or infinite
2	fractional uncertainty of the luminosity > 1.5
3	the FWHM or its uncertainty are zero, NaN or infinite
4	the FWHM $\leqslant 910$ km s^{-1}
5	the fractional uncertainty of the FWHM > 2
6	the velocity offset or its uncertainty are zero, NaN or infinite
7	the velocity offset has reached the lower or upper limit in the fit
8	the uncertainty on the velocity offset > 1000 km s^{-1}

Appendix B

The SOUX sample: Optical/UV/X-ray SEDs and the Nature of the Disc

B.1 The SOUX sample changing mass at fixed luminosity

Figure B.1 shows the changing mass spectra for each luminosity bin covered by the SOUX sample for any L_{3000} bin spanning more than three mass bins. The broad line profiles clearly broaden with increasing mass.

B.2 The wider SDSS parameter space

Here we present composite spectra for the entire SDSS parameter space with z$\leq$0.8 or 2.0 $\leq$ z $\leq$ 2.15.

Figure B.2 shows the composite spectrum corresponding to each $M_{\rm BH}$, L_{3000} gridpoint coloured by the number of sources in each bin. The gridpoints populated by the SOUX AGN sample (enclosed in blue) inhabit the centre of and the most highly populated section of the parameter space. This is expected as the SOUX AGN sample is comprised of sources with high quality data *XMM-Newton* data which are relatively rare and therefore more likely to be found in the more highly populated bins.

Figure B.3 shows the same bins and composites as Figure B.2 but coloured with the resultant log($\dot{m}$) value from the fitting of QSOSED to a 500Å window of constant flux set at the central L_{3000} value for each respective bin.

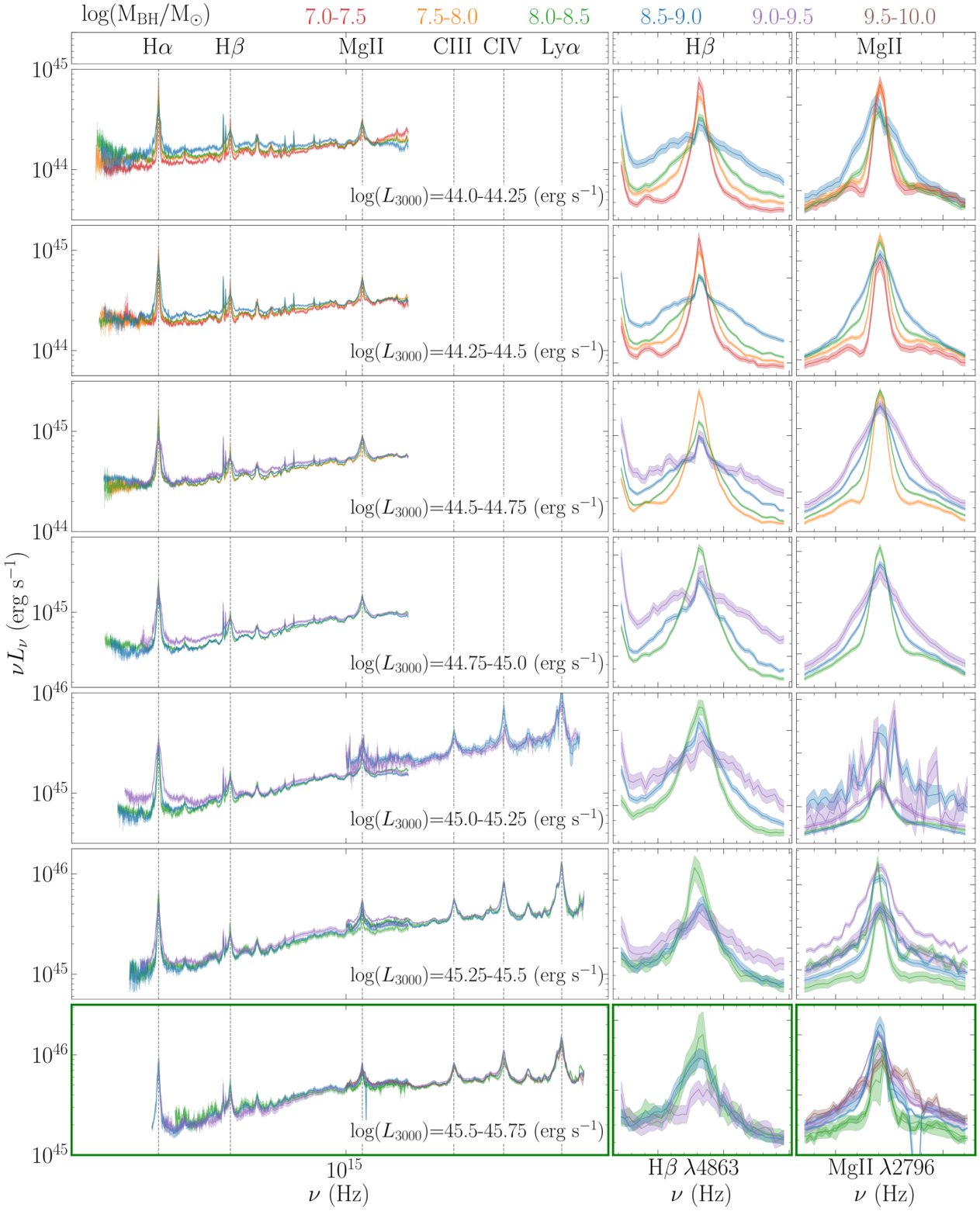

Figure B.1: All luminosity bins shown in Figure 4.9 that span more than 3 populated mass bins over plotted, showing the unchanging spectral shapes across the mass ranges. The right hand panels show the H β and Mg II emission line regions. A separate Mg II profile is for shown for both the low and the high z composite in any L_{3000} with overlap.

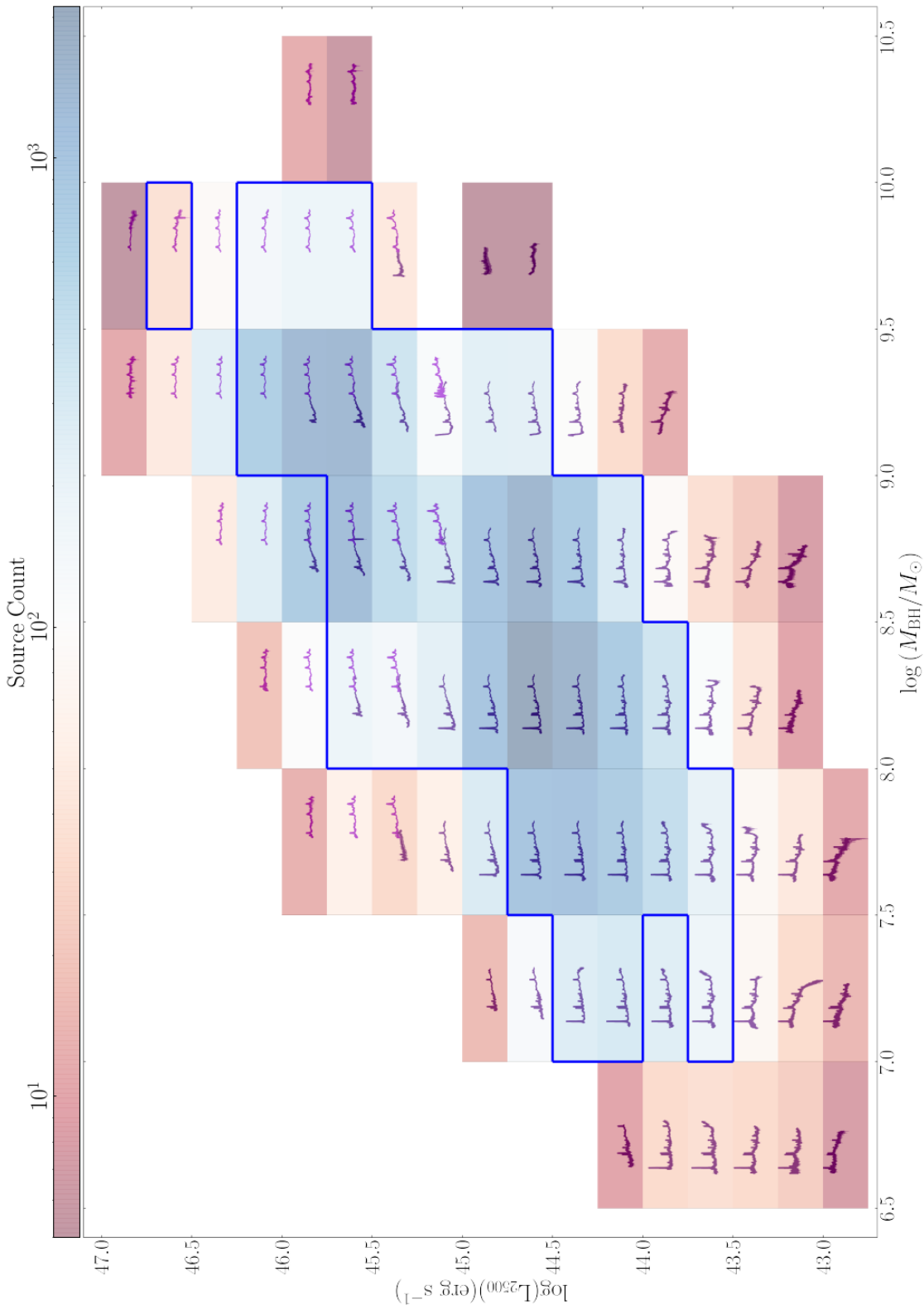

Figure B.2: SDSS composites for the entire parameter space $\log(M_{\mathrm{BH}}/M_\odot) = (6.5\text{-}10.5)$ and $\log(L_{3000}) = (42.75\text{-}47.0)$ ergs^{-1}. Each bin is coloured by the number of sources contained within.

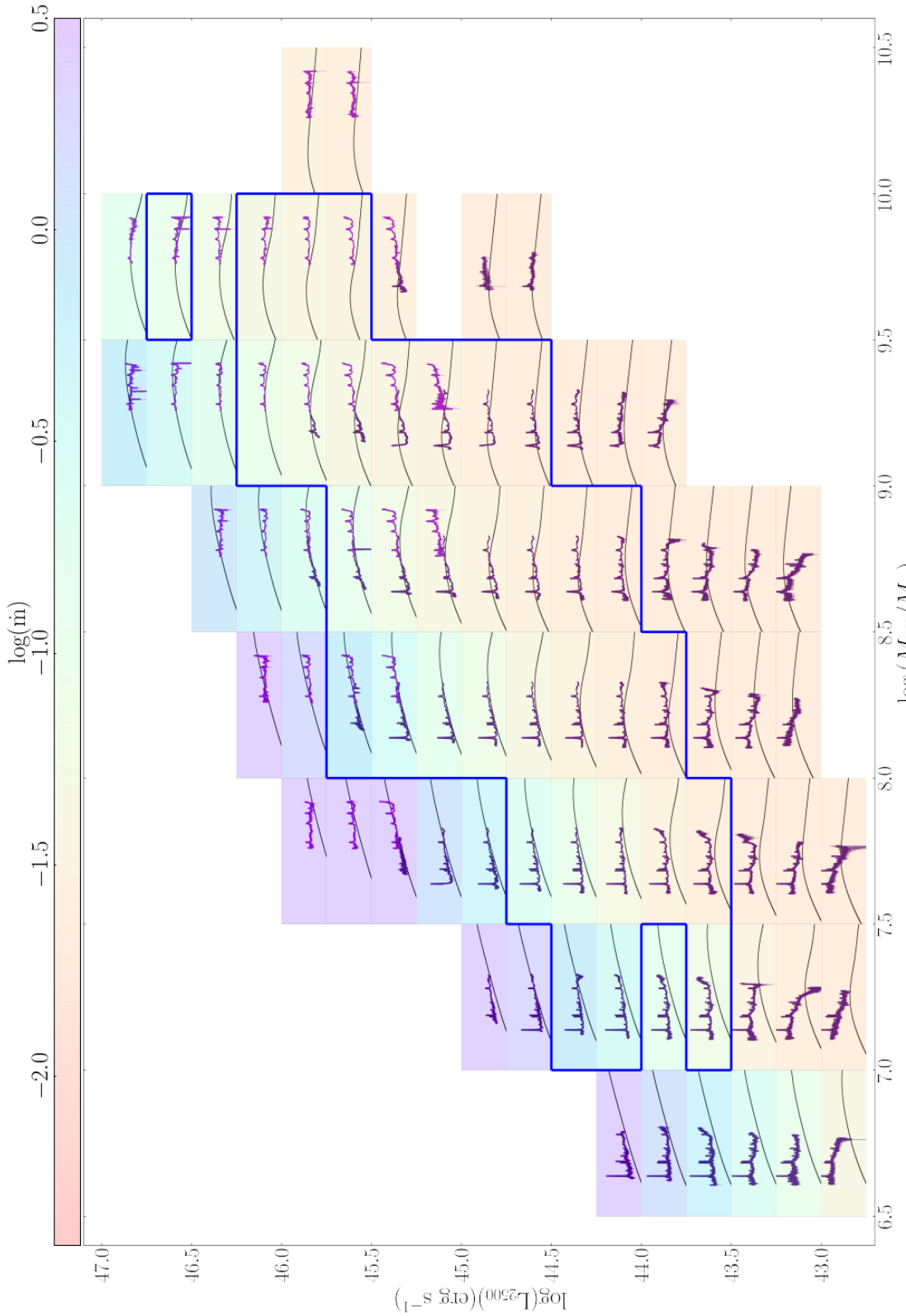

Figure B.3: SDSS composites for the entire parameter space $\log(M_{\rm BH}/M_\odot) = (6.5\text{-}10.5)$ and $\log(L_{3000}) = (42.75\text{-}47.0)$ ergs^{-1}. A QSOSED model has been fit to a 500Å wide bandpass of the central L_{3000} luminosity for each bin. Each bin is coloured with the resultant $\dot{m}$ values.

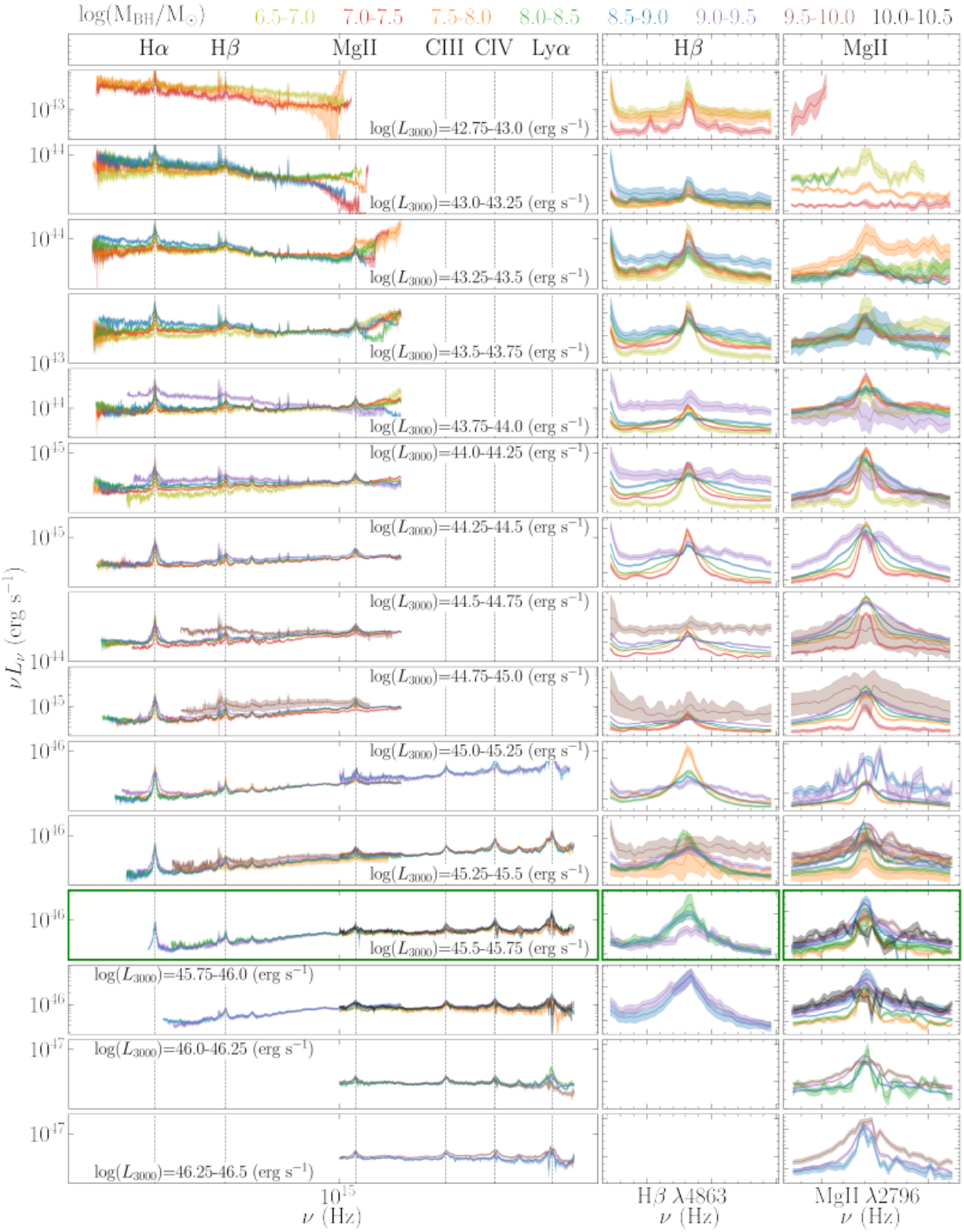

Figure B.4: All luminosity bins shown in Figure B.3 that span more than 3 populated mass bins over plotted, showing the unchanging spectral shapes across the mass ranges.

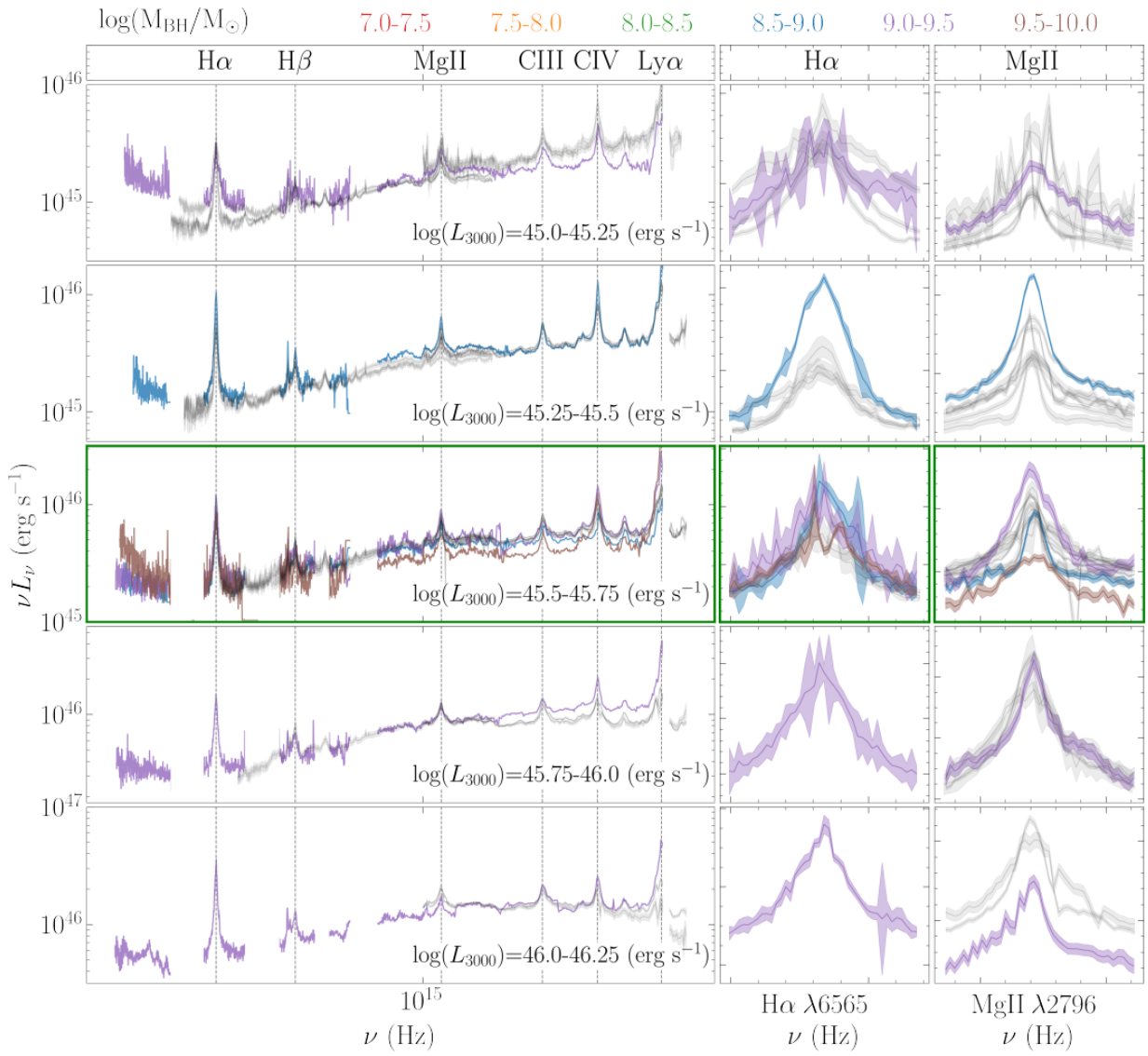

Figure B.5: Four L_{3000} bins with the SDSS data from Figure B.4 plotted in grey. Composite or individual X-shooter spectra taken from Capellupo et al. (2015) and Fawcett et al. (2022) are overplotted and coloured by mass.

B.3 X-Shooter Spectral Comparison

We take a sample of sources from Capellupo et al. (2015) and Fawcett et al. (2022) observed with X-shooter, to compare our results with a wider sample of AGN, not observed with SDSS. These spectra were obtained through private communication. The X-shooter spectra defined as the control sample in Fawcett et al. (2022), were separated into the L_3000 bins shown in Figure B.4. A geometric composite was created for any mass - luminosity bin containing more than one source and individual spectra considered for bins only containing one object, these spectra are shown in Figure B.5.

The spectral shape of the X-shooter data spanning 1.5 dex in mass, are well matched across

all luminosity bins to the shape of the SDSS composites displayed in Figure B.4 and plotted in grey in Figure B.5.

The X-shooter sample display no significant change in spectral shape with changing mass when compared to the SDSS spectra. The central panel of Figure B.5 representing the $log(L_{3000}) = 45.5-45.75(\mathrm{ergs}^{-1})$ bin, shows three X-shooter spectra spanning 1.5 dex in mass between 8.5-10.0 $\log(M_{\mathrm{BH}}/M_{\odot})$. These three spectra do not show any significant change in spectral shape.

B.4 Sources removed from SOUX AGN Sample

We performed a visual inspection on the *XMM-Newton* UV and X-ray data for each source and removed any object displaying a spectral shape indicative of intrinsic absorption. In Table B.1 we list all of the 54 sources that were removed before the fitting procedure.

Table B.1: Sources removed upon visual inspection due to evidence of intrinsic absorption in the *XMM-Newton* data. Columns: (1) SDSS-ID composed of PLATE, MJD and FIBREID from R20, (2) Right Ascension, (3) Declination, (4) Redshift as quoted by R20 and (5) Mass as calculated in K22 using the scaling relations from (Mejía-Restrepo et al., 2016).

SDSS-ID	Right Ascension	Declination	Redshift	Mass ($M_\odot$)
0391-51782-0360	00 22 09.9	+00 16 29.3	0.57	1.74E+08
0412-52258-0400	03 06 39.5	+00 03 43.1	0.11	2.60E+07
0422-51811-0224	01 07 12.0	+14 08 44.9	0.08	2.44E+06
0514-51994-0331	11 40 08.7	+03 07 11.4	0.08	3.19E+06
0539-52017-0171	15 01 48.8	+01 44 05.1	0.48	4.93E+07
0552-51992-0440	09 02 31.2	+52 07 49.6	0.13	8.76E+06
0616-52374-0442	15 36 41.6	+54 35 05.5	0.45	1.86E+08
0623-52051-0257	16 12 32.4	+51 24 01.3	0.36	7.89E+08
0629-52051-0200	16 42 51.3	+44 15 01.3	1.15	2.92E+09
0639-52146-0179	21 18 52.9	-07 32 27.5	0.26	2.57E+07
0709-52205-0040	03 07 07.4	-00 04 24.0	0.66	8.29E+07
1266-52709-0205	08 13 03.8	+25 42 11.0	2.02	1.64E+09
1365-53062-0378	11 17 06.3	+44 13 33.3	0.14	4.72E+08
1423-53167-0603	16 44 42.5	+26 19 13.2	0.14	1.43E+07
1604-53078-0566	11 13 54.6	+12 44 39.0	0.68	4.50E+07
1643-53143-0172	14 14 49.5	+36 12 40.2	0.18	3.89E+06
1744-53055-0630	10 07 26.0	+12 48 56.2	0.24	2.62E+09
1745-53061-0309	10 07 54.9	+12 18 42.1	0.76	2.28E+08
1795-54507-0457	13 09 46.9	+08 19 48.2	0.15	3.22E+08
1978-53473-0039	13 38 07.4	+28 05 09.8	1.09	5.38E+08
1994-53845-0267	12 53 17.5	+31 05 50.6	0.78	8.09E+08
2016-53799-0061	13 12 17.7	+35 15 21.0	0.18	8.98E+08
2020-53431-0051	12 42 10.6	+33 17 02.6	0.04	1.28E+07
2123-53793-0443	14 07 00.3	+28 27 14.6	0.08	5.04E+08
2132-53493-0349	14 23 15.5	+23 58 17.5	0.44	1.03E+09
2288-53699-0617	09 18 48.6	+21 17 17.0	0.15	1.96E+07
2289-53708-0130	09 25 54.7	+19 54 05.1	0.19	1.78E+09
2434-53826-0093	09 07 59.3	+13 51 35.0	0.45	2.49E+08
2501-54084-0384	11 27 36.9	+24 49 23.4	0.06	6.85E+06
2511-53882-0555	11 49 15.4	+23 12 33.6	1.06	9.22E+08
2763-54507-0581	14 53 01.4	+16 44 52.6	0.84	5.08E+09
2784-54529-0010	14 06 21.8	+22 23 46.5	0.1	1.38E+07
3936-55302-0730	15 50 14.8	+21 24 31.4	0.48	6.34E+07
4388-55536-0940	02 33 31.0	-05 45 50.9	0.49	1.47E+09
4558-55569-0008	10 34 16.2	+39 32 40.8	1.41	3.30E+09
4773-55648-0076	10 45 19.8	+04 19 51.2	0.8	7.41E+08
4832-55680-0150	12 19 32.3	+05 40 05.2	1.11	2.56E+09
5333-56001-0534	10 07 51.1	+12 45 26.4	0.21	2.73E+06
5365-55945-0380	11 12 38.0	+13 22 44.9	0.43	1.33E+08
5423-55958-0636	13 09 36.2	+08 28 15.3	0.78	1.19E+09
6370-56238-0807	02 29 29.4	+00 34 41.1	1.25	2.84E+08
6404-56330-0249	11 55 35.8	+23 27 23.1	0.14	6.55E+07
6404-56330-0294	11 55 04.1	+23 31 17.9	1.08	7.20E+08
6676-56389-0806	12 11 46.2	+50 31 04.6	1.16	7.76E+08
6723-56428-0410	15 45 55.5	+48 49 29.8	0.15	1.31E+07
6732-56370-0806	14 48 55.6	+48 30 14.4	0.68	2.20E+08
7235-56603-0274	02 29 28.4	-05 11 25.0	0.31	2.39E+07
7235-56603-0293	02 29 17.1	-05 29 01.4	0.19	9.13E+07
7299-56686-0377	09 47 04.5	+47 21 42.9	0.54	9.06E+07
7339-56772-0577	14 07 59.0	+53 47 59.7	0.17	9.50E+07
7419-56811-0620	12 34 13.3	+47 53 51.1	0.37	2.09E+08
7578-56956-0312	22 18 11.2	+19 30 39.8	1.28	3.99E+09
7841-56960-0809	01 58 46.0	-01 49 40.5	0.51	6.64E+08
7877-56898-0913	01 25 39.4	-01 13 08.4	1.08	2.98E+08

Appendix C

Near-IR Spectroscopic Reverberation mapping: Mrk 509

C.1 JAVELIN - Alias lags

Here we present JAVELIN results for the spectral and emission line lightcurves shown in Sections 5.5.1 and 5.5.2 respectively. These results are obtained via identical JAVELIN runs apart from the extension of the (0-100) day lag limit to (-800,+600) day lag limits. The results clearly show strong aliasing problems due to seasonal gaps in the optical lightcurve. With the exception of the (0.870-0.900) μm and (0.973-0.979) μm bands which show a very strong peak between 0 and 20 days, the other bands, and emission line lightcurves display an almost symmetric pattern of lag peaks indicating the dominance of alias lags due to seasonal gaps in the optical lightcurve.

C.2 Reverberation results using PyCCF

PyCCF is a python package which performs a cross correlation analysis on two unevenly sampled lightcurves using linear interpolation allowing the measurement of a peak and centroid of the cross correlation function (CCF) (Sun et al., 2018). We present the results of PyCCF analysis for all four spectral lightcurves and both emission line lightcurves here. This analysis is presented as a secondary check on the significance of the JAVELIN reverberation results shown in Chapter 5. The results largely agree with the JAVELIN analysis and are discussed in Chapter 5.

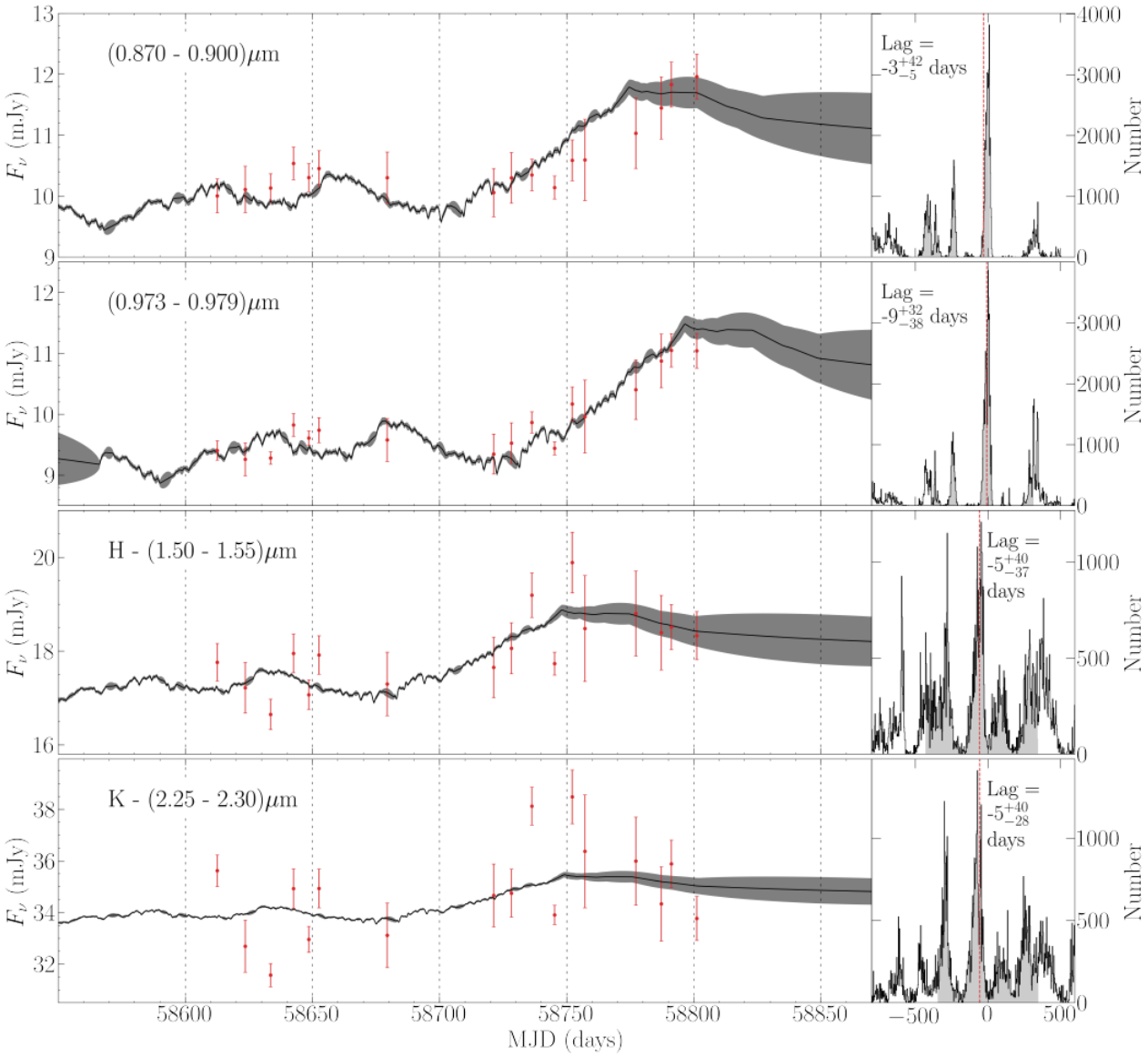

Figure C.1: JAVELIN results for the four spectroscopic bands, (0.870-0.900) μm, (0.973-0.979) μm, (1.50-1.55) μm, (2.25-2.30) μm, displayed from top to bottom. Each run consists of 10^5 iterations, with the determined lags displayed on the right hand of each panel. The 17 red data points represent spectral data points, and the continuous grey curve is the smoothed, shifted and scaled JAVELIN model of the 'driving' g-band lightcurve displayed in Figure 5.2. Lag limits placed at (-800 - +600) days.

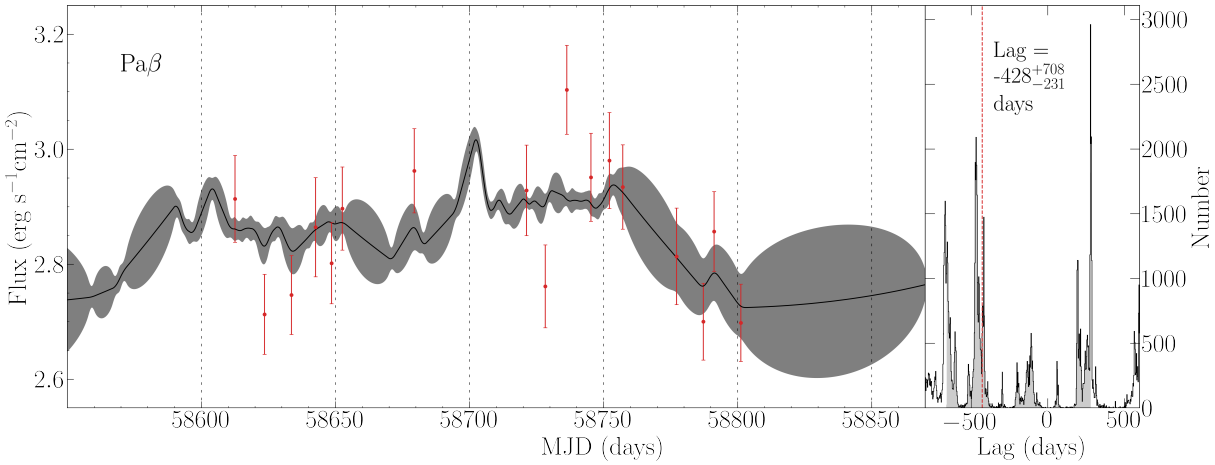

Figure C.2: JAVELIN results for the Pa β broad emission line. Each run consists of 10^5 iterations, with the determined lags displayed on the right hand of each panel. The 17 red data points represent enclosed emission line fluxes, and the continuous grey curve is the smoothed, shifted and scaled JAVELIN model of the 'driving' *g*-band lightcurve displayed in Figure 5.2. Lag limits placed at (-800 - +600) days.

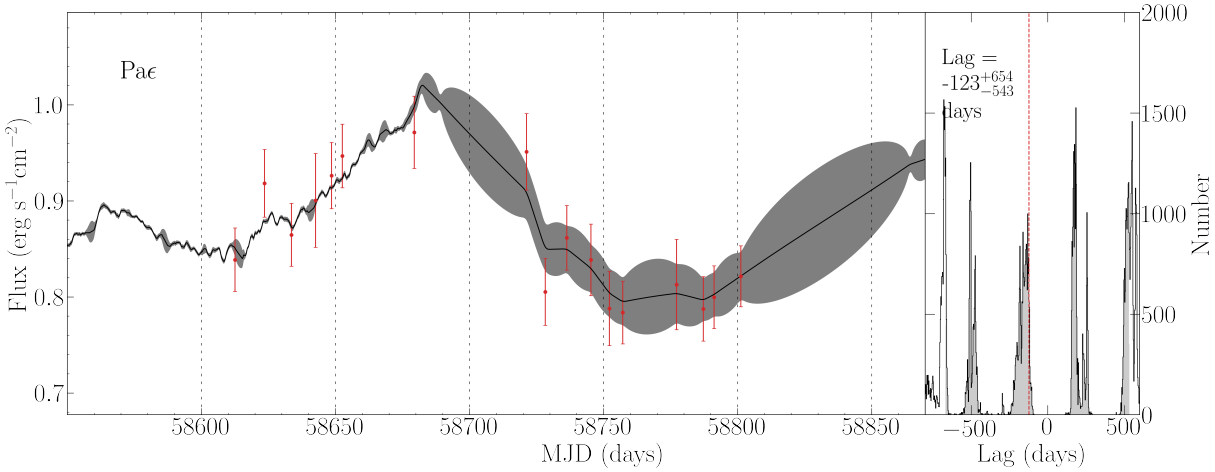

Figure C.3: JAVELIN results for the Pa ϵ broad emission line. Each run consists of 10^5 iterations, with the determined lags displayed on the right hand of each panel. The 17 red data points represent enclosed emission line fluxes, and the continuous grey curve is the smoothed, shifted and scaled JAVELIN model of the 'driving' *g*-band lightcurve displayed in Figure 5.2. Lag limits placed at (-800 - +600) days.

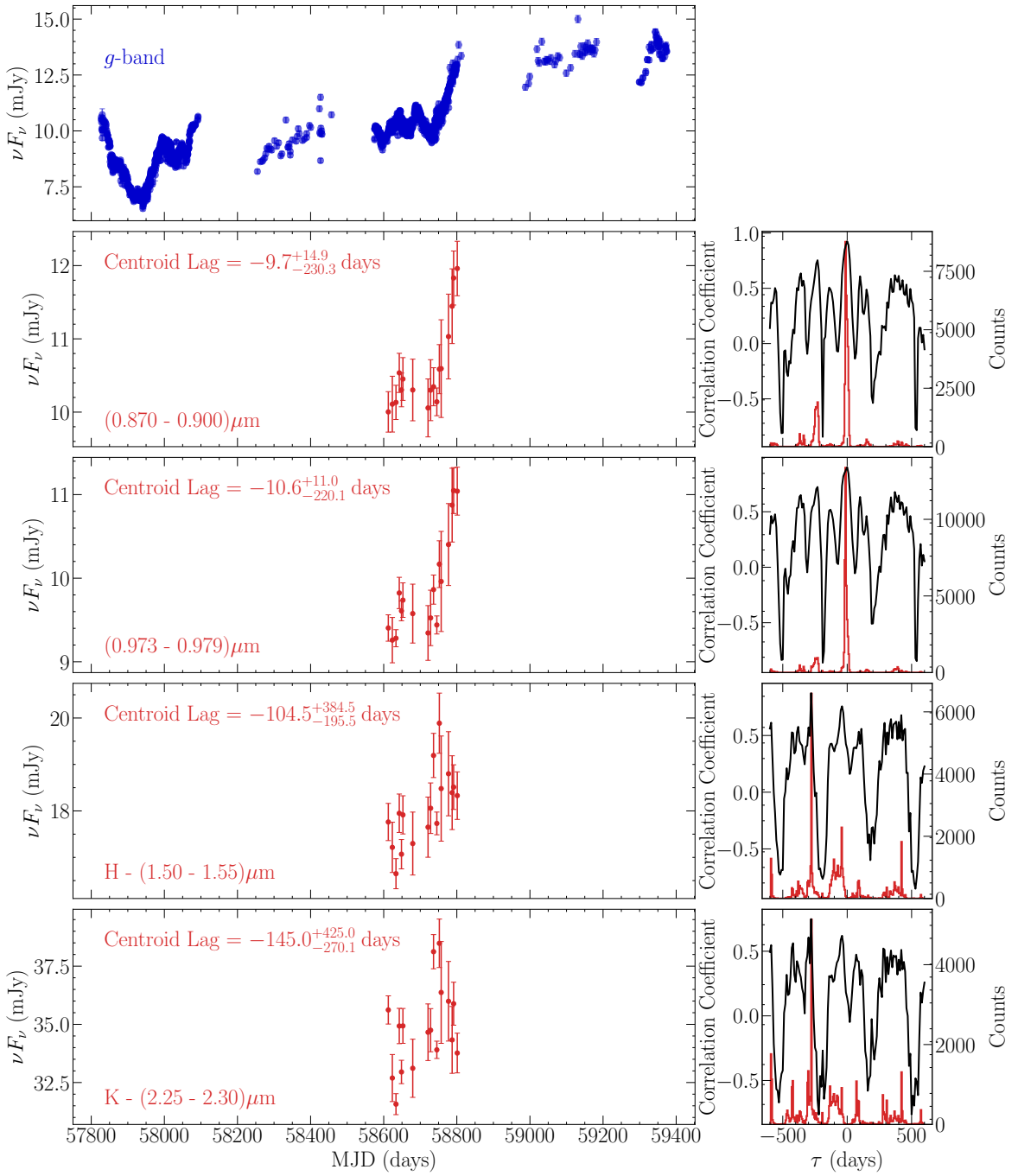

Figure C.4: PyCCF results for the four spectral lightcurves of Mrk 509. The lag limits were set at (-600 - +600) days. The top panel shows the photometric 'g'-band lightcurve. Each of the following four left-hand panels display an unperturbed spectral lightcurve, with the lag distribution and cross correlation coefficient presented in the right-hand panel. We performed 50000 iterations for each lightcurve in order to determine the errors.

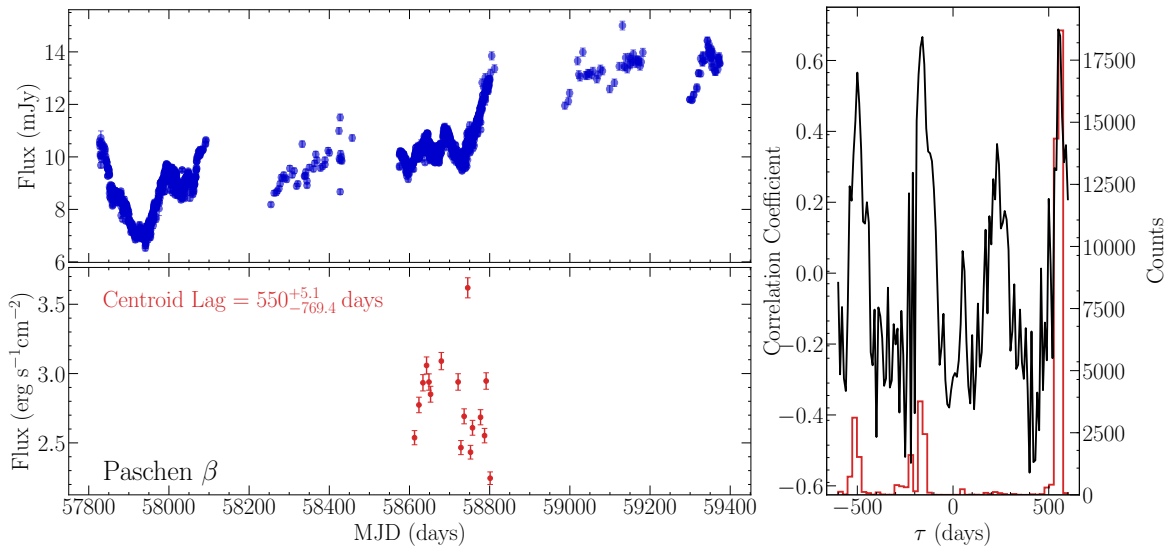

Figure C.5: PyCCF results for the Pa β lightcurve of Mrk 509. The lag limits were set at (-600 - +600) days. The top panel shows the photometric 'g'-band lightcurve. The following left-hand panel displays the unperturbed Pa β lightcurve, with the lag distribution and cross correlation coefficient presented in the right-hand panel. We performed 50000 iterations for each lightcurve in order to determine the errors.

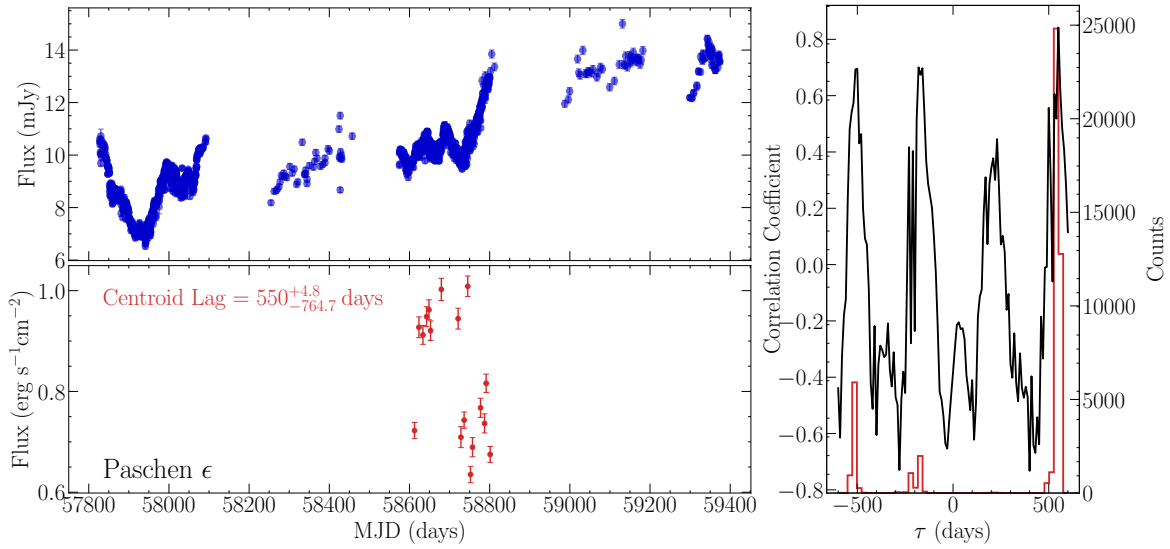

Figure C.6: PyCCF results for the Pa ϵ lightcurve of Mrk 509. The lag limits were set at (-600 - +600) days. The top panel shows the photometric 'g'-band lightcurve. The following left-hand panel displays the unperturbed Pa ϵ lightcurve, with the lag distribution and cross correlation coefficient presented in the right-hand panel. We performed 50000 iterations for each lightcurve in order to determine the errors.

Appendix D

Near-IR Spectroscopic Reverberation mapping: Mrk 817

D.1 Reverberation results using PyCCF

As in the case of Mrk 509 we present PyCCF analysis as a check of the reverberation results obtained with JAVELIN. The results of this analysis are displayed in Figures D.1 and D.2 and discussed in Chapter 6.

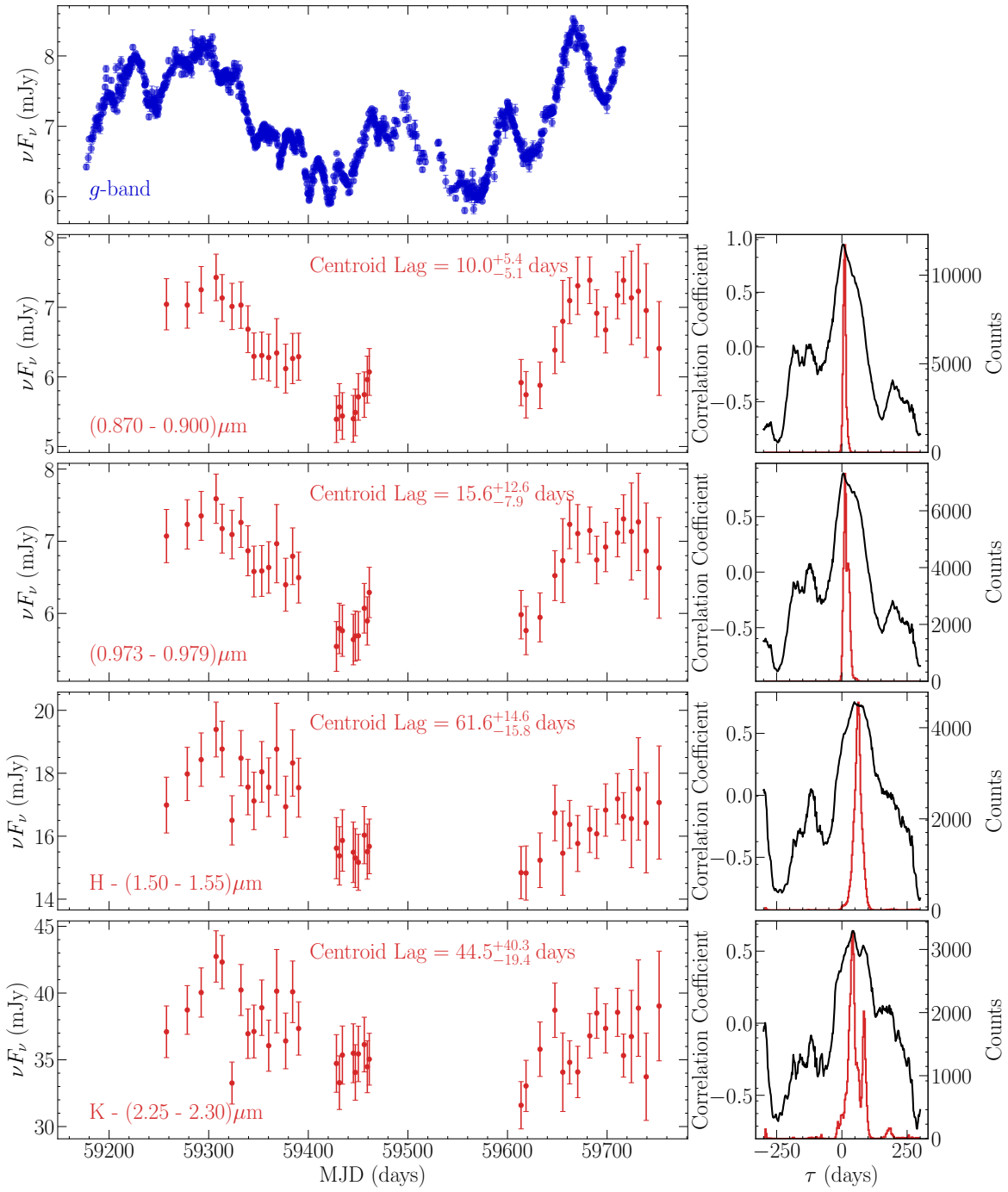

Figure D.1: PyCCF results for the four spectral lightcurves of Mrk 817. The lag limits were set at (-600 - +600) days. The top panel shows the photometric 'g'-band lightcurve. Each of the following four left-hand panels display an unperturbed spectral lightcurve, with the lag distribution and cross correlation coefficient presented in the right-hand panel. We performed 50000 iterations for each lightcurve in order to determine the errors.

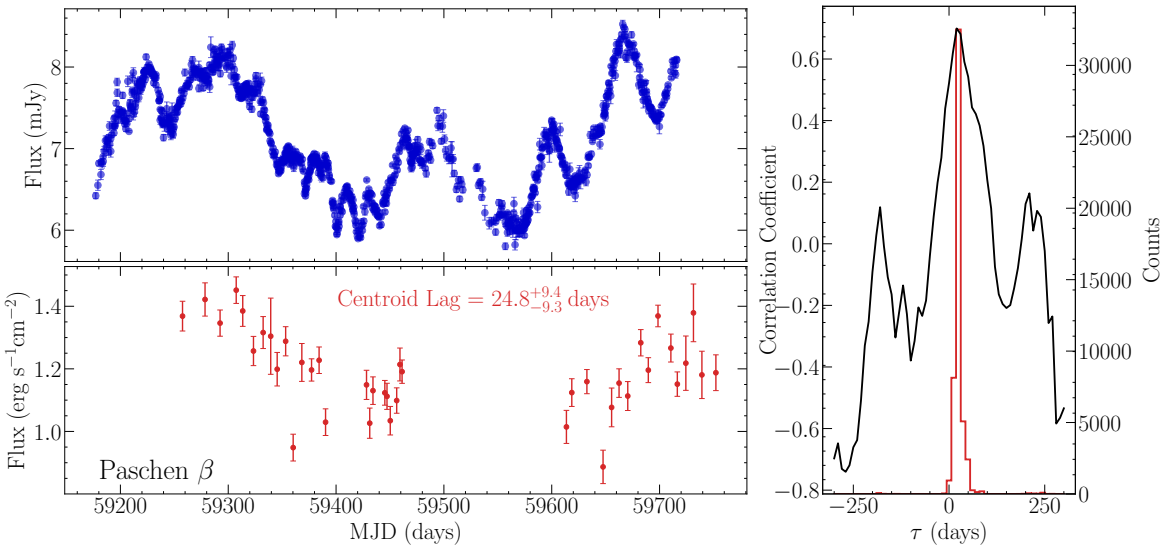

Figure D.2: PyCCF results for the Pa β lightcurve of Mrk 817. The lag limits were set at (-300 - +300) days. The top panel shows the photometric 'g'-band lightcurve. The following left-hand panel displays the unperturbed Pa β lightcurve, with the lag distribution and cross correlation coefficient presented in the right-hand panel. We performed 50000 iterations for each lightcurve in order to determine the errors.

Bibliography

Abazajian K. N., et al., 2009, ApJS, 182, 543

Abramowicz M. A., Czerny B., Lasota J. P., Szuszkiewicz E., 1988, ApJ, 332, 646

Alexander D. M., et al., 2003, AJ, 125, 383

Almeyda T., Robinson A., Richmond M., Vazquez B., Nikutta R., 2017, ApJ, 843, 3

Alonso-Herrero A., et al., 2011, ApJ, 736, 82

Andika I., 2016, PhD thesis, doi:10.13140/RG.2.2.23303.80809

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

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

Arcodia R., Ponti G., Merloni A., Nandra K., 2020, A&A, 638, A100

Astropy Collaboration et al., 2018, AJ, 156, 123

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

Beckmann V., Shrader C. R., 2012, Active Galactic Nuclei

Bellm E. C., 2014, The Zwicky Transient Facility, doi:10.48550/ARXIV.1410.8185, `https://arxiv.org/abs/1410.8185`

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

Bentz M. C., Peterson B. M., Pogge R. W., Vestergaard M., Onken C. A., 2006, ApJ, 644, 133

Bentz M. C., Cackett E. M., Crenshaw D. M., Horne K., Street R., Ou-Yang B., 2016, ApJ, 830, 136

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

Blanton M. R., et al., 2017, AJ, 154, 28

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

Bradley L., et al., 2020, astropy/photutils: 1.0.0, Zenodo, doi:10.5281/zenodo.4044744

Brewer B., Bartok-Partay L., Csányi G., 2009, Statistics and Computing, 21

Brosch N., Polishook D., Shporer A., Kaspi S., Berwald A., Manulis I., 2008, Ap&SS, 314, 163

Brown T. M., et al., 2013, PASP, 125, 1031

Burtscher L., et al., 2013, A&A, 558, A149

Buttiglione S., Capetti A., Celotti A., Axon D. J., Chiaberge M., Macchetto F. D., Sparks W. B., 2009, A&A, 495, 1033

Cackett E. M., Horne K., Winkler H., 2007, Monthly Notices of the Royal Astronomical Society, 380, 669

Cackett E. M., Chiang C.-Y., McHardy I., Edelson R., Goad M. R., Horne K., Korista K. T., 2018, The Astrophysical Journal, 857, 53

Cackett E. M., Bentz M. C., Kara E., 2021, iScience, 24, 102557

Calderone G., Nicastro L., Ghisellini G., Dotti M., Sbarrato T., Shankar F., Colpi M., 2017, MNRAS, 472, 4051

Capellupo D. M., Netzer H., Lira P., Trakhtenbrot B., Mejía-Restrepo J., 2015, MNRAS, 446, 3427

Capellupo D. M., Netzer H., Lira P., Trakhtenbrot B., Mejía-Restrepo J., 2016, MNRAS, 460, 212

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245

Chakravorty S., Kembhavi A. K., Elvis M., Ferland G., 2009, Monthly Notices of the Royal Astronomical Society, 393, 83

Clavel J., Wamsteker W., Glass I. S., 1989, ApJ, 337, 236

Clavel J., et al., 1992, ApJ, 393, 113

Cohen M. H., Ogle P. M., Tran H. D., Goodrich R. W., Miller J. S., 1999, AJ, 118, 1963

Coleman M. S. B., Kotko I., Blaes O., Lasota J. P., Hirose S., 2016, MNRAS, 462, 3710

Coleman M. S. B., Blaes O., Hirose S., Hauschildt P. H., 2018, ApJ, 857, 52

Collinson J. S., Ward M. J., Done C., Landt H., Elvis M., McDowell J. C., 2015, MNRAS, 449, 2174

Collinson J. S., Ward M. J., Landt H., Done C., Elvis M., McDowell J. C., 2017, MNRAS, 465, 358

Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693

Cushing M. C., Vacca W. D., Rayner J. T., 2004, PASP, 116, 362

Czerny B., Hryniewicz K., 2011, A&A, 525, L8

Davis S. W., Woo J.-H., Blaes O. M., 2007, ApJ, 668, 682

De Rosa G., et al., 2015, ApJ, 806, 128

Den Herder J. W., et al., 2001, A&A, 365, L7

Dickey J. M., Lockman F. J., 1990, ARA&A, 28, 215

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

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

Dong X., Wang T., Wang J., Yuan W., Zhou H., Dai H., Zhang K., 2008, MNRAS, 383, 581

Draine B. T., 2016, ApJ, 831, 109

Eckert D., Gaspari M., Gastaldello F., Brun A. M. C. L., O'Sullivan E., 2021, Universe, 7, 142

Edelson R., et al., 2015, The Astrophysical Journal, 806, 129

Edelson R., et al., 2019, The Astrophysical Journal, 870, 123

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

Event Horizon Telescope Collaboration et al., 2019, ApJ, 875, L1

Event Horizon Telescope Collaboration et al., 2022, ApJ, 930, L12

Fabian A. C., 2012, ARA&A, 50, 455

Fanidakis N., Baugh C. M., Benson A. J., Bower R. G., Cole S., Done C., Frenk C. S., 2011, MNRAS, 410, 53

Fausnaugh M. M., 2017, PASP, 129, 024007

Fawcett V. A., Alexander D. M., Rosario D. J., Klindt L., Lusso E., Morabito L. K., Rivera G. C., 2022, Monthly Notices of the Royal Astronomical Society, 513, 1254

Fender R. P., Belloni T. M., Gallo E., 2004, MNRAS, 355, 1105

Fernández-Ontiveros J. A., Muñoz-Darias T., 2021, MNRAS, 504, 5726

Fitzpatrick E. L., 1999, PASP, 111, 63

Foschini L., et al., 2012, A&A, 548, A106

García-Burillo S., et al., 2016, ApJ, 823, L12

Gaskell C. M., 2017, MNRAS, 467, 226

Gaspari M., Tombesi F., Cappi M., 2020, Nature Astronomy, 4, 10

Ghez A. M., et al., 2008, ApJ, 689, 1044

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

Gierliński M., Zdziarski A. A., Poutanen J., Coppi P. S., Ebisawa K., Johnson W. N., 1999, MNRAS, 309, 496

Ginsburg A., et al., 2019, AJ, 157, 98

Gitti M., Brighenti F., McNamara B. R., 2012, Advances in Astronomy, 2012, 1

Glass I. S., 1992, MNRAS, 256, 23P

Glass I. S., 2004, MNRAS, 350, 1049

Goad M. R., Korista K. T., Ruff A. J., 2012, MNRAS, 426, 3086

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

Gravity Collaboration et al., 2017, A&A, 602, A94

Gravity Collaboration et al., 2018, Nature, 563, 657

Gravity Collaboration et al., 2020, A&A, 635, A92

Greene J. E., Ho L. C., 2005, ApJ, 630, 122

Greene J. E., et al., 2010, ApJ, 723, 409

Greiner J., et al., 2008, PASP, 120, 405

Grier C. J., et al., 2013, ApJ, 773, 90

Griffin A. J., Lacey C. G., Gonzalez-Perez V., Lagos C. d. P., Baugh C. M., Fanidakis N., 2019, MNRAS, 487, 198

Guo H., Shen Y., Wang S., 2018, PyQSOFit: Python code to fit the spectrum of quasars (ascl:1809.008)

Guo H., Liu X., Shen Y., Loeb A., Monroe T., Prochaska J. X., 2019, MNRAS, 482, 3288

Guo H., et al., 2020, ApJ, 888, 58

Gültekin K., et al., 2009, The Astrophysical Journal, 698, 198

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

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

Hernán-Caballero A., et al., 2015, ApJ, 803, 109

Hernán-Caballero A., Hatziminaoglou E., Alonso-Herrero A., Mateos S., 2017, Frontiers in Astronomy and Space Sciences, 4

Hernández Santisteban J. V., et al., 2020, MNRAS, 498, 5399

Hernández Santisteban J. V., et al., 2020, Monthly Notices of the Royal Astronomical Society, 498, 5399

Hönig S. F., Kishimoto M., 2017, ApJ, 838, L20

Horne K., 1986, PASP, 98, 609

Hubeny I., Blaes O., Krolik J. H., Agol E., 2001, ApJ, 559, 680

Huško F., Lacey C. G., Schaye J., Schaller M., Nobels F. S. J., 2022, MNRAS, 516, 3750

Hönig S. F., 2019, The Astrophysical Journal, 884, 171

Ichikawa K., et al., 2015, ApJ, 803, 57

Imanishi M., Nakanishi K., Izumi T., Wada K., 2018, ApJ, 853, L25

Jansen F., et al., 2001, A&A, 365, L1

Jiang Y.-F., Blaes O., 2020, The Astrophysical Journal, 900, 25

Jiang Y.-F., Cantiello M., Bildsten L., Quataert E., Blaes O., 2015, ApJ, 813, 74

Jiang Y.-F., Davis S. W., Stone J. M., 2016a, ApJ, 827, 10

Jiang Y.-F., Guillochon J., Loeb A., 2016b, ApJ, 830, 125

Jin C., Ward M., Done C., Gelbord J., 2012a, MNRAS, 420, 1825

Jin C., Ward M., Done C., 2012b, MNRAS, 422, 3268

Jin C., Ward M., Done C., 2012c, MNRAS, 425, 907

Kaastra J. S., et al., 2014, Science, 345, 64

Kalberla P. M. W., Burton W. B., Hartmann D., Arnal E. M., Bajaja E., Morras R., Pöppel W. G. L., 2005, A&A, 440, 775

Kara E., et al., 2021, ApJ, 922, 151

Kawaguchi T., Mori M., 2010, ApJ, 724, L183

Kawaguchi T., Mori M., 2011, ApJ, 737, 105

Kellermann K. I., Sramek R., Schmidt M., Shaffer D. B., Green R., 1989, AJ, 98, 1195

Kelly B. C., 2007, ApJ, 665, 1489

Kishimoto M., Hönig S. F., Beckert T., Weigelt G., 2007, A&A, 476, 713

Kishimoto M., Hönig S. F., Antonucci R., Kotani T., Barvainis R., Tristram K. R. W., Weigelt G., 2009, A&A, 507, L57

Kishimoto M., Hönig S. F., Antonucci R., Barvainis R., Kotani T., Tristram K. R. W., Weigelt G., Levin K., 2011, A&A, 527, A121

Kishimoto M., et al., 2013, ApJ, 775, L36

Komossa S., Voges W., Xu D., Mathur S., Adorf H.-M., Lemson G., Duschl W. J., Grupe D., 2006, AJ, 132, 531

Körding E. G., Jester S., Fender R., 2006, MNRAS, 372, 1366

Korista K. T., Goad M. R., 2001, The Astrophysical Journal, 553, 695

Korista K. T., Goad M. R., 2019, Monthly Notices of the Royal Astronomical Society, 489, 5284

Koshida S., et al., 2014, ApJ, 788, 159

Krawczyk C. M., Richards G. T., Mehta S. S., Vogeley M. S., Gallagher S. C., Leighly K. M., Ross N. P., Schneider D. P., 2013, ApJS, 206, 4

Kriss G. A., et al., 2011, A&A, 534, A41

Kriss G. A., et al., 2019, ApJ, 881, 153

Krumpe M., et al., 2017, A&A, 607, L9

Kubota A., Done C., 2018, MNRAS, 480, 1247

Kubota A., Done C., 2019, MNRAS, 489, 524

Kuraszkiewicz J. K., et al., 2003, ApJ, 590, 128

Kynoch D., Ward M. J., Lawrence A., Bruce A. G., Landt H., MacLeod C. L., 2019a, MNRAS, 485, 2573

Kynoch D., et al., 2019b, MNRAS, 487, 181

La Mura G., Popović L. Č., Ciroi S., Rafanelli P., Ilić D., 2007, ApJ, 671, 104

Landt H., Bentz M. C., Ward M. J., Elvis M., Peterson B. M., Korista K. T., Karovska M., 2008, ApJS, 174, 282

Landt H., Bentz M. C., Peterson B. M., Elvis M., Ward M. J., Korista K. T., Karovska M., 2011a, MNRAS, 413, L106

Landt H., Elvis M., Ward M. J., Bentz M. C., Korista K. T., Karovska M., 2011b, MNRAS, 414, 218

Landt H., et al., 2017, MNRAS, 464, 2565

Landt H., et al., 2019, MNRAS, 489, 1572

Laor A., Davis S. W., 2014, Monthly Notices of the Royal Astronomical Society, 438, 3024

Laor A., Draine B. T., 1993, ApJ, 402, 441

Laor A., Netzer H., 1989, MNRAS, 238, 897

Lawrence A., 1987, PASP, 99, 309

Lawrence A., 2018, Nature Astronomy, 2, 102

Li Y.-R., Wang J.-M., Hu C., Du P., Bai J.-M., 2014, The Astrophysical Journal, 786, L6

Lu Y., Wang T., Zhou H., Wu J., 2007, AJ, 133, 1615

Lu K.-X., Zhao Y., Bai J.-M., Fan X.-L., 2019, MNRAS, 483, 1722

Lumsden S. L., Heisler C. A., Bailey J. A., Hough J. H., Young S., 2001, MNRAS, 327, 459

Lusso E., Risaliti G., 2016, ApJ, 819, 154

Lusso E., Risaliti G., 2017, A&A, 602, A79

Lusso E., et al., 2010, A&A, 512, A34

Lusso E., Fumagalli M., Rafelski M., Neeleman M., Prochaska J. X., Hennawi J. F., O'Meara J. M., Theuns T., 2018, ApJ, 860, 41

Lyke B. W., et al., 2020, ApJS, 250, 8

Magdziarz P., Blaes O. M., Zdziarski A. A., Johnson W. N., Smith D. A., 1998, MNRAS, 301, 179

Malzac J., Belloni T., Spruit H. C., Kanbach G., 2003, A&A, 407, 335

Mason K. O., et al., 2001, A&A, 365, L36

McCarthy I. G., et al., 2010, MNRAS, 406, 822

McCully C., Volgenau N. H., Harbeck D.-R., Lister T. A., Saunders E. S., Turner M. L., Siiverd R. J., Bowman M., 2018, in Guzman J. C., Ibsen J., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 10707, Software and Cyberinfrastructure for Astronomy V. p. 107070K (arXiv:1811.04163), doi:10.1117/12.2314340

McHardy I. M., Koerding E., Knigge C., Uttley P., Fender R. P., 2006, Nature, 444, 730

McHardy I. M., et al., 2014, Monthly Notices of the Royal Astronomical Society, 444, 1469

McHardy I. M., et al., 2016, Astronomische Nachrichten, 337, 500

McHardy I. M., et al., 2018, MNRAS, 480, 2881

McNamara B. R., Nulsen P. E. J., 2007, ARA&A, 45, 117

Mejía-Restrepo J. E., Trakhtenbrot B., Lira P., Netzer H., Capellupo D. M., 2016, MNRAS, 460, 187

Merloni A., Heinz S., di Matteo T., 2003, MNRAS, 345, 1057

Miller L., Peacock J. A., Mead A. R. G., 1990, MNRAS, 244, 207

Minezaki T., Yoshii Y., Kobayashi Y., Enya K., Suganuma M., Tomita H., Aoki T., Peterson B. A., 2004, ApJ, 600, L35

Miniutti G., et al., 2014, MNRAS, 437, 1776

Moravec E., Svoboda J., Borkar A., Boorman P., Kynoch D., Panessa F., Mingo B., Guainazzi M., 2022, A&A, 662, A28

Morganson E., et al., 2012, AJ, 143, 142

Morganson E., et al., 2014, The Astrophysical Journal, 784, 92

Mortlock D. J., et al., 2011, Nature, 474, 616

Murray N., Chiang J., Grossman S. A., Voit G. M., 1995, ApJ, 451, 498

Narayan R., Yi I., 1995, ApJ, 452, 710

Nardini E., et al., 2015, Science, 347, 860

Nelson B. O., 1996, ApJ, 465, L87+

Nenkova M., Sirocky M. M., Ivezić Z., Elitzur M., 2008, ApJ, 685, 147

Netzer H., 2015, ARA&A, 53, 365

Netzer H., 2019, MNRAS, 488, 5185

Netzer H., 2021, MNRAS, 509, 2637

Noda H., Done C., 2018, MNRAS, 480, 3898

Novikov I. D., Thorne K. S., 1973, in Black Holes (Les Astres Occlus). pp 343–450

Oknyanskij V. L., Horne K., 2001, in Peterson B. M., Pogge R. W., Polidan R. S., eds, Astronomical Society of the Pacific Conference Series Vol. 224, Probing the Physics of Active Galactic Nuclei. p. 149

Oppenheimer B. D., Babul A., Bahé Y., Butsky I. S., McCarthy I. G., 2021, Universe, 7

Osterbrock D. E., Ferland G. J., 2006, Astrophysics of gaseous nebulae and active galactic nuclei

Osterbrock D. E., Pogge R. W., 1985, ApJ, 297, 166

Osterbrock D. E., Shuder J. M., 1982, ApJS, 49, 149

Packham C., Radomski J. T., Roche P. F., Aitken D. K., Perlman E., Alonso-Herrero A., Colina L., Telesco C. M., 2005, ApJ, 618, L17

Page M. J., et al., 2012, MNRAS, 426, 903

Pâris I., et al., 2018, A&A, 613, A51

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266

Peterson B. M., Wandel A., 1999, ApJL, 521, 95

Peterson B. M., et al., 2004, ApJ, 613, 682

Peterson B. M., et al., 2013, ApJ, 779, 109

Petrucci P. O., Ursini F., De Rosa A., Bianchi S., Cappi M., Matt G., Dadina M., Malzac J., 2018, A&A, 611, A59

Porquet D., Reeves J. N., O'Brien P., Brinkmann W., 2004, A&A, 422, 85

Pott J.-U., Malkan M. A., Elitzur M., Ghez A. M., Herbst T. M., Schödel R., Woillez J., 2010, ApJ, 715, 736

Proga D., Stone J. M., Kallman T. R., 2000, ApJ, 543, 686

Rakshit S., Stalin C. S., Chand H., Zhang X.-G., 2017, ApJS, 229, 39

Rakshit S., Stalin C. S., Kotilainen J., 2020, ApJS, 249, 17

Ramos Almeida C., et al., 2011, ApJ, 731, 92

Rayner J. T., Toomey D. W., Onaka P. M., Denault A. J., Stahlberger W. E., Vacca W. D., Cushing M. C., Wang S., 2003, PASP, 115, 362

Reichard T. A., et al., 2003, AJ, 126, 2594

Reynolds C. S., Fabian A. C., 1995, Monthly Notices of the Royal Astronomical Society, 273, 1167

Risaliti G., Lusso E., 2015, The Astrophysical Journal, 815, 33

Risaliti G., Lusso E., 2018, Cosmological constraints from the Hubble diagram of quasars at high redshifts, doi:10.48550/ARXIV.1811.02590, `https://arxiv.org/abs/1811.02590`

Ro S., 2019, ApJ, 873, 76

Rodríguez-Pascual P. M., et al., 1997, ApJS, 110, 9

Rosas V. G., et al., 2022, Nature, 602, 403

Rose T., et al., 2019, Monthly Notices of the Royal Astronomical Society, 489, 349

Ruan J. J., Anderson S. F., Eracleous M., Green P. J., Haggard D., MacLeod C. L., Runnoe J. C., Sobolewska M. A., 2019, ApJ, 883, 76

Sacchi A., et al., 2022, A&A, 663, L7

Salpeter E. E., 1964, ApJ, 140, 796

Salvestrini F., Risaliti G., Bisogni S., Lusso E., Vignali C., 2019, A&A, 631, A120

Schaye J., et al., 2015, MNRAS, 446, 521

Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525

Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337

Shen Y., Liu X., 2012, ApJ, 753, 125

Shen Y., et al., 2011, ApJS, 194, 45

Shen Y., et al., 2016, ApJ, 818, 30

Shen Y., et al., 2019, ApJS, 241, 34

Sitko M. L., Sitko A. K., Siemiginowska A., Szczerba R., 1993, ApJ, 409, 139

Skrutskie M. F., et al., 2006, AJ, 131, 1163

Steele I. A., 2004, Astronomische Nachrichten, 325, 519

Stern B. E., Poutanen J., Svensson R., Sikora M., Begelman M. C., 1995, ApJ, 449, L13

Strüder L., et al., 2001, A&A, 365, L18

Suganuma M., et al., 2006, ApJ, 639, 46

Sun M., et al., 2015, ApJ, 811, 42

Sun M., Grier C. J., Peterson B. M., 2018, PyCCF: Python Cross Correlation Function for reverberation mapping studies, Astrophysics Source Code Library, record ascl:1805.032 (ascl:1805.032)

Svoboda J., Guainazzi M., Merloni A., 2017, A&A, 603, A127

Tananbaum H., et al., 1979, ApJ, 234, L9

Torrealba J., Chavushyan V., Cruz-González I., Arshakian T. G., Bertone E., Rosa-González D., 2012, Rev. Mexicana Astron. Astrofis., 48, 9

Tran H. D., 2003, ApJ, 583, 632

Tristram K. R. W., et al., 2009, A&A, 502, 67

Urry C. M., Padovani P., 1995, PASP, 107, 803

Van Groningen E., Wanders I., 1992, PASP, 104, 700

Vasudevan R., 2008, The Effective Eddington Limit for AGN, XMM-Newton Proposal ID 06050903

Vasudevan R. V., Fabian A. C., 2007, MNRAS, 381, 1235

Vazquez B., et al., 2015, ApJ, 801, 127

Vestergaard M., Peterson B. M., 2006, ApJ, 641, 689

Volonteri M., Sikora M., Lasota J.-P., 2007, ApJ, 667, 704

Wang F., et al., 2021, The Astrophysical Journal Letters, 907, L1

Waters T., Proga D., Dannen R., Dyda S., 2022, The Astrophysical Journal, 931, 134

Webb N. A., et al., 2020, A&A, 000, 0

Woo J.-H., Yoon Y., Park S., Park D., Kim S. C., 2015, ApJ, 801, 38

Woo J.-H., Le H. A. N., Karouzos M., Park D., Park D., Malkan M. A., Treu T., Bennert V. N., 2018, preprint, (`arXiv:1804.02798`)

Yip C. W., et al., 2004a, AJ, 128, 585

Yip C. W., et al., 2004b, AJ, 128, 2603

York D. G., et al., 2000, AJ, 120, 1579

Zajaček M., et al., 2021, The Astrophysical Journal, 912, 10

Zel'dovich Y. B., 1965, Soviet Ast., 9, 221

Zhou H., Wang T., Yuan W., Lu H., Dong X., Wang J., Lu Y., 2006, ApJS, 166, 128

Zhu D., Sun M., Wang T., 2017, ApJ, 843, 30

Zhu S. F., Brandt W. N., Luo B., Wu J., Xue Y. Q., Yang G., 2020, MNRAS, 496, 245

Zu Y., Kochanek C. S., Peterson B. M., 2011, ApJ, 735, 80