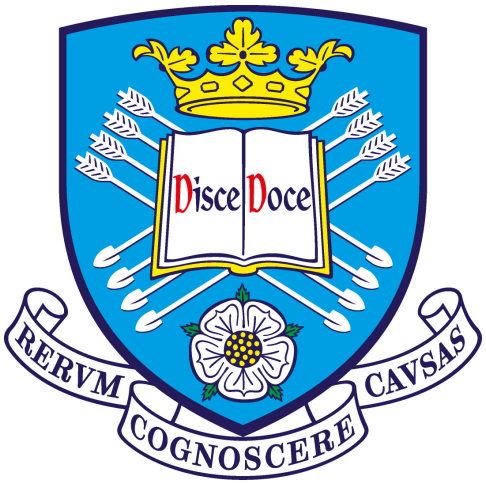
A statistical investigation into the relationship between supermassive black hole growth and star formation

Liam Philip Grimmett

Department of Physics & Astronomy
The University of Sheffield



A dissertation submitted in candidature for the degree of Doctor of Philosophy at the University of Sheffield

September 2020

Contents

List of Figures

List of Tables

Declaration

I declare that no part of this thesis has been accepted, or is currently being submitted, for any other qualification at the University of Sheffield or elsewhere.

This thesis is the result of my own work unless otherwise stated.

The chapters below are based on my own publications:

- Chapter 3 is adapted from Grimmett et al. (2019) - *"Revealing the differences in the SMBH accretion rate distributions of starburst and non-starburst galaxies"*
- Chapters 4 and 5 are adapted from two publications:
  - Bernhard et al. (2019) - *" Inferring a difference in the star-forming properties of lower versus higher X-ray luminosity AGNs"*
  - Grimmett et al. (2020) - *"A binning-free method reveals a continuous relationship between galaxies' AGN power and offset from the main sequence'*

Note that I was not the primary author for Bernhard et al. (2019) - I was second author. For that work, I helped with the initial development of the paper, helped with the statistical modelling of the $R_{\rm MS}$ distribution, calculated the stellar masses for the sample and assisted with the analysis and interpretation of the results.

Beyond the content of this thesis, during my PhD I have also contributed to the following publications:

- Bernhard et al. (2018) - *"Evidence for a mass-dependent AGN Eddington ratio distribution via the flat relationship between SFR and AGN luminosity"*
- Schoettler et al. (2019) - *"Dynamical evolution of star-forming regions: III. Unbound stars and predictions for Gaia"*
- Delvecchio et al. (2019) - *"The Galaxy's Gas Content Regulated by the Dark Matter Halo Mass Results in a Superlinear M_{BH} -M_*Relation"*
- Delvecchio et al. (2020) - *"The Evolving AGN Duty Cycle in Galaxies Since $z \sim 3$ as Encoded in the X-ray Luminosity Function"*
- Bottrill et al. (2020) - *"Exoplanet Detection and Its Dependence on Stochastic Sampling of the Stellar Initial Mass Function"*

Acknowledgements

I have thoroughly enjoyed my PhD experience. But of course, there are a vast number of people who have helped significantly along the way.

I must thank my primary supervisor James M. for his continued support throughout my PhD. I understand taking a statistician was a risky play, but hopefully (notwithstanding me successfully defending this document) it was a risk that has paid dividends. I must also thank my secondary supervisor Kevin for his important contributions to my work. I also have to pay tribute to Manu, who has often provided advice and support on a less-formal, but nonetheless important basis. Outside of these, I must thank the Sheffield AGN group of Clive, Jonny and Lydia, and my office-mates throughout the years, Becky, Christina, Dave, Emma, Gemma, Héloïse, James B., James W., Katie and Umar, who have all provided much needed support when I was asking the dumbest of Astro-related questions. I'd like to also thank Richard and Simon for their weekly sustenance-based support meetings on Fridays at 3pm (usually earlier, never later).

Of course, I have to mention the Sheffield University Rugby League Club too. Throughout my time at Sheffield (including the four years before my PhD), that club has provided me with two housemates, Cameron and Veazey, as well as many lifelong friends (of which there are simply too many to mention) and unlimited high quality "memriz" (or indeed, lack of). I thank everyone in the Club over my many years for their support. The colour scheme of this thesis (black and gold) was chosen to match the Club's colours.

Outside of Sheffield completely, I've always had the privilege of a strong group of friends back home in Wakefield, and in particular Fainty, Gretts, Lodgey, Malley and Richie, who I've managed to stay close with during my time away. I thank them for their constant and important friendship.

Finally, I must thank my family – specifically my parents Lynden and Tracy. It has been a rough ride, but it wouldn't be the Grimmett way if it wasn't.

P.S. I'd like to also thank James A. and Simon for taking the time to read this document and for conducting my PhD defence.

Abstract

In an attempt to reveal whether any coevolution between a central SMBH and its host galaxy exists, the literature has sought to identify whether the growth rates of the two are connected (witnessed as Active Galactic Nuclei, AGN, and star-formation respectively). However, depending on the sample selection method, there appears to be a contradiction in the results, with those studies selecting a galaxy sample reporting a positive correlation and those studies selecting an AGN sample reporting a flat relationship. In order to include non-detections in the analysis, the majority of these studies resort to a binning-and-averaging approach and thus investigate how the average star-formation changes across AGN power bins, or vica versa. However, binning and averaging both have limitations.

In this thesis, we conduct a detailed statistical analysis of the relationship between the SMBH accretion rate and the star formation rate (SFR) of the host galaxy. We firstly investigate how the full distribution of SMBH accretion rates changes between galaxies with excess star formation (i.e., starburst galaxies) and those with lower levels of star formation. Secondly, we investigate how the full distribution of star-forming properties changes between high and low power AGNs, before moving on to present a binning-free methodology to investigate how the same distribution changes continuously with AGN power. Therefore, conducting analyses that moves beyond the binning-and-averaging approach. In general, we find a statistical connection between SMBH accretion rate and SFR likely exists such that more rapidly accreting SMBHs tend to reside in galaxies with higher levels of star formation. We propose that this scenario is consistent with a proposed coevolution of SMBH and galaxy growth in that they are both cogoverned by cold molecular gas availability in the host galaxy.

Chapter 1

Introduction

I'm going on an adventure!

Bilbo Baggins

1.1 The discovery of quasar-like objects

In the mid 20th century, studies had begun to explain the nature of a series of extragalactic sources that, as a result of their stellar-like brightness at galaxy-like distances, did not seem to match any known astronomical phenomenon. These sources, now dubbed Active Galactic Nuclei (AGN), were first systematically studied in 1943, when Seyfert (1943) identified six AGNs that had both broad emission lines in the spectrum and excessive nuclear emission; two unusual features for the galaxy population as it was known then. In a key discovery twenty years later, Schmidt (1963) measured the distance to another AGN (namely, 3C 273) and determined that, if it was at the distance inferred from the redshifted emission lines in it spectra (i.e., 500 megaparsecs), its luminosity must be around ≈10 times larger than a typical galaxy (i.e., $\approx 2 \times 10^{47}$ erg s^{-1}, Courvoisier 1998)[1]. By 1969, 44 of these AGNs had been identified (see the review by Schmidt, 1969, and ref-

[1]During my PhD, I quickly learned that "typical" is an unusual term to describe a galaxy. They are similar to humans in that they share common properties, but on the whole every galaxy is likely unique.

erences therein) and whilst identification of such sources continued significant amounts of research began to try and explain what physical processes were responsible for these excessively bright and distant sources (e.g., Sandage, 1965; Burbidge, 1967; Schmidt, 1969).

It is now widely accepted that, as well as the many other AGNs discovered since, those 44 sources and their excessive nuclear emission is likely powered by accretion of gas and dust on to a central supermassive black-hole (SMBH, e.g., Salpeter, 1964; Zel'dovich, 1964; Lynden-Bell & Rees, 1971; Pringle et al., 1973; Ruffini & Wilson, 1975; Lynden-Bell, 1978; Rees, 1984, 1998, but also see Section 1.2 for more details). Salpeter (1964) first suggested that the accretion of matter on to a small, but incredibly massive object ($> 10^6$ $M_\odot$) was likely an incredibly efficient process (with 5-20% of the rest mass being converted into radiation), meaning vast amounts of energy could be released, with not unrealistic levels of accretion. For example, accretion of 0.002 $M_\odot$ of gas per year at 10% efficiency would result in a luminosity of 10^{43} erg s^{-1}, which would be enough to outshine the Milky Way. Whilst SMBHs have, therefore, long been suspected to be responsible for powering AGNs, arguably the most definitive piece of evidence came in 2019, when the Event Horizon Team directly imaged the SMBH present in the centre of the nearby AGN M87 (see Figure 1.1).

Large numbers of central SMBHs, like the ones powering the aforementioned AGNs, have also been "detected" in their inactive state in nearby massive galaxies (e.g., Lynden-Bell, 1969; Sargent et al., 1978; Tonry, 1984, 1987; Dressler & Richstone, 1988; Kormendy & Richstone, 1995; Bender et al., 1996; Kormendy et al., 1997). These SMBHs are not actively accreting and are therefore, unlike their active counterparts, dark (or at least not bright enough for us to detect). As a result, identifying these SMBHs relies on indirect techniques, such as measuring the velocities (or orbits) of nearby gas or stars. Indeed, by tracing the proper motions of the stars within the central 0.1 parsecs of the Milky Way, a dormant SMBH (with mass $\approx 2.5 \times 10^6$ $M_\odot$) has been identified in the

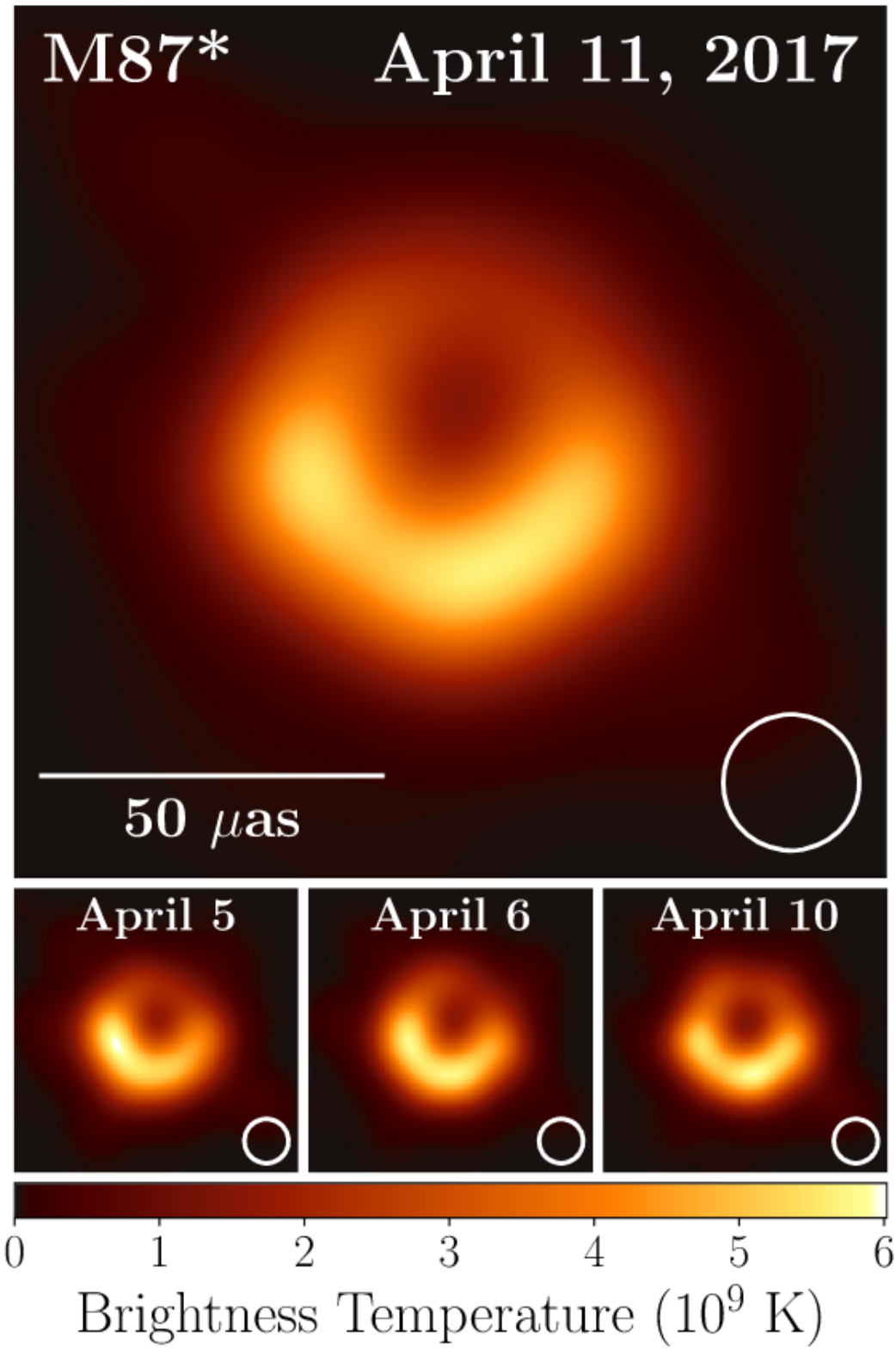


Figure 1.1: The shadow of the SMBH in the local AGN M87. This image was taken by the Event Horizon Telescope collaboration, using many different radio observatories. The use of many telescopes allows for incredibly long baselines to be used, giving the telescope the resolution required to resolve down to the scale of the central SMBH. For size comparison, the solar system would easily fit in the shadow. This Figure was originally presented in Event Horizon Telescope Collaboration et al. (2019).

centre of our own Galaxy (e.g., Lynden-Bell & Rees, 1971; Eckart & Genzel, 1996, 1997; Ghez et al., 1998; Lo et al., 1998). Given the frequency at which these inactive SMBHs are detected in nearby galaxies (see the review by Kormendy & Richstone, 1995, and references therein), it is now widely accepted that most galaxies host a SMBH in their centre.

Interestingly, some recent studies have demonstrated that the mass of these SMBHs tends to correlate with properties of the host galaxy they reside within (e.g., Magorrian et al., 1998; Ferrarese & Merritt, 2000; Gebhardt et al., 2000; Merritt & Ferrarese, 2001; Tremaine et al., 2002; Marconi & Hunt, 2003; Häring & Rix, 2004; Wyithe, 2006; Hu, 2008; Gültekin et al., 2009; McConnell & Ma, 2013; de Nicola et al., 2019; Ding et al., 2020a,b). Figure 1.2 (originally presented in McConnell & Ma, 2013) shows the correlation between the SMBH mass and the stellar velocity dispersion of the host galaxy for a sample of 72 nearby galaxies. However, similar correlations exist for AGNs and more distant galaxies. Figure 1.3 highlights the correlation between SMBH mass and host galaxy bulge mass for a sample of 32 AGNs at $1.2 < z < 1.7$ presented in Ding et al. (2020b) but also contains the works of Bennert et al. (2011) and Häring & Rix (2004), which show the correlations for intermediate redshift AGNs and local non-AGN galaxies, respectively. These correlations tend to suggest that a SMBH and its host galaxy are connected and, given that the host galaxy (out to 10s kpc) lies beyond the gravitational influence of the black hole (< 1 kpc), any connection would not be fully explained by gravity alone. The naturally arising question, is therefore, what "macroscopic" properties of the host galaxy, if any, are dictating the activity levels of the SMBH (as even inactive SMBHs must have been historically-active to grow their mass). Or more succinctly, why is it that some SMBHs are active and some are not, and why does it appear to be related to the host galaxy? Before progressing with our research addressing this question in the forthcoming chapters, the remainder of this introduction aims to provide further context and discuss the findings of the literature thus far. Therefore in Section 1.2 we highlight

the "unified model of AGN" which describes the fundamental components in an AGN system, which is important for considering how to measure SMBH accretion rate. In Section 1.3.1 we summarise the evidence to suggest that the true connection between a SMBH and its host galaxy may be in the way they have co-evolved. In the same section, we then discuss how the literature has provided inconclusive results, depending upon on the analysis methods used, before highlighting areas of improvement in the statistical analysis of the relationship between SMBH growth and galaxy growth in Section 1.3.2.

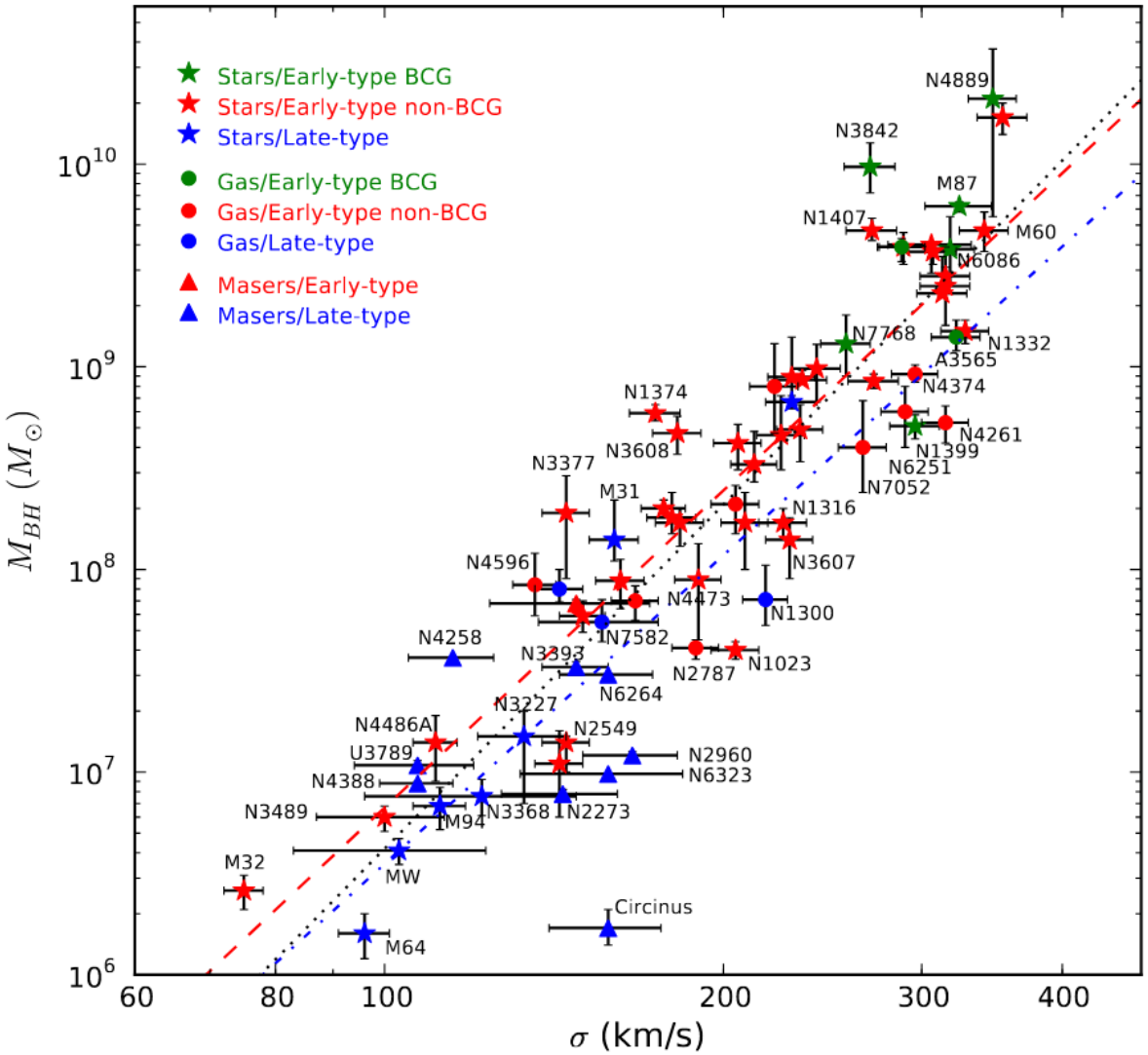


Figure 1.2: The $M_{\rm BH} - \sigma_{\rm vel}$ relation; the observed correlation between the mass of a central SMBH and the stellar velocity dispersion of the host galaxy for a sample of 72 nearby galaxies. The differences in points represent both the morphological differences of the galaxy (colour) or the method used to estimate $M_{\rm BH}$ (marker type). This figure was originally presented in McConnell & Ma (2013).

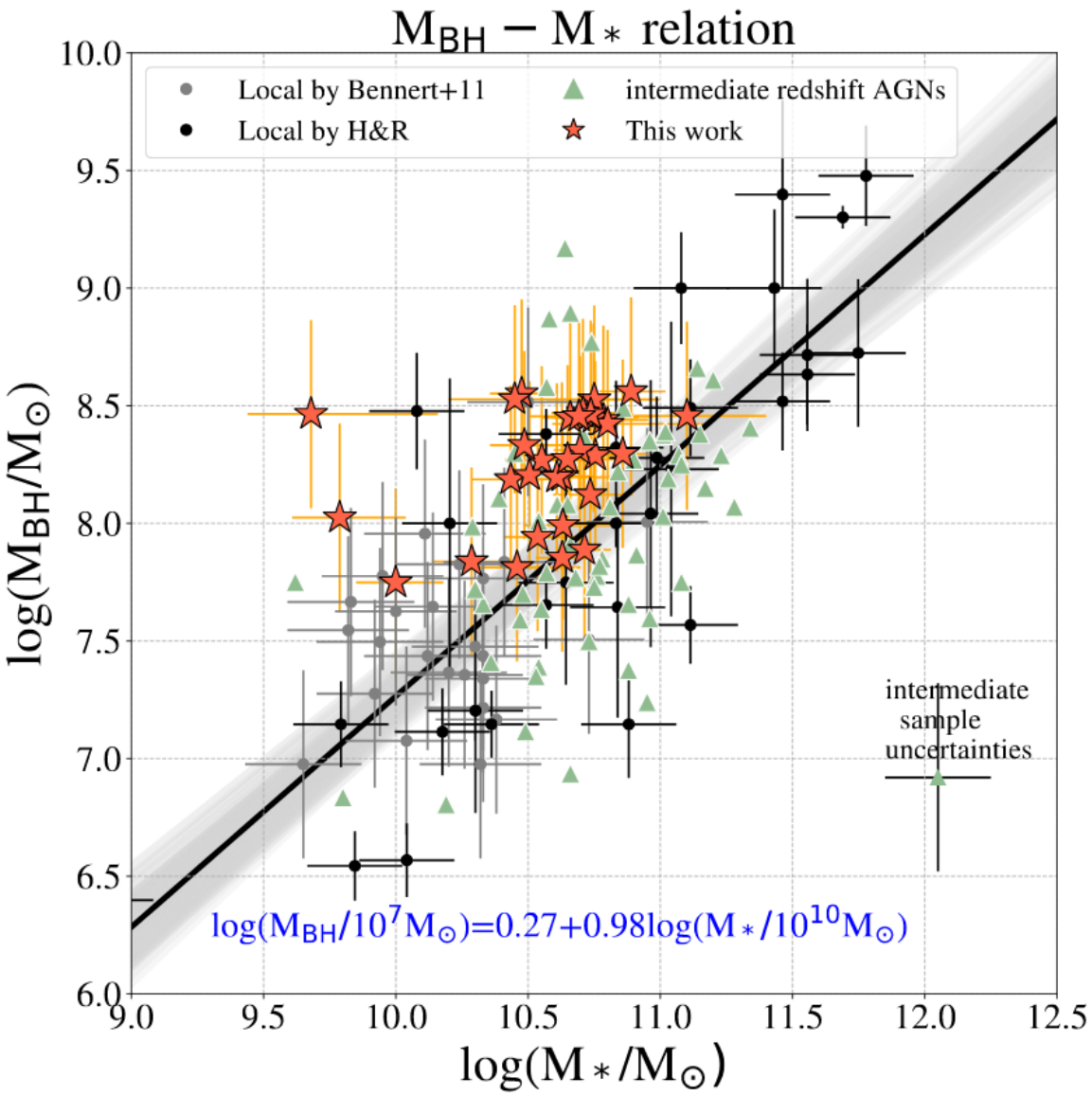


Figure 1.3: The $M_{\rm BH} - M_{\rm bulge}$ relation; the relationship between host galaxy bulge mass and SMBH mass for a sample of 32 AGNs that were X-ray selected between the redshift range $1.2 < z < 1.7$ (shown as red stars). Over-plotted are the results from non-AGN galaxies at lower redshifts and intermediate redshift AGNs from the works of Bennert et al. (2011) and Häring & Rix (2004), implying correlations between SMBH and host galaxy hold for different galaxy populations. This figure was originally presented in Ding et al. (2020b).

1.2 The AGN system

As was previously mentioned, the most widely accepted mechanism for powering AGNs is the accretion of gas and dust on to a SMBH. However, the full system is somewhat more complicated. As this thesis aims to investigate the connection between the SMBH accretion rate and properties of the host galaxy, we need to understand the inner-workings of the AGN system. It is also important to understand the system in order to choose an accurate SMBH accretion rate tracer (which we cover in Section 2.2). In short, all AGNs are considered to be intrinsically similar in the way they work, with the majority of observed differences (such as the presence or absence of broad emission lines or lack of nuclear emission) explained by either changes in the viewing angle or the accretion rate. A sketch of this "unified model" of AGN is presented in Figure 1.4.

During a growth phase, a SMBH is thought to be surrounded by an accretion disk (with a radius of $\approx$ 10 light days, Hawkins 2007) which, as a result of thermal emission, is bright at UV-optical wavelengths (e.g., Shakura & Sunyaev, 1973; Blandford & Znajek, 1977; Shields, 1978; Ulrich et al., 1980; Malkan & Sargent, 1982; Blaes, 2007). In addition to this UV-optical emission, low energy X-ray photons may also be produced in the innermost, and hottest, regions of the accretion disk (e.g., Shakura & Sunyaev, 1973; Mushotzky et al., 1993; Reynolds & Nowak, 2003; Sobolewska et al., 2004; Fabian et al., 2006; Turner & Miller, 2009; Done & Diaz Trigo, 2010; Gilfanov & Merloni, 2014; Kubota & Done, 2018; Petrucci et al., 2018). Higher energy X-ray photons can also be produced by a small corona (a few tens of light minutes in size but see Dovčiak & Done 2016 for a discussion on the size of the corona) that resides just above the accretion disk (e.g., Vaiana & Rosner, 1978; Haardt & Maraschi, 1993; Fabian et al., 2015). These X-ray photons are produced as a result of lower energy photons being up-scattered by high energy electrons in a process called inverse Compton scattering (e.g., Liang & Price, 1977; Galeev et al., 1979; Haardt & Maraschi, 1991; Haardt et al., 1994; Stern et al.,

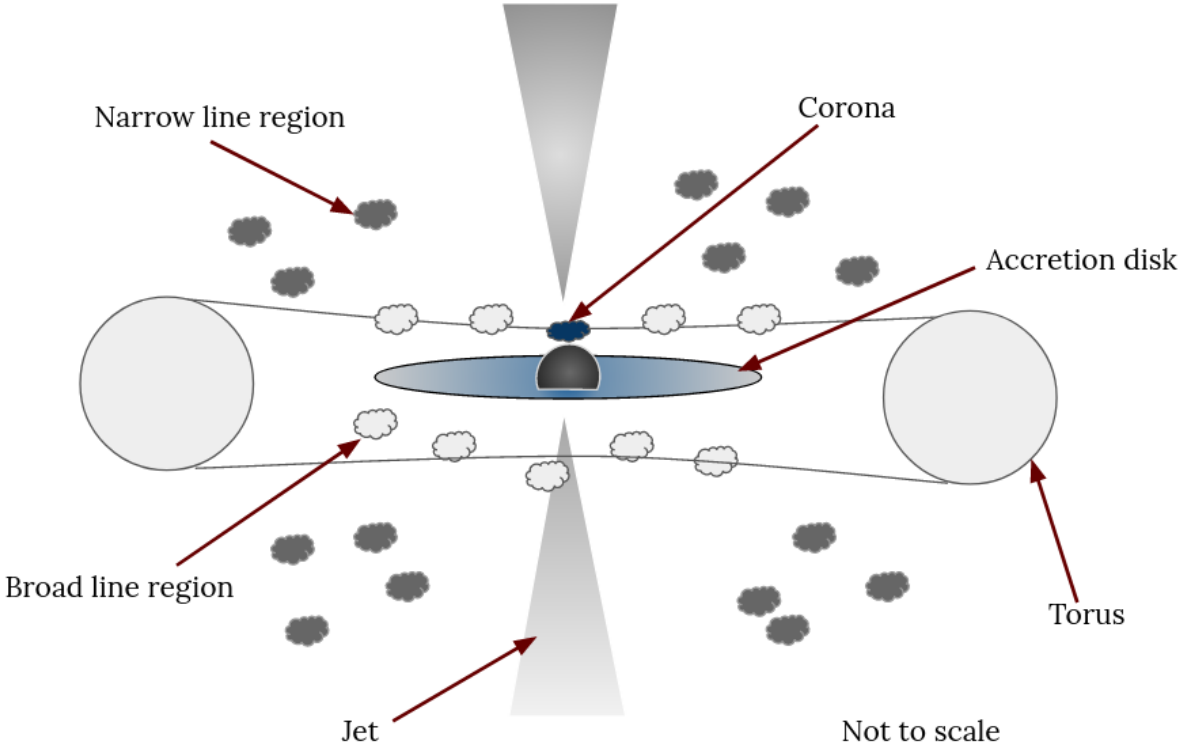


Figure 1.4: The unified model of AGN. Theory predicts that SMBHs in the centre of every galaxy are built from the same components, whilst orientation and accretion rate can explain all the observational differences we see. Surrounding the SMBH is an accretion disk (the properties of which are likely constrained by the precise accretion rate). Just beyond the accretion disk are small clouds of gas, referred to as the broad line region. A small X-ray emitting corona also sits just above the accretion disk. Surrounding the accretion disk and the broad line region, is likely a dusty torus, which can obscure some emission from the central components. Beyond the torus are other clouds of gas (narrow line region), but being further away from the gravitational influence of the black hole, have narrower emission lines than the broad line region. Some AGNs also show the presents of radio jets.

1995).

Beyond the accretion disk are small clouds of gas (extending out to ≈ 100 light days depending on AGN luminosity, Pozo Nuñez et al. 2015) that are illuminated by the accreting SMBH. As a result of being close to (hence under the gravitational influence of) the SMBH, these gas clouds tend to have broad optical emission lines ($\approx$ 1000s of km s^{-1} and are hence named the broad line region, BLR, e.g., Antonucci & Miller 1985). Surrounding the inner-part of the system is a structure referred to as the dusty torus, which can, if it lie in the light of sight, obscure emission from the accretion disk and the BLR (e.g., Miller & Goodrich, 1986; Krolik & Begelman, 1988). As a result of obscuration, the dusty torus is likely to be heated by emission from the accretion disk and can, therefore, re-emit photons at mid-infrared wavelengths. Gas clouds can also reside beyond the torus (out to kpc scales) which, being lesser influenced by the gravity of the SMBH, have narrower emission lines than the BLR ($\approx$ 100s km s^{-1}) and are thus named the narrow line region (NLR). The final structure in the system is the presence of radio jets. These are, however, only seen in $\approx 10\%$ of AGNs and are thought to be dependent on the precise nature of the accretion, rather than orientation (see Heckman & Best, 2014, and references therein).

The luminosity of the corona, accretion disk, BLR, dusty torus and NLR are all, at least fundamentally, dependent upon the level of accretion of the SMBH. Our ability to detect different levels of emission, however, depends upon our line of sight orientation. Whilst the unified model is unlikely to be the true description of all AGN systems (one example being the discussion around the precise shape, or clumpiness, of the torus, e.g., Elitzur & Shlosman 2006), it is likely to be a viable approximation of reality. It also allows us to pinpoint the different observational signatures of AGNs to different components within the system. Again, this is important when we consider how best to measure the growth rate of a SMBH.

1.3 Connecting the SMBH to the host galaxy

Whilst subtle improvements to the unified model are still being made (see the recent review by Netzer 2015), an overwhelming amount of AGN research now aims to investigate, *whether* an AGN is connected to its host galaxy and if so, *how?* In order to investigate any potential connection between an AGN and its host galaxy, two potential research-philosophies can be adopted. Firstly, the statistical analysis of large samples can help by revealing *whether* a statistical connection between properties of the AGN and properties of the host galaxy exists. Secondly, more detailed examination of (usually fewer) sources is required to identify *which physical processes* are responsible for any observed connections. This dichotomy in research-philosophy is not only applicable in astronomy. For example, in a medical setting clinical trials determine *whether* a particular drug has an impact on the general population (i.e., does this drug have an impact on the human body and if so, what changes?), whereas more specific studies aim to identify *which biological processes* are responsible (i.e., identify the direct impact of the drug on specific human cells). Analogously, in observational AGN astronomy, large statistical studies uncover potential connections between properties of the SMBH and the host galaxy whilst studies of a more precise nature identify potential connecting processes. Only by adopting both research-philosophies can rapid progress be made in our scientific understanding. This thesis adopts the former research philosophy, but given its importance, we summarise what physical mechanisms could be responsible for the results we see in Chapter 7, where we discuss the context of our results against the backdrop of AGN feedback (i.e., the impact of SMBH accretion on the host galaxy) and gas availability (the fuel that drives both SMBH accretion and star formation).

1.3.1 Whether a SMBH is connected to its host galaxy

The aforementioned studies identifying correlations between the properties of SMBHs and their host galaxies were not the only statistical sample-based studies seeking to determine whether an AGN is connected to its host galaxy. During the late 1990s and 2000s, further evidence for a potential connection between SMBHs and galaxies was found by investigating their relative growth rates. Since $z \approx 3$, the volume-averaged SMBH growth density, witnessed as AGN luminosity per cubic Mpc, has closely traced the volume-averaged galaxy growth density, witnessed as star formation per unit Mpc (e.g., Boyle & Terlevich, 1998; Heckman et al., 2004; Merloni et al., 2004; Silverman et al., 2008; Aird et al., 2010). The two are systematically offset (star formation is around 3-4 orders of magnitude more prevalent than SMBH accretion) but they both appear to have peaked at $z \approx 2$ and have both declined towards more recent redshifts. Figure 1.5, which appeared in the white paper by Pope et al. (2019), shows the results of recent works of Madau & Dickinson (2014); Delvecchio et al. (2014); Aird et al. (2015); Vito et al. (2018). The results show the star formation rate (SFR) density and SMBH accretion rate density are in good agreement up to $z \approx 3$. Whilst beyond that redshift they appear to deviate from one another, for the vast majority of the Universe's lifetime they have appeared to follow a similar evolutionary track. These results, taken with the correlations between SMBHs and host galaxies shown previously, suggest that the connection between a SMBH and its host galaxy may be connected is in the way they have evolved over time, i.e., the connection may be between SMBH growth, witnessed as AGN power and galaxy growth, witnessed as star formation.

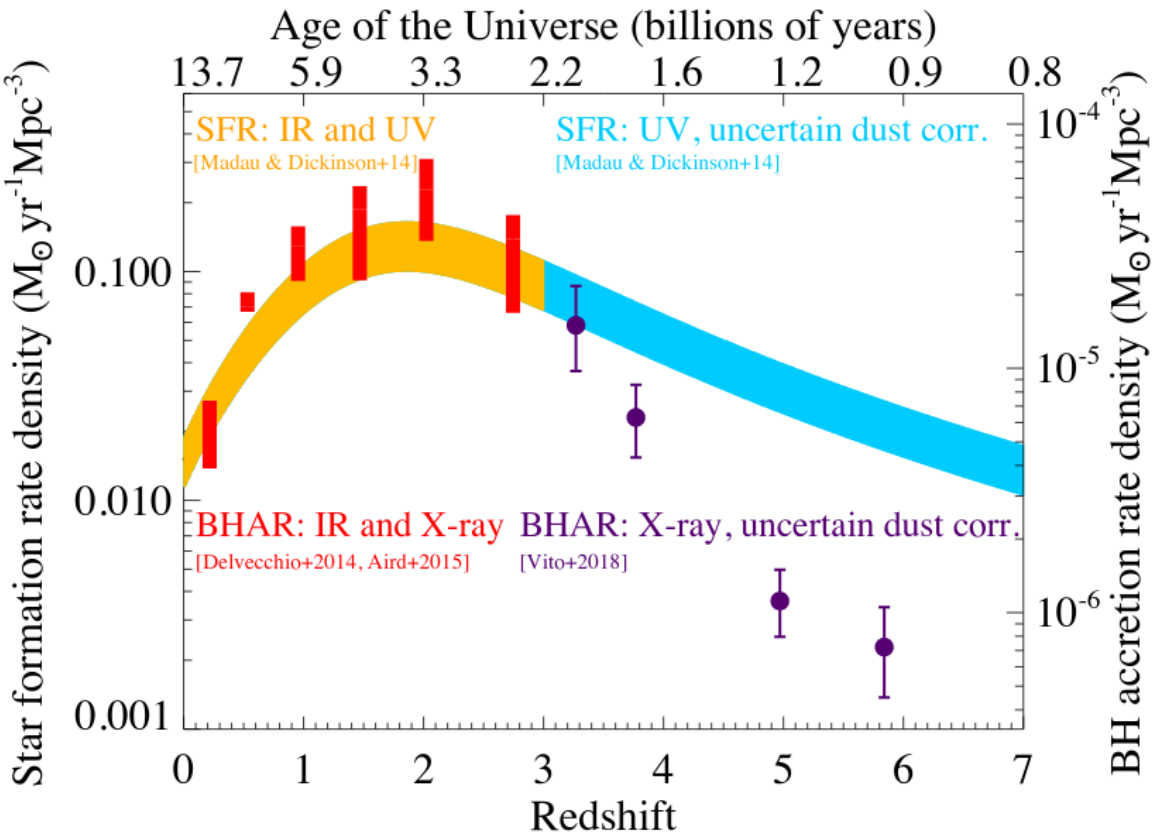


Figure 1.5: The SMBH accretion rate density and the SFR density as a function of redshift. From a redshift of $z \approx 3$ (i.e., the vast majority of the Universe's lifetime) the two trace each other remarkable closely, implying that whenever SMBHs have been accreting, galaxies have also been growing. This provides further evidence that a SMBH and a host galaxy may be connected. This figure was originally presented in Pope et al. (2019).

1.3.2 The correlation between SMBH accretion and star formation

In an effort to determine whether the growth rate of a SMBH and the growth rate of a galaxy are interconnected, a number of recent studies have adopted detailed statistical analyses of the two processes. These studies, however, often take one of two approaches:

- **Approach A:** Investigate how the SMBH accretion rate varies across a sample of galaxies that are classified based on their star-forming properties.
- **Approach B:** Investigate how the SFR varies across a sample of AGN-hosting galaxies that are classified based on their SMBH accretion rates.

The key difference in the two approaches is the way in which the sample is selected. Approach A selects a sample of galaxies that are not specifically AGNs, although some will be, and investigates how the SMBH accretion rate varies as a function of star formation of the host galaxy. Therefore, Approach A asks the specific question - *How does the SMBH accretion rate change as a function of the galaxy population's SFR?* Approach B differs in that a sample of AGNs is initially selected and investigates how the SFR of these AGNs change, as a function of their accretion rate. Thus, Approach B asks a subtly different question - *How do the star-forming properties of AGNs change as a function of their accretion rate?* One potential reason for this mixed-approach is that studies investigating the relationship between SMBH growth and SFR often adopt a binning-and-averaging approach to help include non-detections in their sample (i.e., bin the data in one axis, and average in the other). Whilst a process called stacking[2] can be used to help with non-detections in the averaging axis, there is no consistent way to accurately bin a source in which only an upper (or lower) limit is known, therefore

[2]By adding together the flux of non-detected sources (i.e., stacking them), the signal from the sources increases linearly whilst the background noise increases slower than the signal. The resulting stack therefore has a 'detectable' signal-to-noise ratio. The flux from this stack is then divided by the number of sources in the stack to create an average measurement from the non-detections.

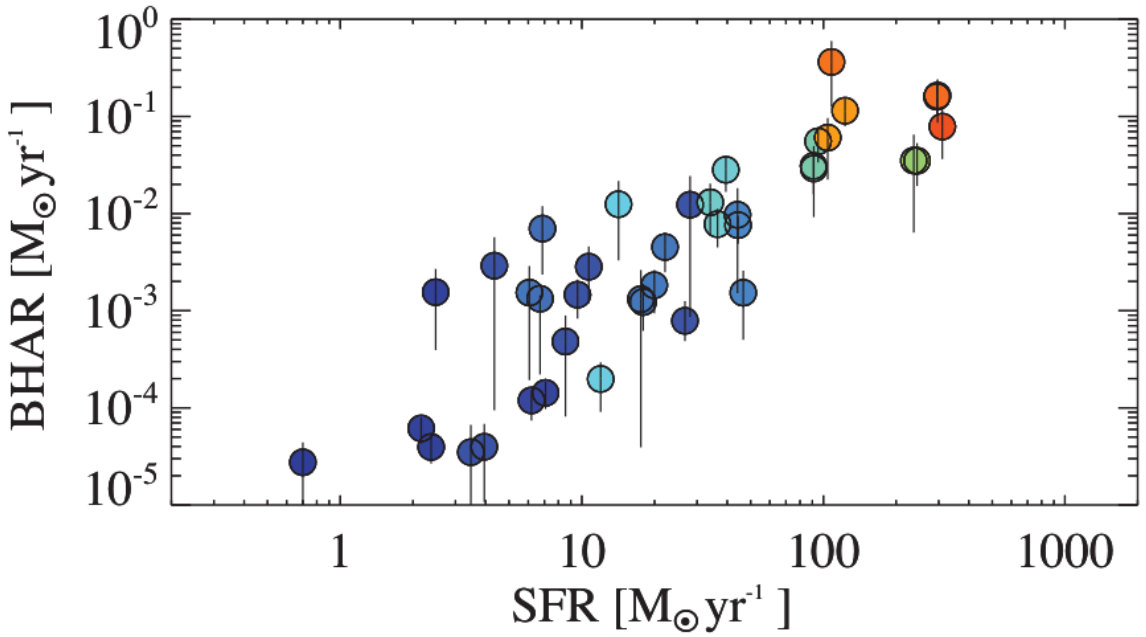


Figure 1.6: The positive correlation witnessed between average SMBH accretion rate in bins of SFR for a sample of ≈ 8600 star-forming galaxies. This figure was originally presented in Delvecchio et al. (2015).

explaining the differences in the sample selection method previously mentioned (as one axis must be fully-detected in order to bin the data).

Recent studies that have investigated how the *average* AGN power (tracing accretion rate) changes between groups of galaxies *binned* by their SFR (i.e., Approach A, e.g., Rafferty et al., 2011; Rosario et al., 2012; Chen et al., 2013; Azadi et al., 2015; Delvecchio et al., 2015; Harris et al., 2016; Lanzuisi et al., 2017; Shimizu et al., 2017; Stemo et al., 2020) have generally found that the average AGN power increases as a function of SFR. Figure 1.6 shows the correlation between average SMBH accretion rate in bins of SFR for ≈ 8600 star-forming galaxies up to a redshift of $z \sim 2.5$ from the work of Delvecchio et al. (2015). Conversely, however, those studies that investigated how the *average* SFR changes between groups of AGNs *binned* in terms of AGN power tend to find less evidence of a correlation (i.e., Approach B, e.g., Harrison et al., 2012; Rosario et al., 2012; Stanley et al., 2015, 2017; Suh et al., 2017; Ramasawmy et al., 2019). Figure 1.7

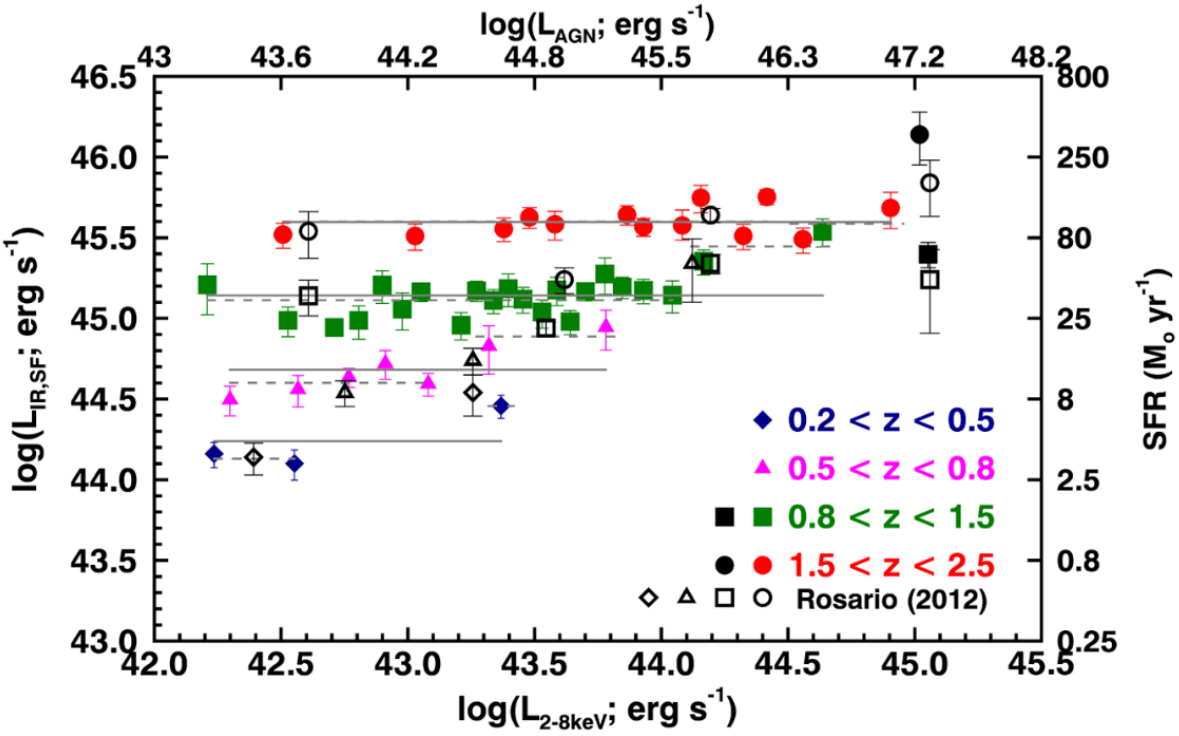


Figure 1.7: The flat relationship witnessed between average SFR in bins of AGN power derived using a sample of ≈ 2000 X-ray detected AGNs. This figure was originally presented in Stanley et al. (2015).

shows the flat correlation between average SFR in bins of AGN power from Stanley et al. (2015) as shown in Figure 4 of that paper. When taken as a whole, these results provide a complicated reality. Studies adopting Approach A suggest that the average SMBH growth rate is higher in galaxies with more star formation (i.e., the growth rates of a SMBH and a host galaxy are likely connected), whilst studies adopting Approach B suggest that the average SFR does not change with AGN power (i.e., the growth rates of a SMBH and a host galaxy are likely not connected). If we are to understand the true nature of how SMBH accretion is connected to host galaxy star formation, we need to make sense of this apparent contradiction in results seen when adopting a binning-and-averaging approach. One potential explanation for these contradictory results is the uncertainties associated with both binning and averaging, which we cover in the forthcoming subsections.

AGN variability and the use of binning

The accretion of gas on to a SMBH is known to be a highly-variable process. In a recent study, McAlpine et al. (2017) mapped out the SMBH accretion rate for an individual galaxy in the Evolution and Assemble of GaLaxies and their Environments simulation (EAGLE, Schaye et al., 2015)[3]. The SMBH accretion rate traced by McAlpine et al. (2017) is shown in the bottom plot of Figure 1.8. This particular galaxy in the simulation not only had a SMBH accretion rate that varied by ≈ 10 orders of magnitude during the lifetime of the simulation, but one that, even at its most stable, could vary by ≈ 4 orders of magnitude over the course of a few megayears. Simulations have been useful for allowing us to track variability over such (relative to humans) long timescales, but even incredibly short term (≈ 20 years) observations have identified such rapid variability in AGN-hosting systems (e.g., Mushotzky et al., 1993; Ulrich et al., 1997). The process of star formation, however, is thought to be much more stable than SMBH accretion and tends to vary considerably less on the same timescales (e.g., Gao & Solomon, 2004; Wu et al., 2005; Krumholz & Thompson, 2007; Wong, 2009). The top plot of Figure 1.8 shows the SFR evolution for the same galaxy as was shown for SMBH accretion (McAlpine et al., 2017). In contrast to the SMBH accretion rate, the SFR only varied by ≈ 3 orders of magnitude throughout the lifetime of the simulation, and at its most stable did not change by an order of magnitude within 100Myr. This means that, when we compare AGN power to SFR we are comparing a process that can vary rapidly in the short-term, against one that is more stable.

Interestingly, it is those studies that bin in the highly-stochastic AGN power axis that tend to find little evidence of a correlation. Whilst this could, of course, be because no intrinsic connection exists, it could also be because binning in such a highly variable axis can wash out potential long-term correlations (Hickox et al., 2014). Demonstrating this

[3]The EAGLE simulation is a large-scale cosmological simulation of the Universe (containing over 10,000 massive galaxies) designed to simulate galaxy formation and evolution.

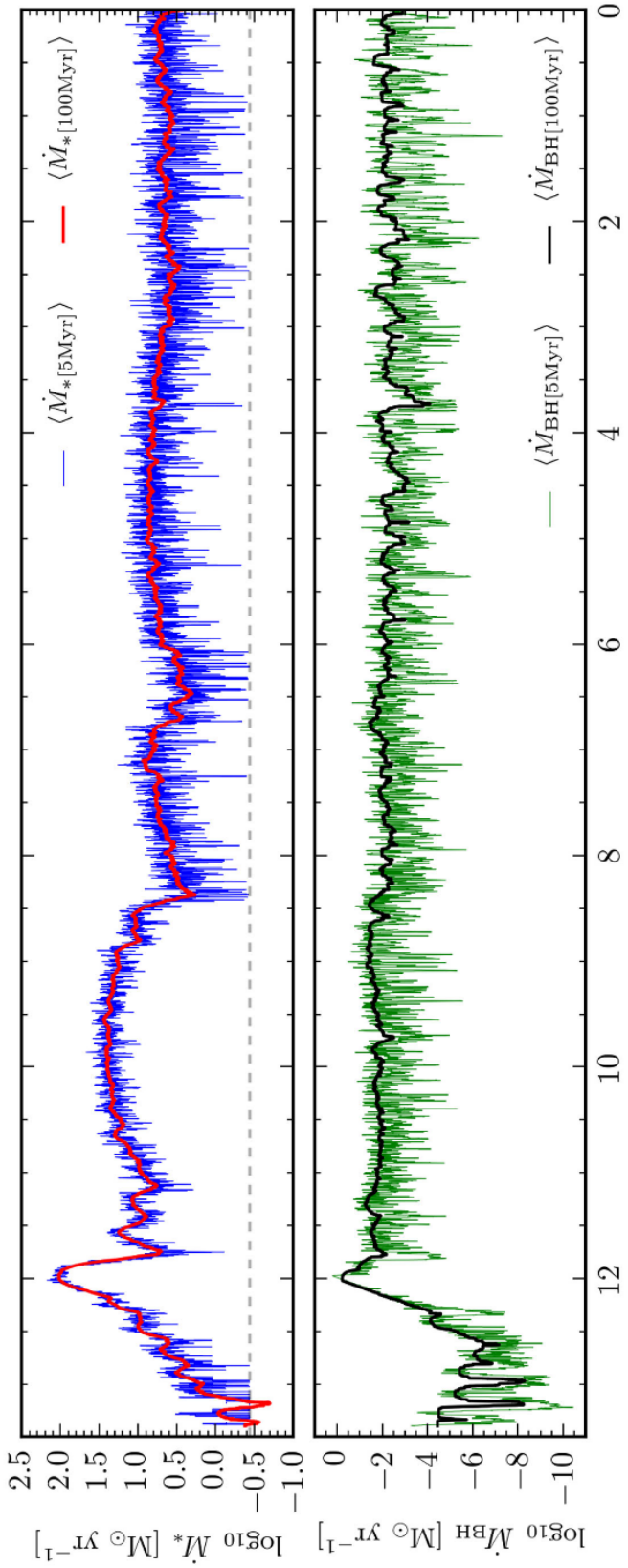


Figure 1.8: Top: The SFR (or galaxy growth rate) of an individual galaxy from the EAGLE simulation as a function of simulation time. The blue line represents the 5Myr running average, whilst the red line represents the 100Myr running average. Optically-derived SFRs tend to be representative on timescales of 10-20Myr, whilst FIR-derived SFRs tend to have longer representative timescales (up to 100Myr, but see Section 2.3). Bottom: Similar to above but for the accretion rate of the SMBH in the same galaxy as above. This time, the green line represents the 5Myr running average, whilst the black line represents the 100Myr average. Most measures of accretion rate (by measuring the AGN luminosity) are close to instantaneous measures of accretion. These figures suggest that over the same timescales, SMBH accretion is likely far more rapid than star formation. This figure was originally presented in McAlpine et al. (2017).

further, Lanzuisi et al. (2017) reported that, even within the same dataset, correlations between AGN power and SFR could change significantly depending upon the chosen binning axis. In addition to binning in a variable axis, there are, however, two further problems associated with binning data generally. Firstly, how is a source classified into a bin if, within errors, it could fall in to two or more bins? Some of the most recent studies adopting binning often take the measured value only (e.g., Delvecchio et al., 2015; Aird et al., 2018), which ignores uncertainties. A second limitation is the implied assumption that sources within one particular bin have similar (or even identical) properties but sources across bins are significantly different (or at least, it is hoped they are). Both of these limitations have the potential to lead to inconsistencies in results. To our knowledge, there has yet to be a study that, when investigating the relationship between AGN power and SFR has completely removed the need for binning in both the AGN power or SFR axis.

The use of averages

As was previously discussed, the need for the binning-and-averaging approach seen in some studies is motivated by the need to account for non-detections in the sample (e.g., Shao et al., 2010; Harrison et al., 2012; Mullaney et al., 2012b; Rosario et al., 2012; Chen et al., 2013; Azadi et al., 2015; Delvecchio et al., 2015; Stanley et al., 2015; Lanzuisi et al., 2017; Stanley et al., 2017). Including these non-detections is critical to ensuring that we capture the activity of the entire population, not just those SMBHs that are actively accreting. The most commonly used method for dealing with non-detections in these studies is by adopting a stacking approach. Whilst stacking is capable of incorporating non-detections into the analysis, it can only provide us with an average. However, averages are summary statistics and therefore only provide us with a simplified view of reality. This has been demonstrated by Mullaney et al. (2015), who investigated the full distribution of star-forming properties of AGN-selected galaxies and compared it to that

of non-AGN galaxies. Whilst having similar averages, that study demonstrated that AGNs had a significantly different star formation distribution than that of non-AGN galaxies (specifically that AGNs have a wider range of star-forming properties than 'main-sequence' galaxies, i.e., galaxies that have star formation proportional to their stellar mass, but see Section 2.3 for more details on main-sequence galaxies.) Indeed, there have also been a series of studies that have adopted the use of full distributions to reveal more detailed properties of the AGN population (e.g., Aird et al., 2010, 2012; Bernhard et al., 2016; Wang et al., 2017)

1.4 This thesis

As a result of the limitations associated with both binning and averaging, it is likely that the true relationship between SMBHs and their host galaxies is not fully revealed by studies adopting the binning-and-averaging approach. As a result of AGNs being incredibly stochastic in the short term, studies must be wary that binning sources by their AGN power has the ability to contaminate results. Likewise, only considering the average SFR (or average AGN power) can sometimes fail to provide us with the full picture. This thesis, therefore, aims to investigate the relationship between SMBH accretion rate and star-forming properties of the host galaxy in a more detailed statistical way.

In Chapter 2 we describe our techniques for measuring SMBH accretion rate and other relevant host galaxy properties. In an attempt to explain what is causing the average increase in SMBH accretion rate per unit star formation, in Chapter 3 we determine how the full distribution of SMBH accretion rates changes for galaxies with extreme levels of star formation and compare them to the general galaxy population. In Chapter 4 we investigate the distribution of star-forming properties between two samples binned in terms of AGN power, so that when remove the need for binning Chapter 5 we can

compare the improvements made from removing the need to bin. In Chapter 6 we discuss potential improvements to the binning-free methodology and provide a discussion and provide thesis conclusions in Chapter 7.

Chapter 2

Using multi-wavelength data to determine host galaxy properties

In theory, theory and practice are the same. In practice, they are not.

Albert Einstein

GM Ben Finegold

2.1 COSMOS survey and multi-wavelength data

Astronomical surveys are incredibly useful for statistics-based studies. As a result of surveys having limited pre-selection criteria, the galaxies within a survey field are likely a more representative sample of the wider galaxy population than a targeted galaxy sample[1]. The galaxies within a survey field would hopefully have very similar (if not the same) properties of the wider galaxy population (effectively matched in morphology, environment, stellar mass, star formation rate and redshift). However, as different galaxies are bright at different wavelengths, our ability to detect those galaxies can introduce potential selection effects. This problem is somewhat (but not entirely) mitigated by

[1] Assuming that the Universe is similar in all directions and the survey has sufficient depth and width.

the availability of multi-wavelength data, as detections in one wavelength can be utilised to obtain upper limits in another (i.e., we know the galaxy exists because it is detected at one wavelength, but we know it is not luminous enough at others to have been detected). The inclusion of non-detections is important to ensure that any differences we detect arise from the wider population and are not influenced by selection biases.

For the data used in this thesis we start with the 2 degree^2 Cosmic Evolution Survey (COSMOS, Scoville et al., 2007). Primarily COSMOS is chosen for its sufficient depth, meaning we can probe higher redshift sources back to the epoch of peak SMBH and galaxy growth (i.e., $z \approx 1-2$, Madau & Dickinson 2014), but also its width, ensuring we can achieve adequate sample sizes. However, COSMOS also benefits from extensive coverage at various different wavelengths (instruments) such as ultraviolet (UV, *GALEX*), optical to near-infrared (Canada-France-Hawaii Telescope, *Subaru*, *VISTA*), near to mid-infrared (*Spitzer*) and far-infrared (FIR, *Herschel*). All such data has been compiled in the catalogue presented by Laigle et al. (2016, L16 from hereon in)[2]. In this thesis, we compliment the L16 catalogue with the *Chandra*-COSMOS-legacy survey Marchesi et al. (2016); Civano et al. (2016, C16 from hereon in), which provides us with additional data at X-ray wavelengths (which we use to measure SMBH accretion rates, see Section 2.2).

Throughout the work presented in this thesis, we need to repeatedly estimate three galaxy properties: SMBH accretion rate, SFR and host galaxy stellar mass. From the multi-wavelength data presented in L16 (matched with the *Chandra*-COSMOS-legacy survey), we use X-ray data to trace SMBH accretion, FIR data to estimate SFR and we fit UV-infrared spectral energy distributions (SEDs) to derive host galaxy stellar masses. In the rest of this chapter we discuss our reasons behind adopting these aforementioned processes to calculate the necessary host galaxy properties. However, as there are slight differences in the SFR calculations between Chapter 3 and Chapters 4 and 5, we reserve specific details for the relevant chapters. It should also be noted that we do not review

[2]The full catalogue is accessible here.

all possible methods for estimating these three galaxy properties. As spectroscopy requires targeted observations, useful spectroscopic data for the entire COSMOS field is not readily available. Although we do briefly mention some spectroscopic techniques, preference is given to methods relying on photometric data as they are more readily available for the COSMOS survey (and probably deep surveys in general).

2.2 Measuring the SMBH accretion through AGN luminosity

For the purposes of this thesis, we need to accurately estimate the SMBH accretion rate. In this section, we discuss how AGN luminosity is proportional to the the accretion rate and how the AGN luminosity can be estimated from various wavelength measurements. We then explain why we choose to use X-rays for the studies presented in the thesis (for AGN detection reviews see LaMassa et al. 2010; Brandt & Alexander 2015; Padovani et al. 2017).

Before covering which wavelengths are most appropriate for estimating the luminosity of an AGN, we must understand why we are measuring the luminosity at all. As previously mentioned in Section 1.1, Salpeter (1964) first suggested that, during an accretion event, a vast amount of energy could be released as the infalling matter's gravitational potential energy is converted into radiation (the vast majority of which would be light). The luminosity of an AGN ($L_{\rm AGN}$) is therefore thought to be a direct function of mass accretion such that,

$$L_{\rm AGN} = \eta \dot{M} c^2, \tag{2.1}$$

where $\dot{M}$ is the mass accretion rate, c is the speed of light and η is the efficiency by which infalling mass is converted to light energy. Again, as previously mentioned η is

thought to be between $5 - 20\%$ (see Raimundo et al. 2012 and references therein). The efficiency is thought to differ between AGN systems depending on properties of both the black hole (such as spin) and the infalling matter (such as mass). However efficiency measurements for individual AGN are very difficult to accurately derive and estimating them often relies on indirect connections (e.g., Fabian & Iwasawa, 1999; Volonteri et al., 2005; King & Pringle, 2006; Berti & Volonteri, 2008; Raimundo & Fabian, 2009). Whilst the work in this thesis does not assume any given individual efficiency, we do hold the assumption that changes in AGN luminosity are driven by changes to the accretion rate, rather than the efficiency. This assumption does however apply to most works in the literature on the connection between AGN luminosity and star formation.

During an accretion event, an AGN can be intrinsically bright across the EM spectrum (e.g., X-ray, UV, optical, mid-infrared and radio). Whichever wavelength is used to accurately trace SMBH accretion must, however, attempt to meet three particular criteria:

1. The emission traced must be intrinsically ubiquitous to SMBH accretion, thus removing any intrinsic selection biases.

2. Ideally, the emission traced would be a direct function of accretion, removing as many secondary dependencies as possible (such as obscuration).

3. Contamination from the host galaxy would be minimal.

The first AGN luminosity tracer we consider are those in the UV-optical luminosity range. Photons are produced at these wavelengths in the accretion disk (e.g., Shakura & Sunyaev, 1973; Shields, 1978; Malkan & Sargent, 1982), which means they are likely ubiquitous to all accreting SMBHs, satisfying the first criterion. However, UV-optical emission is prone to dust obscuration from the surrounding dusty torus (see Figure 1.4). For this reason, rather than using the the UV-optical luminosity directly, it is far more

common that a correction to the [OIII] emission line luminosity be used to trace AGN luminosity at these wavelengths (e.g., Bassani et al., 1999; Heckman et al., 2005; Panessa et al., 2006; LaMassa et al., 2010). It is thought that this emission originates in the NLR and is therefore independent of torus-obscuration. Nonetheless, it would not be exempt to obscuration by dust in the NLR or host galaxy itself (e.g., Lamastra et al., 2009). The [OIII] luminosity also likely depends on the properties of the torus (i.e., covering factor, or how much light escapes beyond the torus) and properties of the NLR itself (Lamastra et al., 2009). Therefore, the correction from [OIII] luminosity (and indeed any other dust-obscured measure) can be largely uncertain (e.g., Heckman et al., 2004). Fundamentally, however, the problem of obscuration can cause a series of problems that are difficult to fully overcome and account for. This means the second criterion is not met by considering tracers in the UV-optical luminosity range as there is a large secondary-dependency on obscuration. With regards to the host galaxy (and the third criterion), significant contamination from stellar populations can occur at these wavelengths (more so in the optical as, aside from the AGN, only massive, recently formed stars can emit UV photons). And whilst it is likely true in the quasar-regime that the host galaxy emission is insignificant compared to the emission from the AGN, for the majority of AGNs it is likely the host galaxy has a significant contribution to the UV-optical luminosity. Therefore, without applying a likely-uncertain correction for the host galaxy contamination, the third criteria is not met.

As a result of obscuring thermally emitted photons, the dusty torus is likely heated to temperatures such that it radiates at MIR wavelengths ($1000 - 1900K$, Barvainis 1987; Suganuma et al. 2006). Excess MIR emission is, therefore, often associated with the presence of an AGN (e.g., Laurent et al., 2000; Lacy et al., 2004; Stern et al., 2005; Alonso-Herrero et al., 2006; Fiore et al., 2008; Georgantopoulos et al., 2008; Donley et al., 2012; Eisenhardt et al., 2012; Mateos et al., 2013). Many studies have found that the MIR can adequately sample both obscured and unobscured AGNs (e.g., Rowan-

Robinson et al., 2005; Martínez-Sansigre et al., 2006; Hickox et al., 2007; Horst et al., 2008; Hao et al., 2010, 2011; Lacy et al., 2015; Suh et al., 2019) even if there is an apparent discrepancy in the amount of MIR emission between the two (e.g., see the difference in SEDs in Ramos Almeida et al., 2011; Hickox et al., 2017). This suggests that the MIR could be used to trace accretion across a representative sample of all AGNs, satisfying the first criterion. MIR photons do not suffer the same levels of dust obscuration as UV-optical ones do, meaning they are also less prone to obscuration in the host galaxy. However, whilst not being significantly impacted by dust, the MIR emission from an AGN is not only a function of accretion. Properties of the torus, such as the covering factor, need to be carefully considered when converting MIR emission to AGN luminosity (e.g., Stalevski et al., 2016). Additionally, star formation in the host galaxy can heat intergalactic dust to similar temperature meaning emission from stellar-heated and AGN-heated dust is difficult to disentangle. This disentangling often requires SED fitting to accurately compute a MIR luminosity that relates to SMBH accretion, which can, for large samples, be relatively model dependent (see e.g., Fritz et al., 2006; Alonso-Herrero et al., 2011; Mullaney et al., 2011; Lira et al., 2013, for examples of infrared SED fitting). So whilst the MIR can be useful for identifying AGNs, and can compliment obscuration-dependent techniques extremely well, for constraining the precise accretion rate they can be fairly uncertain as the third criterion is not met.

Another prominent technique used to trace SMBH accretion rates is an AGN's X-ray luminosity. It is widely accepted that X-rays are produced during most (if not all) accretion events (e.g., Avni & Tananbaum, 1986; Brandt et al., 2000; Gibson et al., 2008; Brandt & Alexander, 2015) by a hot ($\sim 10^9 K$) corona that resides just above the accretion disk (see Figure 1.4). This corona upscatters UV photons that were originally emitted by the accretion disc up to X-ray energies (e.g., Haardt & Maraschi, 1991; Mushotzky et al., 1993; Done & Diaz Trigo, 2010; Gilfanov & Merloni, 2014; Fabian et al., 2015), meaning X-rays can be used to accurately probe the intrinsic accretion

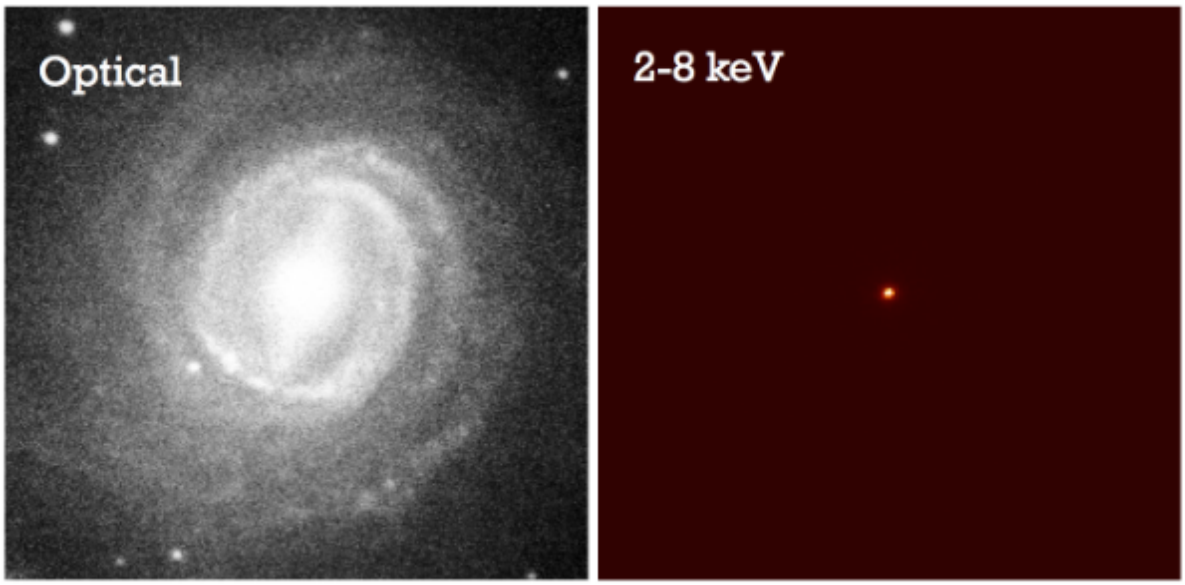


Figure 2.1: Optical and hard ($2 - 8$ keV) X-ray image of the local AGN NGC 3783. The image shows how the host galaxy, clearly visible in the optical, provides very little contamination at X-ray wavelengths. This figure was originally presented in Brandt & Alexander (2015).

rate of AGNs. Additionally, as shown in Figure 2.1, the host galaxy can be largely insignificant if there is an AGN, although this may change at the very lowest X-ray luminosities, (i.e., $L_{2-10\mathrm{keV}} = 10^{39-41}$ erg s^{-1}, see Aird et al. 2018). Indeed, some studies compare other AGN luminosity tracers (e.g., MIR) against the X-ray luminosity in order to measure their performance (e.g., Horst et al., 2008; Gandhi et al., 2009). So whilst the first and third criteria are met, using the X-ray luminosity to trace SMBH accretion does suffer from one serious drawback; X-ray photons can be obscured by large column densities of gas ($N_H = 10^{21-24.5}$ cm^2). The X-ray luminosity, therefore, depends largely upon two factors: the true SMBH accretion rate (which we are trying to estimate) and obscuration from gas in the AGN vicinity or the host galaxy. The problem of obscuration is partially overcome by only using harder X-rays (i.e., those with higher energies, > 2 keV), as they are considerably less prone to obscuration than softer X-rays (< 2 keV, see Lansbury et al., 2017, and references therein). Whilst the most heavily obscured AGNs may be missed (i.e., Compton-thick AGNs), we still believe X-rays are the most

reliable tracer of SMBH accretion. In an effort to overcome the uncertainties associated with obscuration (even at harder wavelengths), throughout this thesis we adopt the 2-10 keV absorption-correct X-ray luminosity from the *Chandra*-COSMOS-legacy survey (Marchesi et al., 2016; Civano et al., 2016) in order to trace SMBH accretion rate.

2.3 Measuring galaxy growth through star formation

In order to investigate the connection between SMBH accretion rate and host galaxy growth, we need to accurately estimate the rate of star formation in host galaxies. In a broad sense, star formation within a galaxy occurs during the gravitational collapse of its cold dense molecular gas clouds. This gravitational collapse usually occurs when the cloud is massive enough, but can also be instigatated by smaller regions reaching a critical density. Upon reaching this critical mass (or density), the gravitational force exceeds the forces supporting the gas cloud (which can be both motion and temperature based), the cloud collapses as a result of its own gravity, prompting star formation, and adding to the stellar mass of the galaxy. Star formation is, therefore, the process by which the galaxy grows. It should be noted this is a very simplistic view of star formation and the more intricate details can be found in the comprehensive reviews of Larson (2003); McKee & Ostriker (2007); Larson (2010); Kennicutt & Evans (2012); Girichidis et al. (2020) and the references therein.

As for estimating the level of star formation in the host galaxy, we utilise the fact that the most massive stars ($\geq$ 10 $M_{\odot}$) go supernova relatively soon after formation ($\leq$ 100 Myr), meaning a stellar population can be aged by determining the luminosity contribution from its massive-star subgroup (i.e., a young stellar population has a larger contribution from massive stars). This is the fundamental principle behind most of the

commonly used SFR indicators. However, in the forthcoming chapters, we go beyond just comparing AGN activity to the raw SFR. As the SFR is known to correlate with host galaxy mass and redshift, discovered connections could just be a byproduct of the relationship between AGN activity and stellar mass or redshift. In this section we first briefly compare the most common SFR indicators and then describe how we account for mass and redshift evolution in forthcoming chapters.

2.3.1 A summary of the SFR tracers

Before covering specific SFR tracers, it is worth highlighting a general cause of uncertainty when measuring the SFR of a galaxy using any tracer. As we try to estimate the abundance of massive stars, a conversion must be used to estimate the mass of both the massive and less-massive newly formed stars. This is often done by assuming an initial mass function (IMF), which describes the theoretical mass distribution of a population of recently formed stars. However, its precise form, accuracy and universality are widely debated (e.g Salpeter, 1955; Scalo, 1986; Kroupa, 2001; Chabrier, 2003; Bastian et al., 2010; Kroupa et al., 2013; Krumholz, 2014; Offner et al., 2014; Hopkins, 2018). Different assumptions about the IMF will almost certainly lead to different estimates of the level of star formation, even with the same measured flux. It is therefore important to consider, for comparing studies investigating SFRs, the choices of IMF. Throughout this thesis, we assume the IMF as described in Chabrier (2003).

As for tracing star formation, we consider a similar set of criteria that was mentioned in Section 2.2 regarding AGN luminosity tracers. That is, an ideal estimator of star-formation traces emission that is ubiquitous to recent star formation, with few, if possible, secondary dependencies. However, this time, the contamination from the AGN would ideally be negligible as well as minimal contamination from older, less massive stars. The first tracer we consider are those at UV wavelengths. In the absence of AGN emission,

the dominant source of UV photons in a galaxy are hot, massive stars, which immediately provides us with an intrinsic SFR tracer (e.g., Lequeux et al., 1981; Donas & Deharveng, 1984; Kennicutt, 1998; Gallagher et al., 1989; Madau et al., 1998). However, the UV luminosity is – as it was for tracing the SMBH accretion rate – both obscured by dust, and in the presence of unobscured AGN, likely dominated by the accreting SMBH. The UV luminosity can be, therefore, unreliable without significant assumptions about dust and AGN contamination, which can both be difficult to disentangle and often require comparison to other SFR estimators (e.g., Calzetti et al., 2000; Calzetti, 2001; Daddi et al., 2007; Salim et al., 2007; Kennicutt et al., 2009; Reddy et al., 2010; Wuyts et al., 2011). It is also likely unreliable to measure SFRs from optical photometry as it is difficult to disentangle emission from older and younger stellar populations. However, optical emission lines (such as Hα) can be used to trace the aforementioned ionising UV photons and are commonly used in local, non-AGN systems (e.g., Calzetti et al., 2004). It is also possible to estimate the SFR from SED fitting over the UV-MIR range (e.g. Aird et al., 2017, 2018). This is a common approach in studies that do not have sufficient FIR coverage. We do use SED fitting to derive our stellar masses (as we explain in Section 2.4) but choose to use the FIR for our SFRs as to remove any potential degeneracy of the stellar mass and SFR as result of them being derived using the same models. Our SED fitting code is also tailored to providing the most accurate stellar masses (when the performance of the code was tested against a mock sample in Ciesla et al. 2015) and fine tuning for the most accurate SFRs would mean slight altering of the SED fitting parameters.

As was the case with the torus obscuring the AGN, the obscuration of UV photons causes the obscuring interstellar dust to warm to temperatures such that photons are re-emitted at FIR wavelengths. FIR photons are far less sensitive to dust obscuration and can be relatively clean from AGN contamination, except in exceptional cases (e.g., Elvis et al., 1994; Richards et al., 2006; Schweitzer et al., 2006; Netzer et al., 2007;

Hatziminaoglou et al., 2010). Note that, as FIR emission only arises from dust-obscured star formation, a correction is still necessary for light that has not been obscured (and hence not detected at these wavelengths). The fact that a significant fraction of young starlight is obscured by dust (at least in star forming galaxies) is however, well established (e.g., Armus et al., 1987; Buat & Xu, 1996; Sanders & Mirabel, 1996; Goldader et al., 2002; Buat et al., 2005; Riguccini et al., 2011; Penner et al., 2012; Oteo et al., 2013). The most significant limitation for using FIR data to estimate SFRs, however, is the lack of sensitivity from FIR instruments, meaning we can only probe higher levels of star formation. For the studies in this thesis, we largely use measurements at FIR wavelengths in order to derive SFRs and as we include upper limits where possible, we can somewhat overcome the problems associated with poorer sensitivity. There are slight differences between the SFR calculations in different studies in this thesis so the specific calculations are fully explained in individual chapters.

2.3.2 A reliable star formation measure

Some studies that investigate AGN activity have found that the incidence of AGN increases out to higher redshifts (i.e., $z \approx 2$, see Aird et al. 2012; Bongiorno et al. 2016; Aird et al. 2018). This trend appears to hold for increasing stellar mass too, although care must be exercised for selection effects (e.g., Kauffmann et al., 2003; Best et al., 2005; Aird et al., 2012; Mullaney et al., 2012a; Aird et al., 2018; Kaviraj et al., 2019). However, the relationship between AGN prominence and stellar mass could be particularly complicated. Aird et al. (2018) noticed that, whilst the AGN detection probability changed as a function of stellar mass, is did so differently for star-forming galaxies when compared to quiescent ones. However, that AGN activity depends on redshift and stellar mass seems a widely accepted proposition.

The SFR of a galaxy, seemingly regardless of SFR tracer used, is also known to

generally (at least for star-forming galaxies) increase with stellar mass and redshift (e.g Brinchmann et al., 2004; Daddi et al., 2007; Elbaz et al., 2007; Noeske et al., 2007; Santini et al., 2009; Karim et al., 2011; Whitaker et al., 2012; Kashino et al., 2013; Steinhardt et al., 2014; Tomczak et al., 2016; Santini et al., 2017; Bisigello et al., 2018; Boogaard et al., 2018). This strong relationship, dubbed the star-forming main sequence, is shown in Figure 2.2 for a sample of 72,858 star-forming galaxies as originally presented in Schreiber et al. (2015). It is worth noting that there are certain populations of galaxies that appear to deviate from the main sequence. Firstly, a sub-population of galaxies appear to reside well below the main sequence (often called the red sequence) which have high stellar masses, but considerably less star formation than those galaxies that lie on the main sequence. The majority of galaxies appear to reside on the main sequence, or in this redder sequence, giving a bimodal star formation distribution. However, two smaller groups do exist. Firstly, a 'green valley' group connects the main sequence to the red galaxies with little star formation and secondly a series of starburst galaxies, that have considerably more star formation than that of the majority of star-forming galaxies. These groups are shown in the schematic in Figure 2.3.

As a result of both AGN activity and SFR being connected to the redshift and stellar mass of the host galaxy, it is important to attempt to control for these potential codependencies. To do this, throughout this thesis, rather than consider the raw SFR of our sources we consider the SFR relative to the main sequence shown in Figure 2.2 (which we herein refer to as the "starburstiness" or $R_{\rm MS}$), such that $R_{\rm MS}= \frac{\rm SFR}{\rm SFR_{MS}}$, where $\rm SFR_{MS}$ is the SFR on the main sequence for a galaxy with the same stellar mass and redshift. As per the prescription of Schreiber et al. (2015), $\rm SFR_{MS}$ can be calculated by using,

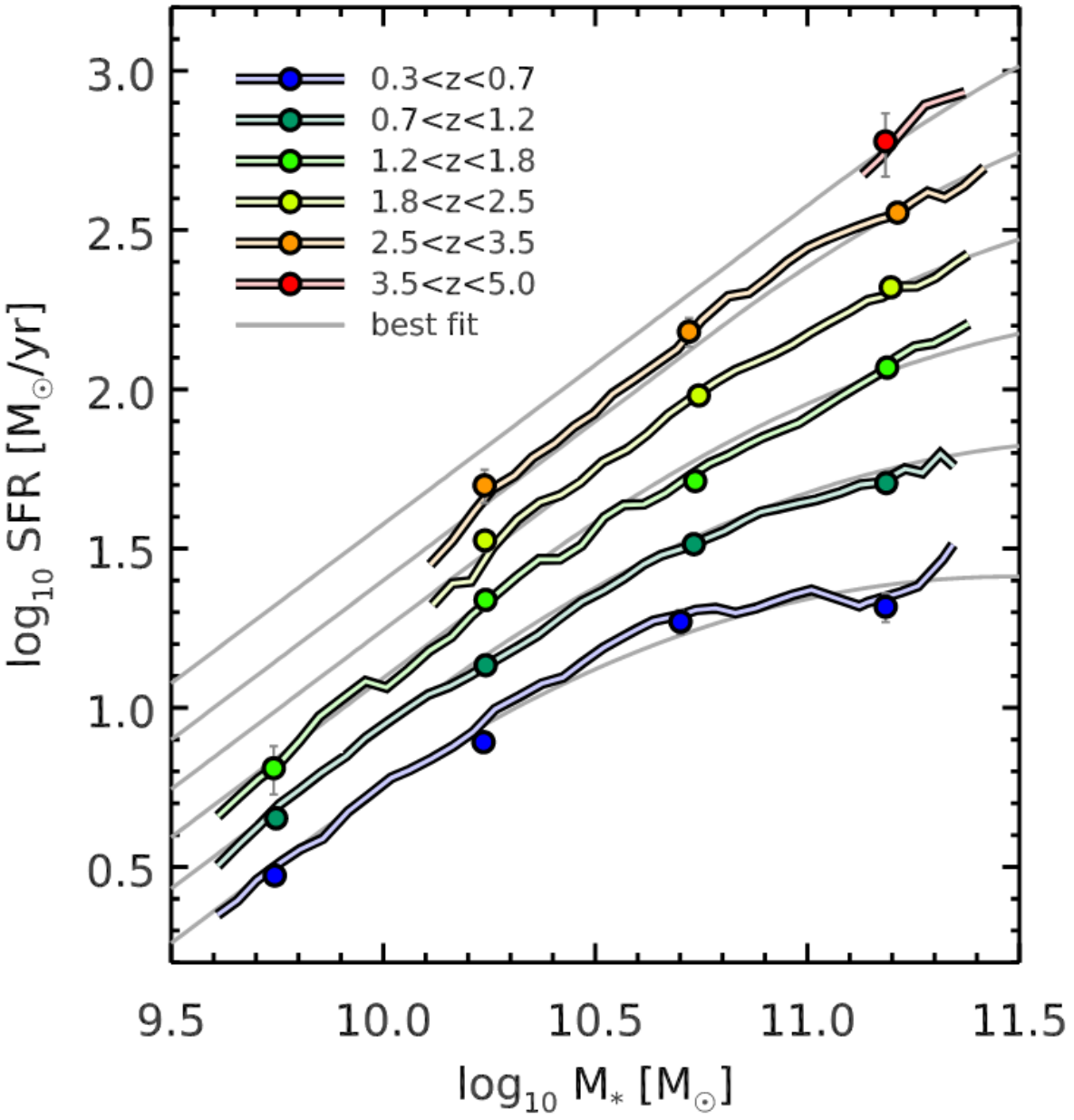


Figure 2.2: The average SFR of star-forming galaxies as a function of mass and redshift, calculated using by stacking of non-detections. Coloured lines represent a "continuous" measurement, which are calculated by changing the bin boundaries. The grey lines represent the a quadratic equation fit to the data, allowing for main sequence SFRs to be calculated. At higher redshifts, it is worth noting that the performance of the fit can not be judged, except at the highest stellar masses. This figure was originally presented in Schreiber et al. (2015).

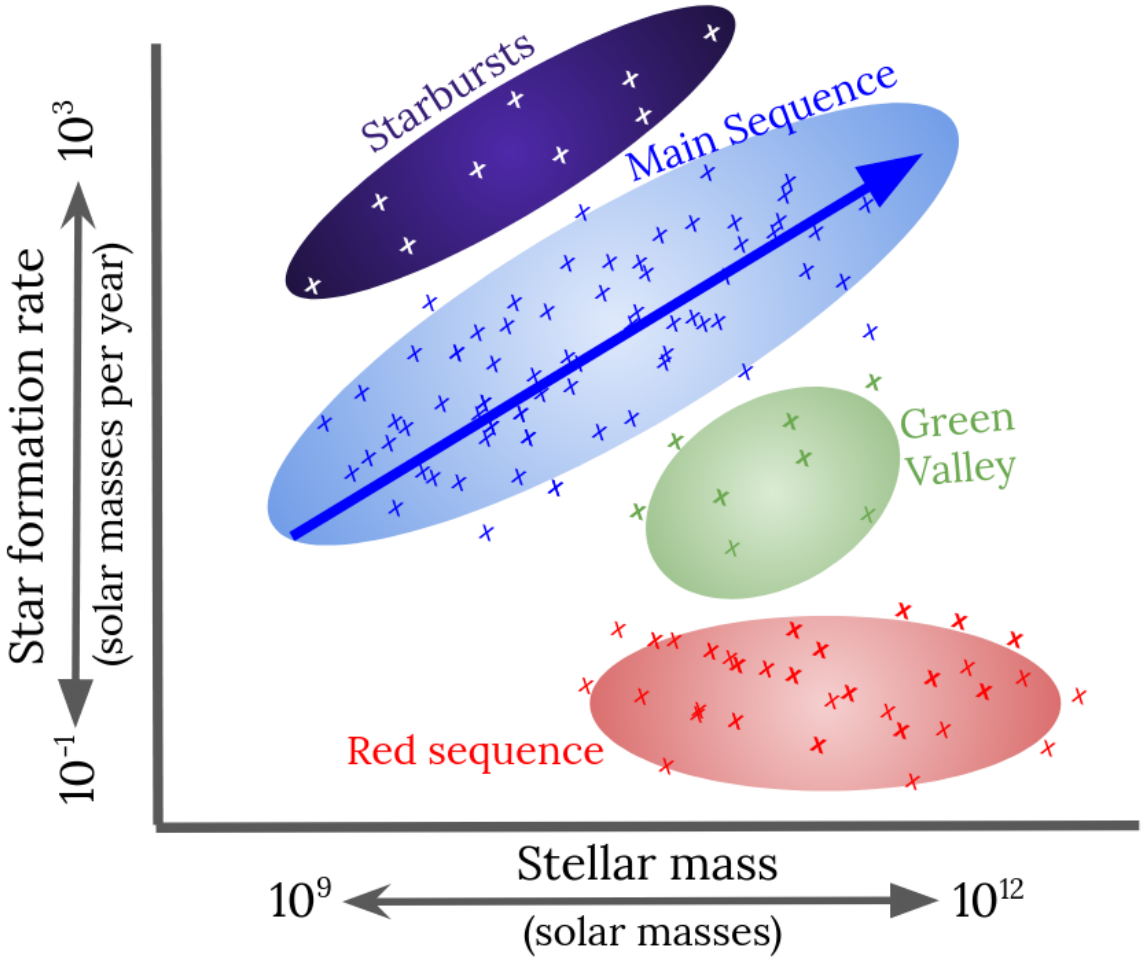


Figure 2.3: Sketch highlighting the star-forming groups of the galaxy population. The main sequence and starburst galaxies are often referred to as the star-forming galaxies, whereas the green valley and red sequence may be referred to as quiescent. The area of the ellipse nor the density of points are to scale with fractions residing in each group. Main sequence and red sequence galaxies form the vast majority of the population. Studies suggest up to 3% of galaxies may be starburst (e.g., Schreiber et al., 2015). For star-forming galaxies, there is a strong connection between stellar mass and SFR, although this does not appear to hold for red sequence galaxies.

$$\log_{10}(\mathrm{SFR_{MS}}) = \log_{10}\left(\frac{M_*}{10^9}\right) - 0.5 + 1.5\log_{10}(1+z) - 0.3[\max(0, \log_{10}\left(\frac{M_*}{10^9}\right) - 0.36 - 2.5\log_{10}(1+z))]^2, \tag{2.2}$$

where $\mathrm{SFR_{MS}}$ is in units of $\mathrm{M_\odot yr^{-1}}$ and M_* is in units of $\mathrm{M_\odot}$. R_{MS} can therefore be thought of as the excess (or discrepancy) in star formation that is not explained by redshift and stellar mass. This approach allows us to derive a measure of star formation that is largely stellar mass and redshift independent, reducing the potential for host galaxy properties to contaminate results from studies comparing AGN activity to star formation. However, calculating the R_{MS} values obviously requires us to calculate stellar masses for large samples, which we cover in the next section.

2.4 Using SED fitting to derive stellar masses

The stellar mass of a galaxy represents the star formation history integrated throughout the lifetime of a galaxy. A lot of studies in extragalactic astronomy carefully consider the stellar mass of the galaxy sample, as a range of galaxy properties correlate with stellar mass. In an attempt to mitigate the potential effects of stellar mass on their results, a large number of statistical studies either use mass-controlled samples or divide their results in to stellar mass bins. The studies in this thesis need to show similar levels of caution and therefore we need to accurately derive stellar masses.

Deriving the stellar mass for a large sample of galaxies is most commonly achieved by investigating the UV-FIR SED of the host galaxies; a method which has proved relatively successful when applied to simulated samples (e.g., Wuyts et al., 2009; Ciesla et al., 2015; Hayward & Smith, 2015; Mobasher et al., 2015; Torrey et al., 2015; Price et al., 2017;

Laigle et al., 2019). The (simplified) idea behind SED fitting is that a population of stars is generated which, when the light from individual stars is combined, accurately matches the observed SED. In reality, this is far less simple for two reasons. Firstly, the light emitted from a stellar population depends on factors such as the mass distribution, the metallicity and the age of stars and secondly, the light emitted is modified by a series of host galaxy effects, such as dust obscuration and emission. SED modelling is, therefore, a two stage process. Firstly, simple stellar populations (SSPs) are generated, in which all the stars are assumed to have the same age and metallicity. The light emitted from an SSP can then be calculated by integrating, over an age-altered IMF, the modelled emission from stars of a given mass, allowing a combined SSP spectrum to be calculated (for a detailed discussion of uncertainties within this process see Walcher et al., 2011, and references therein). During the first part of the SED fitting routine (usually referred to as stellar population synthesis, SPS), many SSPs are generated that adequately sample the input parameter space (i.e., for different combinations of metallicity and age-altered mass functions). SPS attempts, therefore, to provide a bank of intrinsic galaxy SEDs that can be used to fit observed galaxy SEDs. The second part of the SED fitting routine is to modify these intrinsic SEDs, by assuming dust attenuation and dust emission models, to attempt to replicate the observed SED. AGN emission can also be included in this part of the SED fitting process.

In reality, SED modelling is a difficult and complicated process. Usually, there is a compromise made between excessive model parameters and realistic modelling of the galaxy emission. Nonetheless for a large sample, SED fitting is still regarded as the current best approach for stellar mass calculations of large samples of galaxies with multi-wavelength data. Whilst a series of SED fitting codes are available (all specialising in varying areas of uncertainty), we use Code Investigating GaLaxy Emission (CIGALE, Noll et al. 2009; Serra et al. 2011; Roehlly et al. 2014). CIGALE has the ability to account for AGN contribution by including the AGN emission models presented by Fritz

et al. (2006), which helps to disentangle AGN emission from the host galaxy stellar population. Ciesla et al. (2015) studied the ability of CIGALE to reproduce the stellar masses of a mock sample of galaxies and reported that, in the presence of an AGN, the predicted stellar masses were in reasonable agreement with the input. More specifically, the three leftmost plots of Figure 11 in Ciesla et al. (2015) highlight the performance of CIGALE for varying quantities of photometric data. Generally, CIGALE performed well in terms of measuring stellar masses (within 40% of the input, with no systematic offset) when given photometric data from across the spectrum. We used CIGALE to derive stellar masses for all our sources, throughout this thesis, irrespective of whether they were previously identified (in the X-rays) as hosting AGNs, so as to mitigate a calculation bias. As L16 report photometric data ranging from the far-UV through to the far-IR, we are confident that we have sufficient data to determine the stellar masses for the sources in the samples used throughout this thesis. The range of possible parameter values that we used for the CIGALE run are shown in Table 2.1. These values are chosen as they were found to be the most successful for reproducing stellar masses in Ciesla et al. (2015) and are the same values chosen by Bernhard et al. (2016), who highlighted a strong correlation between the masses calculated by CIGALE and those in L16.

Modules	
Process/Model	**Theoretical Model Chosen**
Star Formation History	Delayed
Stellar Population Synthesis Model	Bruzual & Charlot (2003)
Dust Attenuation	Calzetti (2001)
Dust Emission	Dale et al. (2014)
AGN	Fritz et al. (2006)
Initial Mass Function	Chabrier (2003)
Parameter Values	
Star Formation History	
e-folding time of the main stellar population model (Myr)	100,1000,3000,10000,1E10
Age of the oldest stars in the galaxy (Myr)	100,1000,2000,3000,4000,5000,6000,7000,8000,9000,10000,11000
Stellar Population Synthesis Model	
Metallicity	0.02
Separation Age (Myr)	10
Dust Attenuation	
E(B-V)* for the old population	0.01,0.05,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0,1.1,1.2,1.4
E(B-V)* reduction factor of the old population	0.44
Central wavelength of the UV bump (nm)	217.5
Width (FWHM) of the UV bump(nm)	35.0
Slope of dust attenuation power law	0.0, 0.25, 0.5, 0.75
AGN	
Alpha Slope	1.5,2.5
Ratio of the maximum to minimum radii of the torus	60
Tau	1.0,6
Beta	-0.5
Gamma	0
Full opening angle of the torus(degrees)	100
Angle between equatorial axis and line of sight(degrees)	0.001, 89.990
Fraction of L_{IR} from AGN	0,**0.05,0.1,0.2,0.3,0.4,0.5,0.7,0.9**

Table 2.1: The various different modules used and the possible parameter values input into the CIGALE SED fitting code to derive host galaxy properties. Note CIGALE was run differently for those sources with AGN detections. The extra possible parameter values for the AGN run are shown in bold.

Chapter 3

Revealing the differences in the SMBH accretion rate distributions of starburst and non-starburst galaxies.

> Somewhere, something incredible is waiting to be known.
>
> Carl Sagan

3.1 Introduction

The finding that the average AGN luminosity increases with SFR (e.g., Rafferty et al., 2011; Rosario et al., 2012; Chen et al., 2013; Azadi et al., 2015; Delvecchio et al., 2015; Harris et al., 2016; Lanzuisi et al., 2017; Shimizu et al., 2017; Stemo et al., 2020) implies that the distribution of AGN luminosity changes as a function of the star-forming properties of the host galaxy. However, averages give little insights into the full shape

of these distributions. For example, does a sample have a higher average AGN luminosity because each AGN is slightly more luminous, or is it due to a small number of extreme, high luminosity AGN pulling the average up? Addressing such questions will provide a deeper understanding of the relationship between SMBH growth and galaxy growth: is the heightened average in star-forming galaxies caused by a slight increase in the activity of all AGN or a greater fraction of extreme cases? A direct way of addressing this is to determine how the AGN luminosity distribution changes as a function of the star-forming properties of their host galaxies. This has been explored in some recent studies (e.g., Aird et al., 2012; Azadi et al., 2015; Wang et al., 2017; Aird et al., 2017) who used rest frame optical to near infrared colours or SED fitting routines to identify samples of star-forming and quiescent galaxies and determined the differences in the stellar mass specific AGN X-ray luminosity. In general, these studies report a suppression of AGN activity in quiescent galaxies, particularly at modest specific AGN luminosities. However, in light of the difficulties associated with SFR estimates derived from optical wavelengths as covered in Section 2.3, it has yet to be determined whether these results are also observed when using FIR-derived SFRs.

In order to measure the SMBH growth relative to the size of the SMBH and remain consistent with the aforementioned works, in this chapter we analyse the specific X-ray luminosity (i.e., $sL_{\rm X} = L_{\rm X}/M_*$) distributions. As aforementioned, s$L_{\rm X}$ is an appropriate tracer of the Eddington ratio ($\lambda_{\rm Edd}$), which is given by,

$$\lambda_{\rm Edd} = \frac{L_{\rm AGN}}{L_{\rm Edd}}, \tag{3.1}$$

where $L_{\rm Edd}$ is the Eddington luminosity. $L_{\rm Edd}$ corresponds to (assuming a constant efficiency) the theoretical maximum luminosity (driven by the theoretical maximum accretion rate) after which the radiation pressure exceeds the inwards gravitational force and thus accretion would be self regulated. The more massive the SMBH, the stronger

the inwards gravitational force and therefore the higher the theoretical maximum accretion rate. As $L_{\rm Edd}$ is therefore a function of SMBH mass, the s$L_{\rm X}$ can be converted to the λ_{rmEdd} by assuming conversion factors: firstly converting $L_{\rm X}$to $L_{\rm AGN}$ and secondly stellar mass to SMBH mass. As these conversions can be highly uncertain the use of s$L_{\rm X}$can help overcome the need to apply uncertain correction factors.

In this chapter, we measure the full (i.e., including detected and undetected sources) s$L_{\rm X}$ distributions of galaxies whose star-forming properties have been measured from FIR data. We then compare these distributions of starbursting galaxies (defined by their $R_{\rm MS}$ values, see Section 2.3.2) against non-starbursting galaxies. We cover our specific sample derivation in Section 3.2. To measure the AGN luminosity distributions we construct a flexible model (see Section 3.3) that allows for both a power law style distribution (with lower and upper exponential turnovers) and a distribution that is more log-normally shaped allowing the data to determine which is more appropriate. Finally, we present the complete results and potential explanations in section 3.4 and possible implications and caveats in section 3.5. Throughout this Chapter we assume a 6-parameter ΛCDM cosmological model, with parameter values best inferred from the WMAP 9-year observations (Hinshaw et al., 2013). We must assume a cosmology for estimating distances when converting between intrinsic flux limits and luminosities.

3.2 Data

We start this section by summarising the process by which we derived our final sample of galaxies before elaborating on the specific calculations in the subsections that follow. Note that, stellar mass calculations were covered in more detail in Section 2.4.

3.2.1 Sample selection

As discussed in Section 2.1, we use the COSMOS survey to derive our samples (Scoville et al., 2007). In this instance, in order to measure the AGN luminosity distributions, it is important that we have as clean and unbiased a sample as possible. This is most easily obtained by using blank field surveys like COSMOS. In addition we also require a large sample, to avoid suffering from small sample size statistics, and comprehensive multi-wavelength coverage (for deriving stellar masses and SFRs). In particular, we also require good X-ray coverage as this provides, arguably, the most uncontaminated measure of AGN luminosity (see Section 2.2 and Brandt & Alexander, 2015). These requirements are well-met by the COSMOS survey, making it a natural choice for this study.

Our sample selection starts with the catalogue presented by L16, which contains photometric data for 1,182,108 sources in the COSMOS field. We supplement this with X-ray data from the catalogue presented by C16, which contains X-ray fluxes from *Chandra* for 4016 sources. We then apply the following steps to derive our final sample:

1. Firstly we ensure that the redshifts between L16 and C16 are consistent. We start with the photometric redshifts presented in L16 for all our sources as default. Then, for those sources present in C16, we adopt the "best" (i.e., spectroscopic if present, otherwise photometric) redshift presented in C16 (of which 1,981 are spectroscopic and 1,307 are photometric). We adopt the C16 redshift to ensure that we can use their derived X-ray luminosities in our analysis. Next, we select galaxies in the redshift range $0.05 \leq z < 2.5$, leaving 783,028 sources. This redshift range includes the vast majority of detections in the *Herschel PEP* survey, as the detection fraction drops off considerably at redshifts greater than $z = 2.5$ (see Figure 12 from Lutz et al. 2011). Importantly, however, this redshift range spans the epoch during which the majority of SMBH and galaxy growth took place (Aird

et al., 2010; Delvecchio et al., 2014).

2. We then derive stellar masses for all our remaining sources by fitting their SED using CIGALE (see Section 2.4). To avoid introducing a bias we also recalculated stellar masses for all the remaining sources rather than use the stellar mass presented in L16 (with the AGN component switched off). We then select only those sources with $\log_{10}(M_*/M_\odot) \geq 10.5$ to ensure the sample is mass-complete across our entire redshift range. This leaves us with a sample containing 58,241 galaxies. In Figure 3.1 we show a comparison between our CIGALE calculated mass and the mass presented in L16 for 4,750 randomly selected non X-ray sources.

3. Next, we obtain 2-10 keV luminosities (or upper limits thereof) for the remaining sources. Where the source is present in C16, we adopt the luminosity (or upper limit) given in that catalogue. If the source is not detected we calculate a 2-10 keV luminosity upper limit using the sensitivity maps of the *Chandra*-legacy survey (Civano, priv. comm.). How these upper limits are calculated is fully explained in Section 3.2.2. Any of the 58,241 sources in our sample that are not covered in the sensitivity map are deemed to have insufficient X-ray data and thus removed, leaving a sample of 40,418 (of which 2,763 have a measured X-ray luminosity).

4. SFRs in this sample are calculated by fitting SED models on IR to radio photometry taken from Jin et al. (2018). The photometry catalogue is produced by a "super-deblending" technique (Liu et al., 2018), including de-confused photometry at MIPS/24μm, *Herschel*, SCUBA2, AzTEC and MAMBO wavelengths, supplemented by NIR Ks, IRAC (L16) and radio data (Smolčić et al., 2017; Daddi et al., 2017). We used the same SED fitting algorithms described in Liu et al. (2018), included AGN models of Mullaney et al. (2011) and the spectroscopic redshifts of C16 to ensure redshift consistency. We then classified the sources according to the starburstiness quantity as described in Section 2.3.2. This calculation is further

	$0.05 \leq z < 0.5$		$0.5 \leq z < 1.5$		$1.5 \leq z < 2.5$	
	Det	UL	Det	UL	Det	UL
Starburst	10	97	54	516	31	227
Non-starburst	90	1868	780	14299	461	7986

Table 3.1: The complete sample sizes for our study, split by redshift bin, starburst classification and whether the sources are X-ray detected or an upper limit on X-ray luminosity had been calculated.

explained in Section 3.2.3. Sources without radio or MIPS/24μm data are omitted as a radio or MIPS/24μm detection is required for the deblending routine. The non-detection at these wavelengths could indicate a lower SFR and such sources are, therefore more likely to be classified as non-starburst. Whilst we could include these sources in our analysis under this assumption, our non-starburst sample is already the larger of the two samples in all of our redshift bins sized and thus does not warrant the introduction of such an assumption. After removing those galaxies without radio or MIPS/24μm detections, our final sample size is 26,419.

5. Finally, in order to investigate any redshift evolution in our $sL_{\rm X}$ distributions we subset our sample into three redshift bins: $0.05 \leq z < 0.5$, $0.5 \leq z < 1.5$ and $1.5 \leq z < 2.5$. The number of detected and upper limits for each redshift bin can be seen in Table 3.1. In addition, Figure 3.2 shows the detected $sL_{\rm X}$ distribution for both the starburst and non-starburst samples for each redshift bin and the cumulative upper limit fraction.

3.2.2 X-ray luminosity upper limits

If we were to include only X-ray detected sources when measuring our $sL_{\rm X}$ distribution we would be introducing a significant selection bias in to our analysis. It is therefore vital that we include galaxies for which we do not have an X-ray detection by calculating upper limits on their specific X-ray luminosity which we can then include in our maximum-

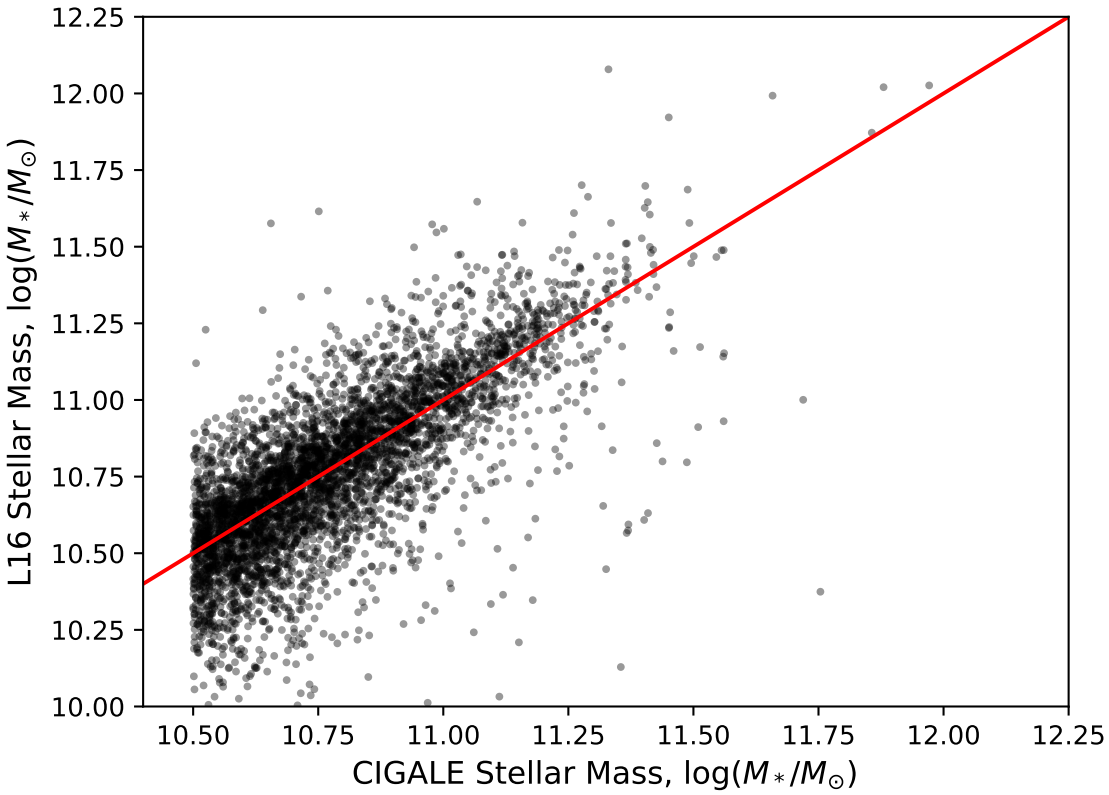


Figure 3.1: Stellar masses presented in L16 compared with the CIGALE derived stellar mass for a sample of 4750 randomly chosen non-AGN sources. The red line corresponds to the one-to-one case. Despite choosing CIGALE so that we can more accurately include the AGN component into the SED modelling, we choose to recalculate all stellar masses using CIGALE (including those without X-ray detected AGN) to mitigate potential systematics. The masses are, however, in good agreement when compared to those derived in L16 using alternative SED fitting codes.

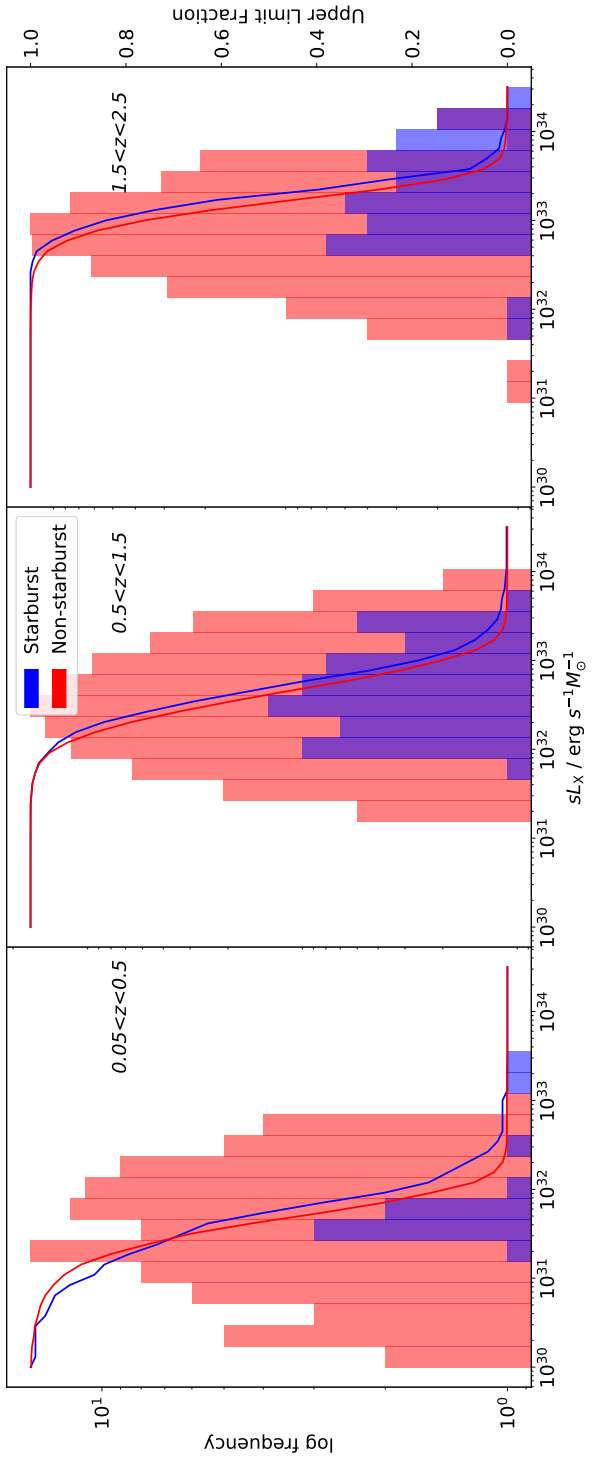


Figure 3.2: The detected sL_X distribution for the starburst (blue histogram) and non-starburst (red histogram) samples. Also shown is the cumulative upper limit fraction for the starburst (blue line) and non-starburst (red line). This illustrates where information about the true distribution is likely to come from (i.e., whether predominantly from the detections or non-detections).

likelihood analysis (see Section 3.3.3).

To calculate upper limits on the X-ray luminosities of our sources, we use the 2-10 keV sensitivity map of the *Chandra*-legacy survey (F. Civano, priv. comm.). This provides 3σ flux upper limits across the whole X-ray coverage of the survey. As such, to obtain flux limits for our non-X-ray detected galaxies we simply extract the flux limit at the position of that galaxy. This corresponds to an *observed* flux limit, whereas for our analysis, we require an intrinsic flux limit that attempts to account for any obscuration due to gas and dust. For detected sources we can use the hard (2-10 keV) to soft (0.5-2 keV) flux ratio to estimate the level of obscuration. This cannot, however, be done for undetected sources so for those we assume an average flux ratio calculated from the detected sources of $Q = 1.13$. We acknowledge the possibility that the undetected sources may have a higher level of obscuration than detected sources. However, the distribution of hard to soft flux ratios (for detected sources) is positively skewed. Therefore, the mean is shifted to higher levels of obscuration when compared to the median (0.74) or mode (0.53) meaning that the mean value we assume is conservative. In addition, we note there was no significant effect on our results when adopting an even higher obscuration level (e.g, $Q = 2$). We then use the following equation to obtain an upper limit on the intrinsic flux based on the upper limit on the observed flux (see Bernhard et al. 2016):

$$\log_{10}\left(\frac{F_{\rm I}}{F_{\rm O}}\right) = \sum_{i=0}^{2} a_i \log_{10}(Q)^i + b_i z^i, \tag{3.2}$$

where $F_{\rm I}$ is the intrinsic flux, $F_{\rm O}$ is the observed flux (i.e., the flux limit) and Q is the average flux ratio from the detected sources, i.e., $Q = 1.13$. Fitting this polynomial on the sample derived in Brightman et al. (2014), Bernhard et al. (2016) found the best fitting coefficients were given by $(a_0,\ a_1,\ a_2,\ b_1,\ b_2,\ b_3) = (0.23,\ 0.61,\ 0.041,\ 0.01,\ -0.11,\ -0.02)$, and we adopt these values. The need for a redshift term in this polynomial is driven by the finding of Brightman et al. (2014) that for a fixed luminosity, the level of obscura-

tion changed with redshift. More specifically, the covering factor of the torus increased, which provides an increased chance of additional obscuration. After calculating an upper limit on $F_{\rm I}$, we then use our adopted redshifts to calculate an upper limit on 2-10 keV luminosities, adopting a conversion of

$$L_x = F_{\rm I} 4\pi D^2 (1+z)^{2-\Gamma}, \tag{3.3}$$

where $\Gamma = 1.8$ is the assumed averaged intrinsic photon index (Burlon et al., 2011).

There are 17,823 galaxies that have insufficient X-ray coverage to calculate a meaningful X-ray upper limit. These are removed from the 58,241 that make up our mass-complete sample leaving 40,418 galaxies, of which 2,763 have a detected X-ray luminosity (the rest have upper limits on X-ray luminosity).

3.2.3 Calculating Starburstiness

Before we can derive and compare the specific X-ray luminosity distributions we need to divide our sample based on their star-forming properties. In order to derive SFRs for our sample in this Chapter we use the catalogue provided by Jin et al. (2018), which provides FIR-based SFRs for the COSMOS field. Jin et al. (2018) adopt a similar deblending routine as that presented in Liu et al. (2018). We use a positional match to identify counterparts in the SFR catalogue to the 40,418 galaxies for our mass-complete sample of galaxies. Since Jin et al. (2018) use mostly K-band positions as priors for their deblending we use a small matching radius of 1" to identify counterparts to that catalogue. From these SFRs, we calculate the $R_{\rm MS}$ statistic for our galaxies using the method outlined in Section 2.3.2. Each source in the sample is then classified as starburst if its $R_{\rm MS} > 3$, and non-starburst otherwise.

Since Jin et al. (2018) provide uncertainties on SFRs, we choose to discard any sources with ambiguous starburst status (i.e., those galaxies whose SFR error bars span

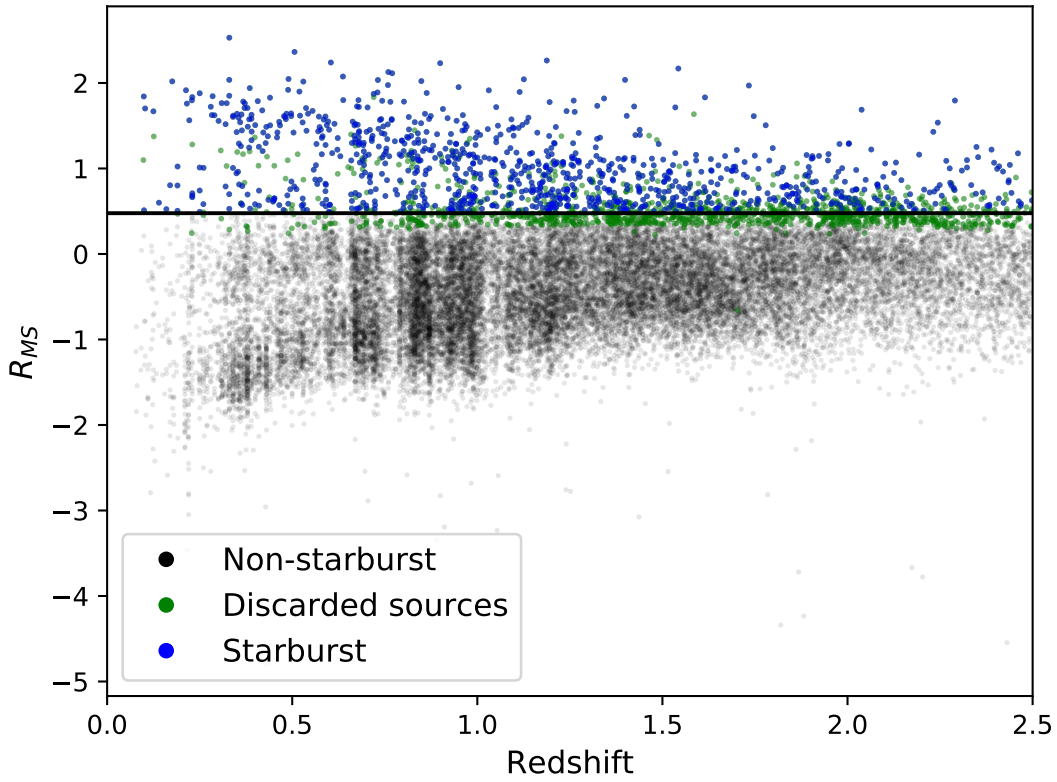


Figure 3.3: The distribution of specific star formation rate to main sequence (at equivalent mass and redshift) ratio (i.e., R_{MS}) as a function of redshift. Sources highlighted in blue are those selected as starburst. Sources in green have been discarded as their uncertainty on SFR estimate could introduce ambiguity into our classification. Including these sources would require a non-binning approach, which we introduce in Chapter 5.

the starburst divide). This prevents the unnecessary introduction of uncertainty. To accurately include information from those sources with ambiguous status a Bayesian hierarchical model would be required, in addition to an analysis without the limitation of binning on SFRs, i.e., an analysis that considers how the $sL_{\rm X}$ changes as a function of SFR, rather than between two bins. Such a model is introduced in Chapter 5. However, as a check, we tested what would happen should we include those sources with ambiguous status assigned based on their calculated starburstiness and noted that it did not have a significant impact on the results. We chose to omit them to minimise the number of potential misclassifications. Figure 3.3 shows the R_{MS} distribution for all our sources, with starburst sources highlighted in blue and discarded sources in green.

3.3 Constructing a flexible model

In this chapter, we aim to model the full $sL_{\rm X}$ distributions (i.e., including detected and undetected sources) of starburst and non-starburst galaxies in a range of different redshift bins. This section starts by describing how we construct a model that is able to incorporate information from undetected sources, whilst retaining the flexibility required to model the different functional forms the $sL_{\rm X}$ distribution may take. After describing the model, we also derive the likelihood function, from which we can infer the $sL_{\rm X}$ distributions by considering the maximum likelihood estimates of the parameters.

3.3.1 Model Selection

Constraining the precise form of the $sL_{\rm X}$ distribution (or its Eddington ratio equivalent) has been the focus of a number of recent studies (e.g., Aird et al., 2017; Bernhard et al., 2018; Aird et al., 2018). These works have suggested a number of different functional forms for the distribution. Currently, the three most popular functional forms are: a power-law with exponential cut-off (similar to a Schechter function, e.g., Hopkins et al.,

2009; Aird et al., 2012; Bongiorno et al., 2012; Hickox et al., 2014; Bernhard et al., 2016, 2018; Wang et al., 2017; Lanzuisi et al., 2017; Georgakakis et al., 2017), a log-normal distribution (e.g., Kauffmann & Heckman, 2009) or a so called "light-bulb" model (i.e., accretion is either on or off, e.g., Conroy & White, 2013). The difference in the observed shape of the distribution has recently, however, been attributed to selection effects with Jones et al. (2016) suggesting that after correcting for such effects a broad distribution is a good representation for the $sL_{\rm X}$ distribution of the AGN population. In this work, we also find that our samples are best modelled by a power-law with exponential cutoff. However, we develop and use a flexible probability distribution that retains the ability to recover both a power-law distribution and, if necessary, a log-normal-like distribution (see Figure 3.4).

In addition to the flexible nature of our model there are a number other criteria that would be desirable for a purpose-built probability distribution. Firstly, we must have a strict probability distribution (i.e., integrates to 1), which enables us to include information from upper limits using the likelihood function (see 3.3.2 for details). Secondly, for a power-law slope distribution, it is desirable to be able to control the power-law index, and the position of the low and high end exponential cut-offs. In the following subsections, we will describe how our model was built and how we included upper limits into this model.

3.3.2 Model construction

Following Aird et al. (2017), we choose to model our specific X-ray luminosity distributions as a sum of 40 unique Gamma distributions where a single Gamma distribution is described by the following equation:

$$Ga(X|\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \tag{3.4}$$

where α, β control the position and shape of the distribution and $\Gamma(\alpha)$ is a normalising constant. The mode of the Gamma distribution is given by $\frac{\alpha-1}{\beta}$. If α is fixed, the mode can be controlled by β. As such, a set of β values can be used to construct a series of equidistant Gamma distributions. If we then take the sum of these Gamma distributions, we recover a flat power-law distribution with lower and upper cut-offs, as seen in the upper-left plot of Figure 3.4. In particular, the minimum value of β controls the position of the left-most gamma distribution and the maximum value controls the mode of the right-most. Therefore, controlling the smallest and largest values for β allows us to control the positions of the turnovers in our model.

With the position of the lower and upper turnovers controlled by β the remaining parameter that we wish to control is the power-law slope. The power-law slope is controlled by the normalisation of the individual gamma distributions. Allocating each gamma distribution with parameter β a normalisation (i.e., a multiplicative constant) of β^{γ} produces a power-law distribution with a slope of γ (see Figure 3.4). The lower-left plot in Figure 3.4 illustrates how, if the minimum and maximum β parameters are close, the model has the ability to fit something similar to a log-normal distribution.

The above model provides us with the flexibility to construct a power-law distribution with appropriate turnovers. Importantly, in addition to this flexibility, summing gamma distributions allows us to easily include information from undetected sources by the incorporation of upper limits. To include upper limits in a likelihood function requires integrating the probability distribution. Using defined parametric distributions, such as the gamma distribution, allows the integrals to be quickly and easily calculated, eliminating the computation time and numerical uncertainties associated with the numerical integration that would be required if we assumed a standard power-law with cutoffs.

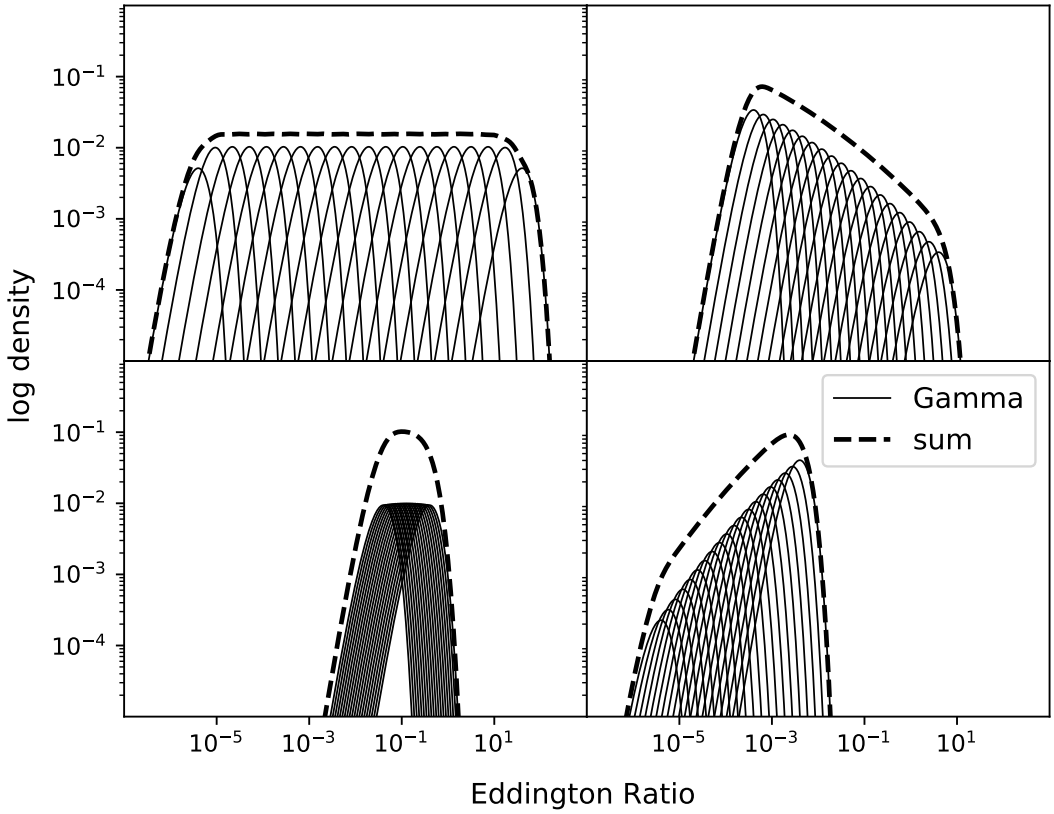


Figure 3.4: Examples of our model built by the summation of 20 independent gamma distributions (40 are used in the actual model for better accuracy). The parameters are as followed: The shape of each gamma distribution is fixed at $\alpha = 3$. *Top left:* $\gamma = 0$, $\log(\beta_{\min}) = -4$ and $\log(\beta_{\max}) = 1$, *Top right:* $\gamma = -1$, $\log(\beta_{\min}) = -5$ and $\log(\beta_{\max}) = 0$, *Bottom left:* $\gamma = 1$, $\log(\beta_{\min}) = -6$ and $\log(\beta_{\max}) = -1$, *Bottom right:* $\gamma = 0.1$, $\log(\beta_{\min}) = -3$ and $\log(\beta_{\max}) = 1$.

3.3.3 Likelihood Function

Now that we have a description for our model, we need to use our data to obtain the most likely parameter values for our model distributions (hereafter, the parameter values are collectively referred to as $\boldsymbol{\theta} = \{\beta_{\rm min}, \beta_{\rm max}, \gamma\}$). For a single X-ray detected galaxy, with $sL_X = X$, the likelihood is given by the probability density function (PDF),

$$f(X|\boldsymbol{\theta}) = \sum_{i=1}^{40} K\beta_i^{\gamma}\beta_i^{\alpha}X^{\alpha-1}e^{-\beta_i X}, \tag{3.5}$$

where K is a global normalisation constant.

For a sample of n X-ray *detected* galaxies the total likelihood can be written as the product of the PDFs, i.e.,

$$L(\theta|\mathbf{x}) = \prod_{i=1}^{n} f(X_i|\theta). \tag{3.6}$$

In our case, however, we have a large number of non-detections for which we have upper limits on their $sL_{\rm X}$. In such cases we must replace the PDF with the cumulative distribution function (CDF). That is, the PDF must be replaced by its integral evaluated up to the point of the upper limit. Mathematically, given data $\mathbf{x} = \{X_1, ..., X_m, X_{m+1}, ..., X_n\}$ where $\{X_1, ..., X_m\}$ are detected sources and $\{X_{m+1}, ..., X_n\}$ are upper limits, the likelihood function can now be expressed as,

$$L(\theta|\mathbf{x}) = \prod_{i=1}^{m} f(X_i|\theta) \prod_{i=m+1}^{n} \int_{-\infty}^{{\rm UL}_i} f(X_i|\theta)dX_i. \tag{3.7}$$

Given an $sL_{\rm X}$ value for each of the sources in our sample it is this likelihood equation that we seek to maximise. To incorporate uncertainties on the detected sources we calculate an error on the X-ray luminosity by calculating the relative error on the flux observed and propagating this through to the relative error on the luminosity (i.e., neglecting uncertainty on photo-z, for example). For each detected source we then replace

the absolute detected value with a randomly sampled value from a Gaussian distribution centred at the observed value with the aforementioned percentage uncertainties. We do this during each step of the maximisation process to accurately account for the uncertainties on sL_{X} throughout the analysis.

3.3.4 Likelihood maximisation

In Section 3.3.3, we derived the likelihood function for our parametric distribution. From here, we can determine which parameter values maximise the likelihood function by using the Markov-chain Monte Carlo (MCMC) Python package EMCEE (Foreman-Mackey et al., 2013). MCMC is required as the likelihood function is too complicated to maximise analytically.

We use MCMC methods to calculate posterior distributions of the parameters of our model, for each redshift bin and both the starburst and non-starburst sample. Our chains each have 200 walkers, each of which are run for 5000 steps (re-sampling the detected values from their uncertainty distributions), with the first 1000 removed for burn-in. This results in a posterior sample of size 800,000 for each parameter. We then choose to thin this sample by selecting every 200th value in the sample. Thinning is used to reduce the sample size to more manageable numbers but also removes the slight dependence between consecutive draws in the chain. On inspection, we noticed the chain converged much more rapidly than the applied burn-in so we are confident we are sampling the posterior parameter space.

3.4 Results

We start this section by presenting the output (i.e., the posterior distributions) from the MCMC algorithm. We then discuss the specific parameter results and their potential implications on the sL_{X} distributions for the starburst and non-starburst samples.

	Starburst	Non-starburst
$0.05 \leq z < 0.5$		
power law slope	-0.406 (-0.571, -0.275)	-0.857 (-0.944, -0.791)
low turnover	-5.161 (-6.01, -4.641)	-4.842 (-4.877, -4.734)
high turnover	-1.194 (-1.429, 0.016)	-1.610 (-1.929, -0.808)
$0.5 \leq z < 1.5$		
power law slope	-1.090 (-1.212, -0.900)	-1.203 (-1.248, -1.160)
low turnover	-3.138 (-3.257,-3.017)	-3.377 (-3.395, -3.328)
high turnover	-1.126 (-1.357, -0.332)	-0.965 (-1.051, -0.799)
$1.5 \leq < 2.5$		
power law slope	-0.902 (-1.077, -0.711)	-2.178 (-2.301, -2.084)
low turnover	-2.518 (-2.781, -2.389)	-2.303 (-2.332, -2.268)
high turnover	-0.051 (-0.314, 0.553)	-0.556 (-0.614, 0.608)

Table 3.2: Modes from the posterior distributions presented in Figures 3.5, 3.6 and 3.7. The errors, displayed in brackets, are the 68% highest posterior density intervals calculated using the *HPDInterval* package in R.

3.4.1 MCMC output

We present the burned-in, thinned, posterior distributions for the three redshift bins, $0.05 \leq z < 0.5, 0.5 \leq z < 1.5$ and $1.5 \leq z < 2.5$ for both starburst and non-starburst sources in Figures 3.5, 3.6 and 3.7, respectively. They show repeated MCMC draws from the posterior distribution of each parameter on the diagonal, as well as the 2D contour plots (shown because of the potential dependence between model parameters) on the off-diagonal, calculated using kernel density estimation (a non-parametric way of estimating a distribution from a histogram using smoothing). In this figure, as well as all further plots, the starburst sample is shown in blue, whereas the non-starburst sample is shown in red. Summary statistics from the posterior samples are shown in Table 3.2.

By randomly selecting from the posterior parameter values we can construct the range of possible $sL_{\rm X}$ distributions. This is shown in Figure 3.8, in which we highlight the median $sL_{\rm X}$ distributions including 1σ error regions, for the three redshifts bins. The errors are calculated by identifying the 16th and 84th percentiles at a given value of $sL_{\rm X}$ for all the sampled parameter values. In the following subsections we discuss, in more

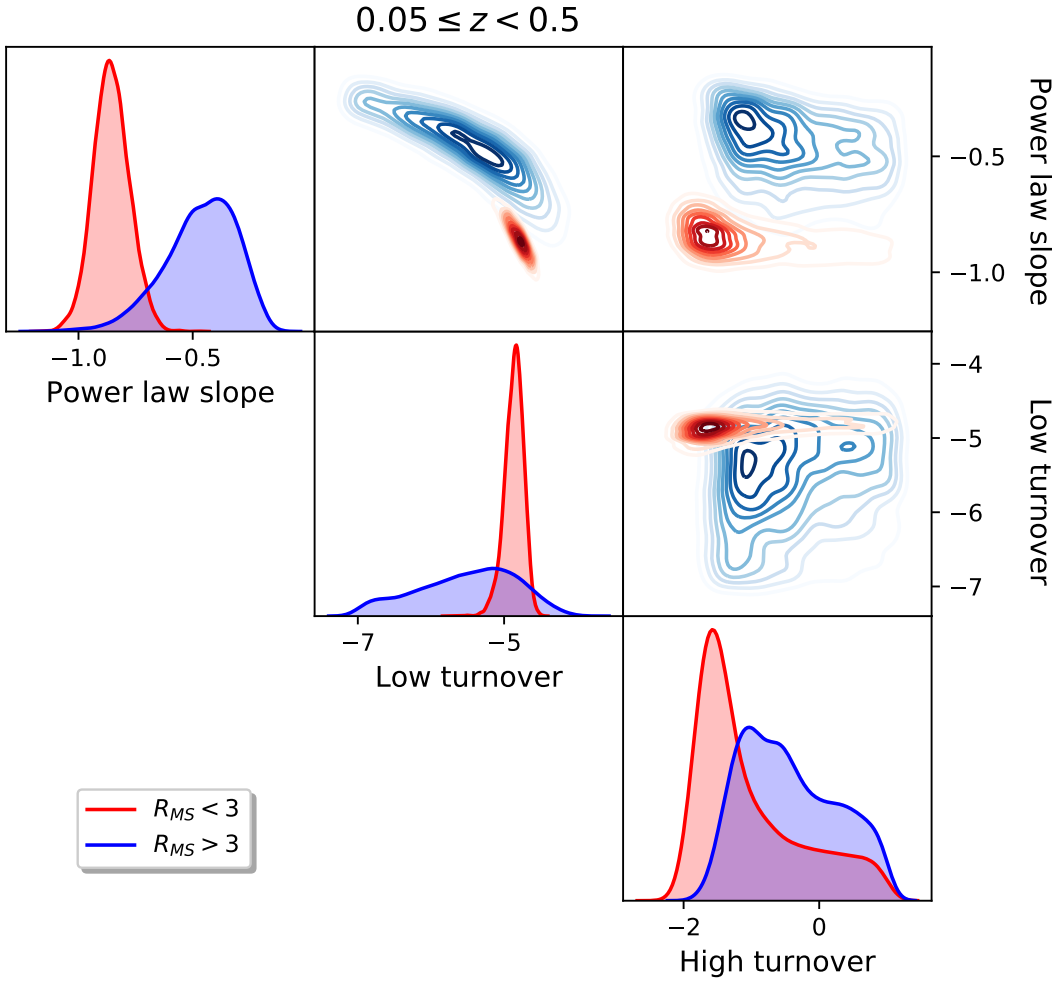


Figure 3.5: The posterior distributions (on diagonal) and the 2-D contour plots, drawn using a kernel density estimation technique for the redshift range $0.05 < z < 0.5$ split between starburst (blue) and non-starburst (red).

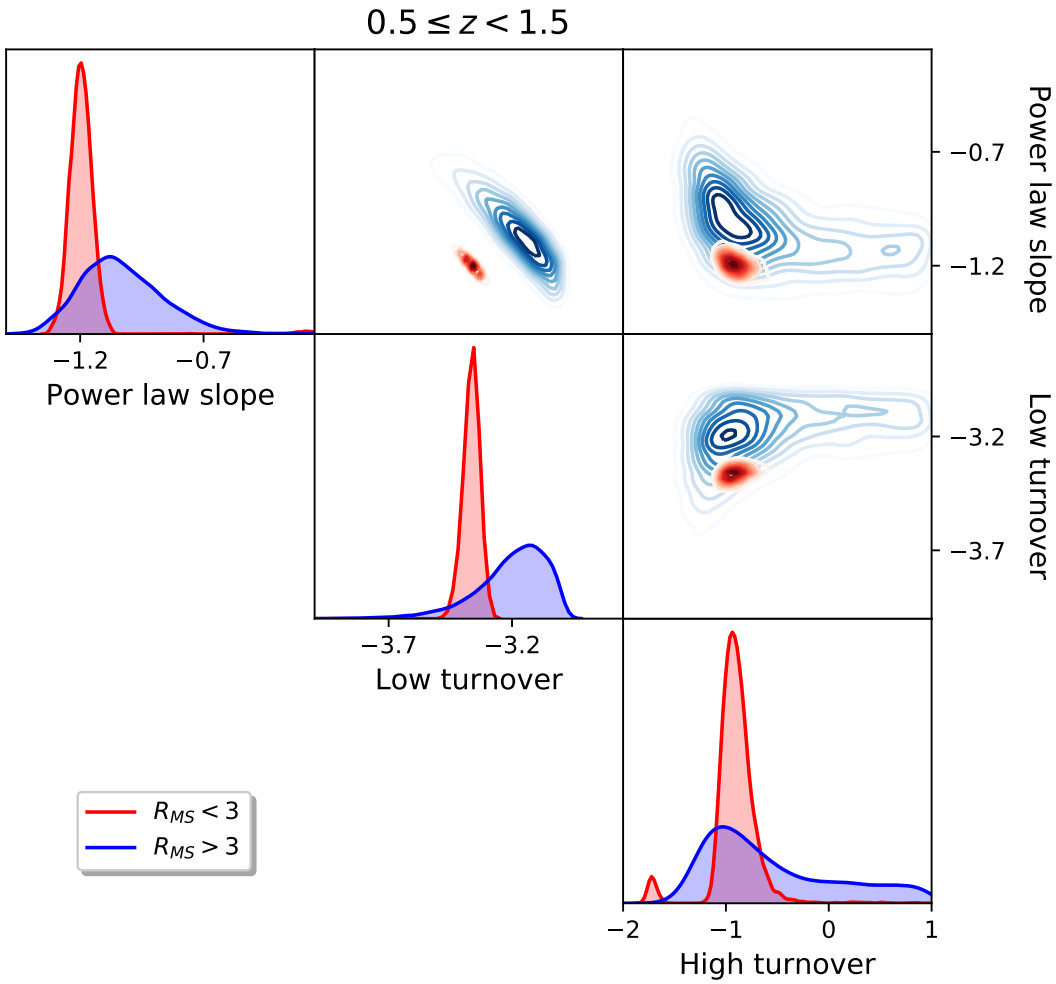


Figure 3.6: Same as Figure 3.5, but for the redshift range $0.5 < z < 1.5$.

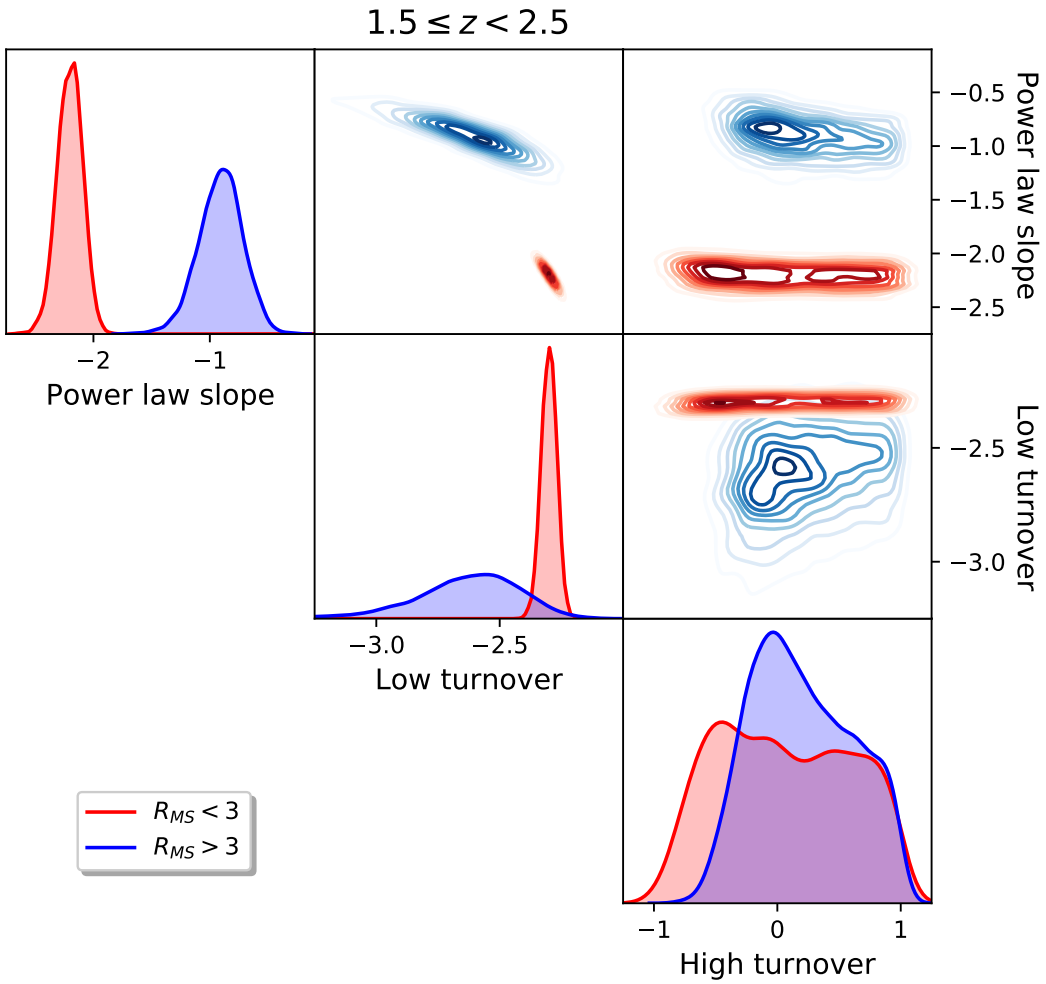


Figure 3.7: Same as Figure 3.5, but for the redshift range $1.5 < z < 2.5$.

detail, the differences between the parameter values for the two starburst samples and as a function of redshift. As is good statistical practice, the posterior distributions displayed in Figure 3.8 are only displayed between the range of the minimum and maximum values of detections.

3.4.2 Power law slope

The power law slope parameter controls the gradient of the model between the low and high exponential turnovers. The steepness of this slope could be indicative of the proportion of very luminous sources in the sample, because the slope largely controls the ratio of higher to lower $sL_{\rm X}$ sources (i.e., above and below the midpoint, respectively). From the posterior distributions presented in the upper-left plots of Figures 3.5, 3.6 and 3.7, we see consistently that the modes of the power law slope distribution are shifted to less negative values for the starburst samples in all three redshift bins. In the lowest, intermediate and highest redshift bins we can state that the power law slope in starburst galaxies is shallower than in non-starburst galaxies at a significance of 97.7%, 80.9% and 98.5% respectively. This could suggest that the proportion of higher $sL_{\rm X}$ sources is greater in the starburst population than the non-starburst population (as a result of having a higher ratio of high to low $sL_{\rm X}$ sources) and we explore this possibility further in Section 3.5.2. The difference in power law slope can also be seen in the full posterior $sL_{\rm X}$ distributions shown in Figure 3.8 with the gradients of the distributions prior to the break displaying a greatest difference in the high redshift bin.

3.4.3 High turnover

Whilst the power law slope indicates the ratio of high to low $sL_{\rm X}$ sources (above and below the midpoint), the high turnover controls the maximum possible values of $sL_{\rm X}$ in the model. From the posterior distributions presented in the lower-rightmost plots

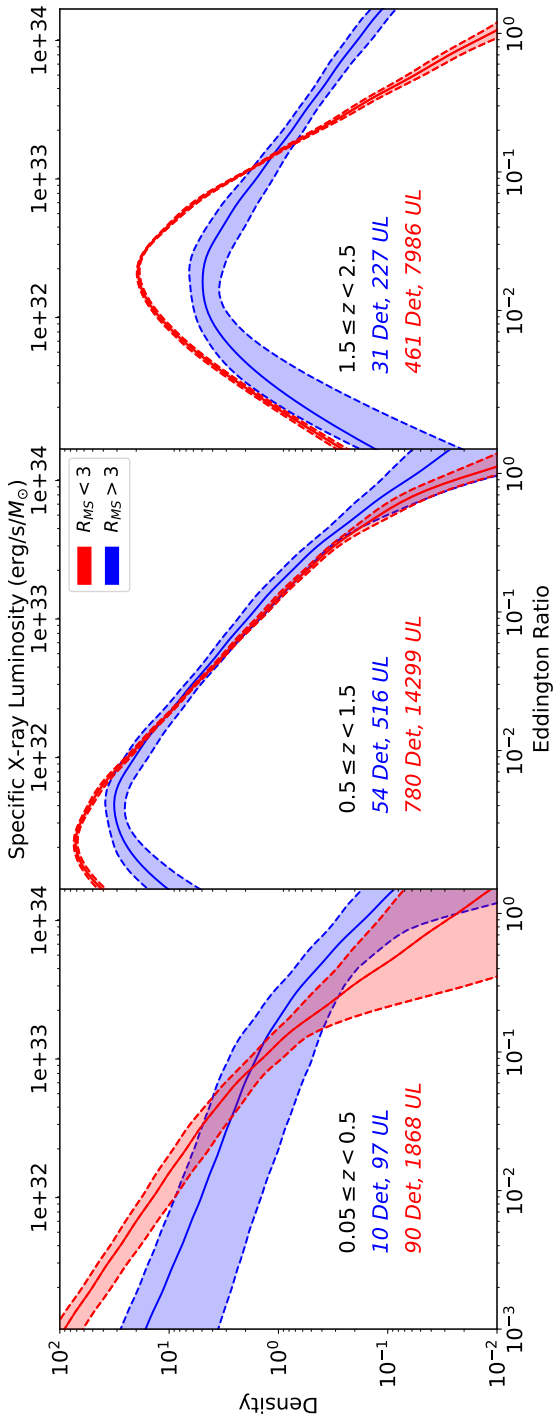


Figure 3.8: The full $sL_{\rm X}$ distributions inferred from our analysis for all three redshift bins. The 1σ error regions are shown by the shaded region (calculated by finding the 16th and 84th percentile at a fixed value of $sL_{\rm X}$). It should be stressed that these error regions are not errors on the whole distribution, rather on a given value for $sL_{\rm X}$. The starburst sample is shown in blue, while the non-starburst are shown in red. The sample sizes are also shown for reference.

of Figures 3.5, 3.6 and 3.7, we see that there is significant overlap between the high turnover distributions in both samples across all the redshift bins. We see a shift in the mode of the posterior distributions in our lowest and highest redshift bins. In addition to this, the high turnover posterior distributions are generally broader than those of the power law slope. We suspect that this is a consequence of this extreme end of the model being constrained by extremely luminous, extremely rare AGN and therefore the inferred posterior distribution is poorly constrained. Having said that, in the highest redshift bin, the significant difference in power-law slope and the inability to recover the high turnover accurately enough combines to create an excess of very high $sL_{\rm X}$ sources in the starburst sample, as shown in Figure 3.8. Therefore, at this high redshift we cannot rule out that SMBHs in starburst galaxies have the ability to accrete at higher maximum thresholds.

3.4.4 Parameter evolution with redshift

As previously mentioned, we subset our sample into three redshift bins to investigate how the various parameters describing our distributions evolve from a redshift of $z \approx 2.5$. In Figure 3.9 we show how the mode of the posterior distributions change for each parameter as a function of redshift. Figure 3.9 shows the mode of the posterior distributions for each parameter (power law slope, low turnover and high turnover in the left, middle and right plots, respectively) plotted against the midpoint of the redshift bin it was inferred from.

The leftmost plot in Figure 3.9 shows how the power law slope has evolved with redshift. This plot suggests that the power law slope for non-starbursts becomes more negative as we go to higher redshifts. As the power law slope may reflect the ratio of higher to lower (i.e., above and below the *mid-point*) $sL_{\rm X}$ sources, the apparent parameter evolution indicates that the proportion of higher $sL_{\rm X}$ sources in the non-starburst

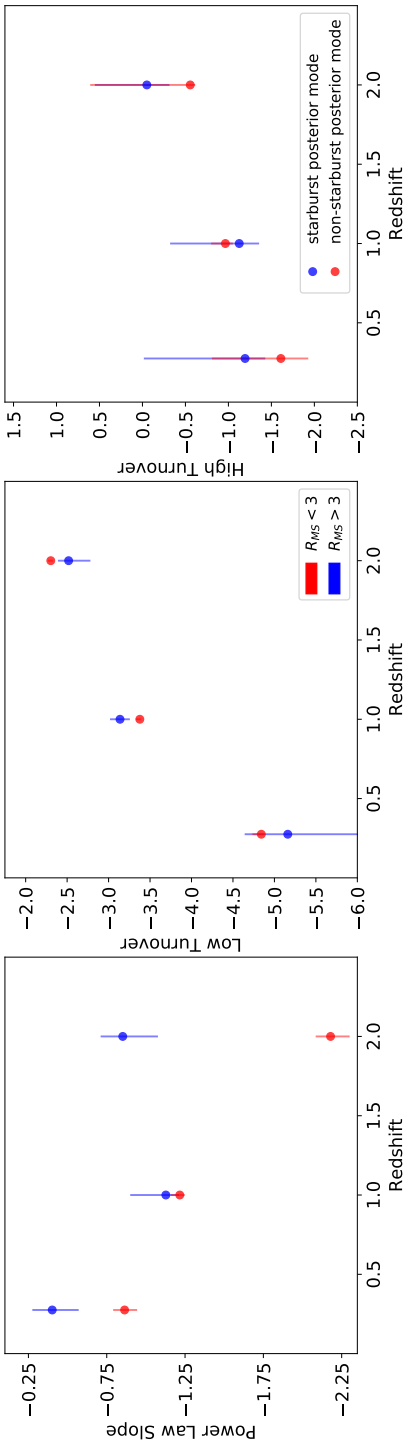


Figure 3.9: Parameter evolution plots for each redshift bin between the starburst (blue) and non-starbursts (red). The posterior mode for each parameter is plotted against the midpoint of the redshift bin it has been inferred from, along with 1σ uncertainties.

galaxy population may have also evolved with redshift. More specifically, as the power law slope has declined out to higher redshifts, the proportion of higher sL_{X} in AGN has also declined. In addition to this, the difference in parameter evolution between the two samples suggests that the proportion of high sL_{X} sources was higher in the starburst population than the non-starburst one, which indicates that a relationship between intense star formation and SMBH growth is likely to exist and that the evolution of this parameter is more dependent on starburst/non-starburst classification than redshift. However, considering how the low and high turnovers (indicating the range of sL_{X}) evolve alongside the power law slope will provide a more complete picture. The low turnover rapidly increases with redshift and whilst the high turnover evolution is poorly constrained (again, due to the rarity of sources at this end of the distribution), it does appear there may be a slight increase with redshift. Should this be the case, it would suggest that while the proportion of higher sL_{X} sources in the population decreases with redshift, the average sL_{X} increases. However, it is worth emphasising that the middle and right plots in Figure 3.9 do suggest that the difference in the low and high turnover between the accretion rate distributions is primarily driven by redshift and not starburstiness, whereas the left plot suggests a greater dependence on starburstiness. Moreover, as the low turnover effectively controls the normalisation (as the probability distribution must integrate to unity), the redshift evolution of the low turnover gives us insight into the normalisation of the distribution with redshift. More specifically, the normalisation increasing with redshift reaffirms the idea that the Universal accretion rate increases with redshift (therefore, we would theoretically expect the low turnover to decrease after $z \approx 2.5$ as we know Universal accretion peaked at $z \approx 2$). We explore the implications of this in Section 3.5.2.

3.5 Discussion

The primary goal of this study is to measure the differences, in the distributions of SMBH accretion rates for starbursting and non-starbursting galaxies, which may be able to help explain why the average SMBH accretion rate increases with SFR. We used the specific X-ray luminosity (i.e., $sL_{\rm X} = L_{\rm X}/M_*$) as a proxy for Eddington ratio and derived the SFR of our sources from *Herschel* FIR photometry via a deblending and SED fitting routine (see Jin et al. 2018 for details). Our sources were split according to their star-forming properties; if their star formation rate placed them a factor of three above the main sequence (using the prescription of Schreiber et al. 2015) they were classed as a starburst galaxy, otherwise they were classed as a non-starburst.

3.5.1 Assumptions and analysis limitations

In order to model the distribution of non-starburst and starburst galaxies as accurately as possible we constructed a flexible parametric model that was able to recover either of the two most popular forms of the $sL_{\rm X}$ distribution reported in the literature (see Section 3.3.1). However, the model is not without limitations and we acknowledge and discuss these further in this section.

Firstly, as with any parametric study, our analysis and interpretation of results are model dependent. A parametric form of the distribution must be assumed (in this case, a power-law with exponential cut-offs or log-normal) in order to account for information from both detections and non-detections. The aim of the study is then to derive the most-likely parameter values for a given model and compare those parameters between samples. From that, we can first pose the question: given our model, do the parameters that describe the underlying distributions differ *significantly* for our starburst and non-starburst samples? If so, then the *underlying* distributions differ. If they do differ, then we can also ask, given our model, how to they differ? It is important, however, to

consider the limitations of our (or any other) model, particularly when considering the latter question. For example, we acknowledge that our model is incapable of replicating the distribution found in Aird et al. (2017), who found a "bump" in the distribution at lower $L_{\rm X}$ values ($10^{39} - 10^{41}$erg s^{-1} depending on mass and redshift) that they attributed to star formation. As such, any differences in our inferred distributions could be as a result of a bump that we do not specifically model. However, were we to include a bump at lower $sL_{\rm X}$ values, it would likely cause the inferred power law slope of our starburst sample to flatten further (as upper limits would occupy the bump) which would strengthen the significance of our results.

Secondly, the data in this study contains a large fraction of non-detections. The reason for this is that we intend to infer our results on the entire galaxy population as opposed to only X-ray detected sources, as the latter would produce biased results. However, aside from the appeal of an unbiased sample, the non-detections do contain information about the underlying distribution. The CDF used in this analysis allows us to incorporate information from the non-detections by fully considering the possible values for them. In Figures 3.5, 3.6 and 3.7, one can see the power law slope and the low turnover are correlated. One possible reason for this is that initially, at the high $sL_{\rm X}$ end of the distribution, the power law slope is inferred from the detected sources and the model then computes whether enough upper limits are introduced to maintain this slope. This indicates that our model is sensitive to the fraction of upper limits in the analysis (the low turnover must occur at the point where upper limits are unlikely to be able to maintain the gradient of the most likely power law slope, which is inferred from the detections). As such, it is likely that the low-turnover at low Eddington ratios is a direct consequence of the combination of large numbers of upper limits in the data with our assumed model shape. This further stresses the importance of ensuring that we have a sample representative of the population with a proportionate fraction of non-detections and a justified choice of model.

As with any population study, it is extremely difficult to rule out all possible systematic effects that could influence our final results. We attempt to mitigate the effects of any unknown systematics by (a) treating starburst and non-starburst samples the same in terms converting X-ray fluxes to accretion rates and (b) comparing starbursts to non-starbursts within the same redshift bin and thus minimising the influence of, e.g., flux limits between the samples. Considering point (a) specifically: one could imagine that starburst galaxies have a higher level of absorption due to enhanced amounts nuclear gas introduced by galaxy interactions. If this were the case, then this would work to enhance the differences we see, as correcting for stronger absorption in starbursts would systematically increase the intrinsic $sL_{\rm X}$ we measure, leading to an even greater number of high $sL_{\rm X}$ AGN amongst starbursts.

3.5.2 Inferring the results

Figures 3.5, 3.6 and 3.7 suggest that the parameter with the largest difference between the starburst and non-starburst samples is the power law slope. Given our model, the probability that the accretion rate distribution for starburst galaxies has a less-negative power law slope in the lowest, middle and highest redshift bins are 97.7%, 80.9% and 98.5% respectively. While these differences are not significantly different at the 3σ level, a difference in the power law slope may indicate that the fraction of higher $sL_{\rm X}$ sources may be different between the starburst and non-starburst samples at a given value of specific X-ray luminosity or Eddington ratio. In order to investigate this further, we calculated the fraction of sources with "high" accretion rates (i.e., greater than $0.1\lambda_{\rm Edd}$) in both the starburst and the non-starburst posterior accretion rate distributions. This is calculated by integrating each of the 4000 posterior $sL_{\rm X}$ distributions for each sample from $0.1\lambda_{\rm Edd}$ upwards. These fractions are presented in Figure 3.10 and show that the starburst sample has a larger fraction of sources with high accretion rates across all

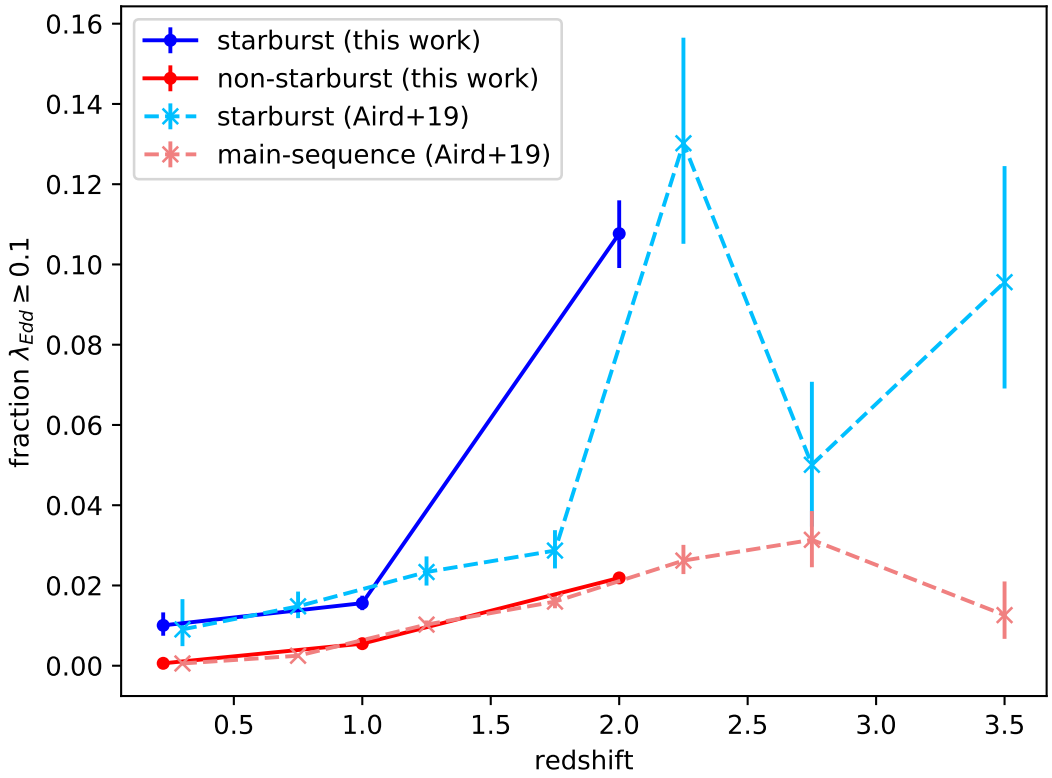


Figure 3.10: Fraction of sources with high accretion rates (i.e., greater than $0.1\lambda_{\rm Edd}$) as a function of redshift for the starburst and non-starburst samples. Uncertainties are 1σ and are calculated by selecting the 99.7% credible interval from the posterior $sL_{\rm X}$ distributions. Over-plotted are the starburst and main-sequence fractions from Aird et al. (2019) with 1σ uncertainties.

redshift bins. Also included in this plot are the ratios of high to low accretion rate AGN for starbursts and main-sequence galaxies derived from the $sL_{\rm X}$ distributions of Aird et al. (2019).[1] We note a remarkable consistency between our results and those of that earlier work.

In order to be able to quantify the difference in these fractions we calculate the probability that a randomly selected posterior $sL_{\rm X}$ distribution from the starburst sample has a higher fraction of high accretion rate sources than a randomly selection posterior distribution from the non-starburst sample. We find that starbursts have a larger fraction of high accretion rate AGN than non-starbursts in 99.6%, 99.97%, and >99.99% of cases in our low, middle, and high redshift bins, respectively. In other words, our inferred distributions suggest one is significantly more likely to identify a high accretion rate AGN in a given starburst compared to a given non-starburst.

The result that the starburst population has a higher fraction of high $sL_{\rm X}$ is consistent with the findings of Georgakakis et al. (2014); Wang et al. (2017); Aird et al. (2017, 2018), who found that the distribution of accretion rates was shifted to lower values in quiescent galaxies compared to star-forming galaxies. By contrast, we also find no strong evidence that the positions of the exponential turnover in the distribution differs between the two populations. Overall, we interpret this in terms of SMBHs in starburst galaxies spending longer at higher accretion rates, but the maximum possible accretion remain broadly the same across the two populations. This could be caused by the SMBH self-regulating at accretion rates close to the Eddington limit. With recent evidence that starburst events are more commonly associated with interactions (Pawlik et al., 2018; Kauffmann, 2018; Dietrich et al., 2018) this could be interpreted as further evidence that interactions also enhance the levels of SMBH accretion (Comerford et al., 2015; Glikman et al., 2015; Ricci et al., 2017).

[1]Aird et al. (2019) used optical to near-infrared SED fits, as opposed to the far-infrared data used in this study, to classify galaxies according to their star-forming properties

At face value, our results seem to show no indication that intense radiation produced from an AGN during an accretion phase negatively impacts star-formation (Di Matteo et al., 2005; Fabian, 2012). Otherwise, we may have expected to find heightened accretion rates within non-starburst galaxies. However, care must be exercised when considering the stochastic nature of AGN variability, since any impact on the FIR-derived SFR will be delayed by roughly 100 Myr (Kennicutt, 1998). Indeed taking the complementary approach of measuring the SFR distribution in X-ray luminosity bins, Scholtz et al. (2018) demonstrates the need for negative feedback in simulations to reproduce the observed X-ray luminosity-dependent stellar mass specific SFR (sSFR) distributions. This demonstrates that the relationship between AGN feedback and SFR requires multiple complementary analysis methods to provide a complete picture. We therefore stress that the above result should not be interpreted as evidence against AGN activity quenching star formation rate, as any study of this nature fails to adequately account for the time-delay between AGN activity and the shutting-down of star formation.

3.6 Conclusions

In this Chapter we have developed a flexible model in order to infer the specific accretion rate distributions of central SMBHs within starburst and non-starburst galaxies. Our model distribution consists of a power-law curtailed by an upper and lower turnover, and allows us to incorporate information from upper limits, thereby allowing our sample to be more representative of the galaxy population in general. We derived the specific accretion rates from the 2-10 keV X-ray luminosities (or upper limits thereof) and used deblended *Herschel* maps to estimate the star formation rates. A source was classified as starburst if it had a SFR a factor of 3 greater than the main sequence at its redshift.

The main conclusions of this work are as follows:

1. Given our assumed model, we find suggestive (i.e., between 1.8 and 3σ) evidence

that the accretion rate distributions for massive galaxies ($\log_{10}(M_*/M_\odot) > 10.5$) are dependent on both the star-forming properties of the galaxies and on redshift.

2. More specifically, when modelled as a curtailed power-law, the gradient of the power law slope of the accretion rate distribution is shallower (i.e., less negative) in starburst galaxies, suggesting there is a slightly higher probability of detecting a high $sL_{\rm X}$ (high Eddington ratio) AGN in galaxies that have recently undergone an intense period of star-formation. This suggests that SMBHs in starburst galaxies spend more time at higher accretion rates than their non-starburst counterparts.

3. We find stronger evidence that starbursts and non-starbursts differ in terms of their specific accretion rates when we use our posterior $sL_{\rm X}$ distributions to calculate the fractions of such galaxies with high accretion rates (i.e., greater than $0.1\lambda_{\rm Edd}$). In doing so, we estimate that the fraction of starbursts hosting high accretion rate AGN is larger than the fraction of non-starbursts at confidence levels of 99.6%, 99.97%, and ¿99.99% for our low ($0.05 \leq z < 0.5$), mid ($0.5 \leq z < 1.5$), and high ($1.5 \leq z < 2.5$) redshift bins, respectively.

4. Within our uncertainties, we find no evidence that the positions of the high end turnover of the accretion rate distribution differs between starburst and non-starburst galaxies. We interpret this as suggesting that, whilst there are a higher fraction of SMBHs accreting at higher rates in the starburst population, the maximum accretion rates do not differ considerably, particularly in our low and middle redshift bins. This suggests that either the SMBHs are being self-regulated as they approach the Eddington limit or at least some other process is preventing accretion at considerably higher rates.

By selecting a galaxy sample and investigating how the full distribution of accretion rate properties changes as a function of star-forming properties (i.e., Approach A, as

described in Section 1.3.2) we have revealed a deeper understanding of the accretion rate properties of galaxies. It is natural, therefore, if we wish to provide a deeper understanding of the star-forming properties of AGN (i.e., Approach B) to investigate how the full distribution of star-forming properties changes as a function of accretion rate properties, which is the aim of the next chapter.

Chapter 4

Finding a subtle difference in the $R_{\rm MS}$ distribution between lower and higher X-ray luminosity AGN

> The Universe is under no obligation to make sense to you.
>
> Neil Degrasse Tyson

Declaration

As was stated at the beginning of the thesis, a large part of this chapter is based off a publication in which I am not the primary author (i.e., Bernhard et al., 2019) - I was second author. For the work I present here, I helped with the initial development of the paper, helped with the statistical modelling of the $R_{\rm MS}$ distribution, calculated the stellar masses for the sample and assisted with the analysis and interpretation of the results. Whilst I did not write the aforementioned paper, all the text in this chapter has been entirely written by myself.

4.1 Introduction

In Chapter 3 we discussed how various differences in the underlying distributions of SMBH accretion rates can result in almost identical changes in the average (and would thus be indistinguishable should only averages be considered). Thus, in an attempt to explain what underlying properties were causing an increase in average SMBH accretion rate per unit star formation, we instead considered the full distributions. The use of these full distributions allowed us to reveal significant characteristics of the connection between SMBHs and their host galaxies in more detail. However, the results presented in that chapter do not fully explain the contradictory results discussed in Chapter 1 between SFR and AGN power (i.e., the lack of a correlation between average SFR per unit AGN power). Explaining these results requires us to take the alternative approach to that taken thus far – investigate the star-forming properties of AGNs as a function of their AGN power. Given that a variety of distributions can have similar averages, the lack of a correlation between average SFR in bins of AGN power does not necessarily corroborate with the notion of no connection between star formation and SMBH growth. Thus, in order to either identify subtle differences, or confirm the lack of connection, we must investigate the *distribution* of star-forming properties.

In this chapter, we derive and compare the full (i.e., including upper limits) $R_{\rm MS}$ distribution between low and high X-ray luminosity ($L_{\rm X}$, tracing AGN power) AGNs. The outline of this chapter is as follows. In Section 4.2 we explain how we construct our sample. In Section 4.3, we explain how we model the $R_{\rm MS}$ distribution, including the parametric form we assume and derive the posterior distribution. In Section 4.4, we summarise our results and discuss possible implications in Section 4.5.

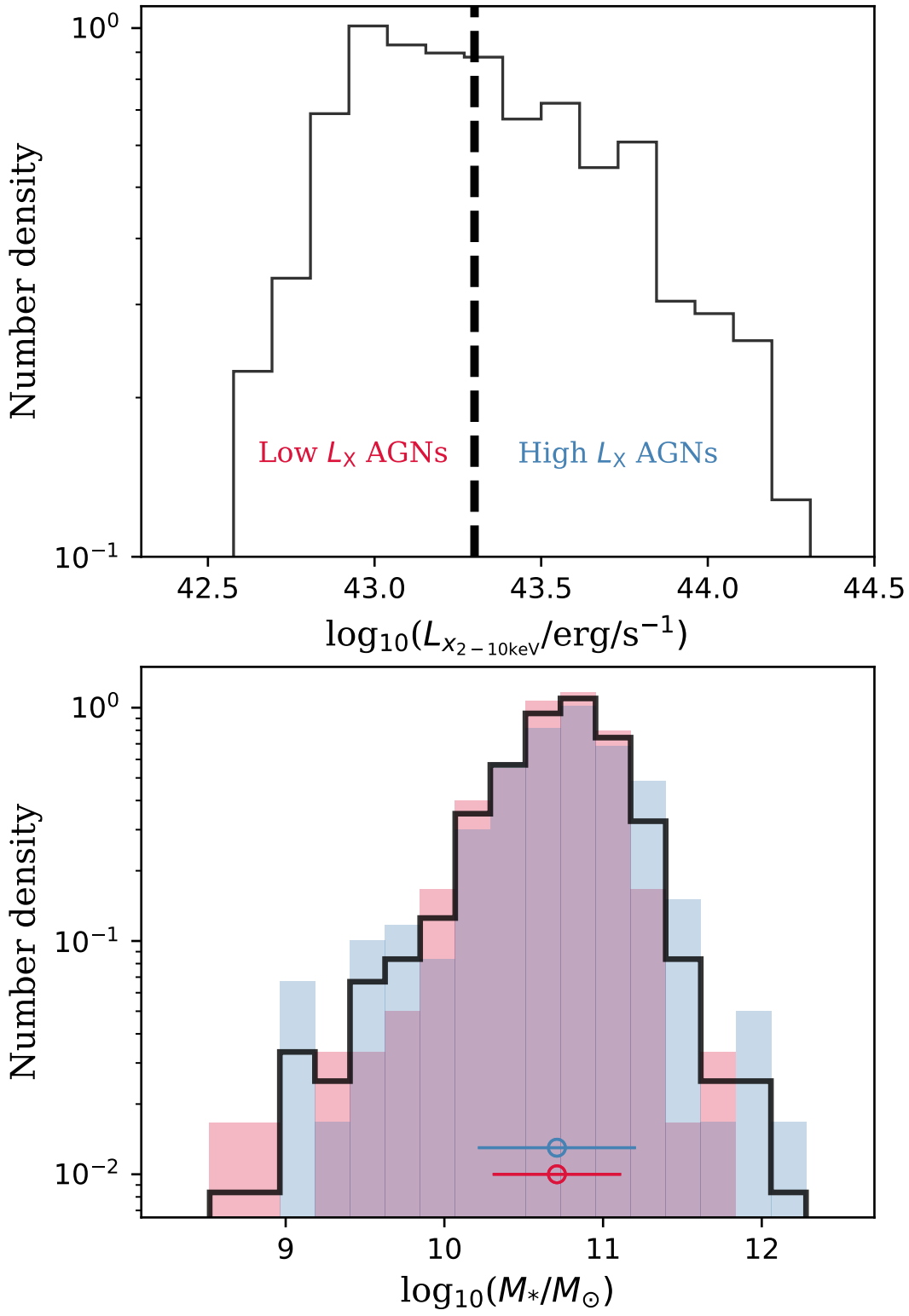


Figure 4.1: *Top*: The distribution of X-ray luminosities of our sample, highlighting the division between low and high L_X AGNs as selected in this study. *Bottom*: The stellar mass distribution for the low (red) and high (blue) L_X samples and the total stellar mass distribution (black). There are no immediately obvious differences in the stellar mass distribution between the two samples.

4.2 Sample derivation

In this study, we chose to adopt the alternative approach to investigating the relationship between SMBH accretion rate and star formation to that taken in Chapter 3. That is, in this chapter, we select a sample of AGNs, and investigate how the distribution of $R_{\rm MS}$ changes as a function of their AGN power. Thus, in order to derive our sample we start with the sources presented in the C16 catalogue. This catalogue contains the absorption-corrected 2-10 keV luminosity for $\approx$ 4000 AGNs, however, in order to mitigate any potential redshift effects, we select only those in the redshift range $0.8< z<1.2$. This redshift range is chosen as a balance between being closer to the peak epoch of both SMBH growth and star formation (e.g., Madau & Dickinson, 2014; Delvecchio et al., 2014; Aird et al., 2015; Vito et al., 2018) and having sufficient *Herschel* detections. There are 776 AGNs in the C16 catalogue that satisfy the redshift cut. In order to derive FIR SFRs, we require adequate *Herschel* data and therefore, we discard 112 sources that do not have adequate *Herschel* coverage (we do keep sources with adequate coverage, but no detections and we explain how we consider them in the next paragraph). We also discard a further 123 sources that do not have a detected $L_{\rm X}$, but instead an upper limit, leaving us with 541 AGNs[1]. The top plot of Figure 4.1 shows the $L_{\rm X}$ distribution for all 541 AGNs derived in the sample.

In order to derive SFRs (and subsequently $R_{\rm MS}$ values) we match these 541 AGNs to the super-deblended FIR photometry catalogue presented in Jin et al. (2018), which is the same catalogue as was used in Chapter 3. Whilst the majority of our AGNs have detections in at least one of the 6 bands presented (i.e., 24μm, 100μm, 160μm, 250μm, 350μm or 500μm), 100 AGNs did not have a detection at any of these wavelengths. For those 100 AGNs, we derive 3σ upper limits in the 100μm and 160μm bands by using the sensitivity maps provided by the PACS Evolutionary Probe team (Lutz

[1]Details of how to extend this work to include $L_{\rm X}$ upper limits and extend beyond our redshift range is fully explained in Chapter 6

et al., 2011). From this FIR photometry, we adopt the multi-component IR SED fitting code DECOMPIR[2] (Mullaney et al., 2011). DECOMPIR reconstructs the IR SED by combining a series of galaxy templates and an AGN template to derive a total IR luminosity for star formation. From there, a conversion is applied to convert the total IR luminosity to a SFR following the prescription of Kennicutt (1998). This applies for the 30% of sources with 3 detected fluxes in any of the 6 aforementioned wavelengths. For the remaining sources, which all have fewer than 3 detections, an upper limit on total IR luminosity (and hence on SFR) is derived by only fitting a host galaxy (i.e., non-AGN) component. By only fitting host galaxy templates, any AGN contribution is ignored. It is therefore appropriate to treat these sources as upper limits (i.e., it would be the true SFR if there was no contamination from the AGN). Once we have derived SFRs (or upper limits thereof) we can use the redshift of the sources (75% of which were spectroscopic) and stellar masses (as calculated in Section 2.4 and displayed in the bottom plot of Figure 4.1) to convert from SFRs to $R_{\rm MS}$ values. We still consider the $R_{\rm MS}$ over SFR (or specific SFR), as there could be subtle changes in the SFR distribution that are driven by redshift, even within our redshift selected sample.

In order to investigate how the $R_{\rm MS}$ changes with AGN power, we split our sample into two bins according to L_X. Those sources with $L_X > 2 \times 10^{43}$ erg s^{-1} are classified into the high L_X bin, whereas those with $L_X < 2 \times 10^{43}$ erg s^{-1} are classified into the low L_X bin. This threshold is chosen so that the sample size between the two groups is equal with 271 sources (with 65 detected $R_{\rm MS}$ and 206 upper limits) in the low L_X sample and 270 sources (with 83 detected $R_{\rm MS}$ and 187 upper limits) in the high L_X sample. We highlight that the choice of L_X in this study (as opposed to sL_X previously) is motivated by the lack of apparent difference in the stellar mass distributions between the low and high L_X samples as shown in Figure 4.1.

[2]DECOMPIR is publicly available at https://sites.google.com/site/decompir/

4.3 Parametric form and posterior distribution

A constant theme throughout this thesis is adequate consideration for sources with upper limits. If our sample was fully detected, the analysis is straight-forward, as most parametric distributions have algebraically-stated maximum likelihood estimations (MLEs) for their parameters. However, these algebraic forms for MLEs do not hold when presented with upper limits. In this section we therefore outline our assumed parametric form for modelling the $R_{\rm MS}$ distribution, how we derive a likelihood function (and subsequently a posterior distribution) such that upper limits are adequately accounted-for and then we describe our techniques for maximising the posterior distribution.

In order to model the $R_{\rm MS}$ distribution we assume a log-normal form. Although recent studies have found the scatter around the main sequence to be well-modelled by a log-normal distribution (Rodighiero et al., 2011; Sargent et al., 2012; Guo et al., 2013; Chang et al., 2015; Mullaney et al., 2015; Caplar & Tacchella, 2019; Davies et al., 2019; Popesso et al., 2019a,b) there may be a "bump" in the high-$R_{\rm MS}$ end of the distribution caused by starburst galaxies. Indeed, it is also true that there is likely an additional component at lower $R_{\rm MS}$ values due to the population of quiescent galaxies. Therefore the accuracy of using a log-normal distribution could be questioned. However, we leave devising a more flexible model to Chapter 6, where we introduce the possibility of multi-component models and discuss their credibility. It is important that we stress that this study – and results arising from it – are working under the assumption that the deviation from the main sequence of star formation is log-normally distributed, at least for AGNs.

As we choose to use a Bayesian approach, we wish to derive the posterior distribution, which is proportional to the product of the data-driven likelihood function (assuming a log-normal $R_{\rm MS}$ distribution) and the prior distributions. We are then interested in sampling parameter values from this posterior distribution. The remainder of this section, therefore, describes how we derive the likelihood function.

The likelihood function is given by the product of the PDFs of all the detected $R_{\rm MS}$ values, and the CDFs of all undetected sources. The PDF of a given detected $R_{\rm MS,i}$ value with parameters μ (representing the mode) and σ (representing the width), is given by

$$f(\log_{10}(R_{\rm MS,i})|\mu,\sigma) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{(\log_{10}(R_{\rm MS,i})-\mu)^2}{2\sigma^2}\right). \tag{4.1}$$

For upper limits (i.e., non-detected $R_{\rm MS}$ values, which ultimately comes from an upper limit on the SFR) the PDF is replaced by the CDF. The CDF is the integral of the PDF and can therefore be written as,

$$\begin{aligned} F(\log_{10}(R_{\rm MS})|\mu,\sigma) &= \int_{-\infty}^{R_{\rm MS}} f(X|\mu,\sigma)dX \\ &= \frac{1}{2}\left(1+\mathrm{erf}\left(\frac{\log_{10}(R_{\rm MS})-\mu}{\sigma\sqrt{2}}\right)\right), \end{aligned} \tag{4.2}$$

where $f(X|\mu,\sigma)$ is given by Equation 4.1. In other words, for a given galaxy, $F(\log_{10}(R_{\rm MS}))$ is close to 1 if most of the $R_{\rm MS}$ distribution with given μ and σ values lies below the value of the upper limit. By contrast, $F(\log_{10}(R_{\rm MS}))$ is close to 0 if most of the distribution lies above the upper limit, meaning those μ and σ values are *incompatible* with that limit.

By combining both our m detections, $R_{\rm MS,1},...,R_{\rm MS,m}$, and $n-m$ non-detections, $R_{\rm MS,m+1},...,R_{\rm MS,n}$, the likelihood function is given by the product of the PDFs (for the detections) and the CDFs (for the upper limits),

$$\begin{aligned} L(\log_{10}(R_{\rm MS})|\mu,\sigma) = &\prod_{i=1}^{m} f(\log_{10}(R_{\rm MS,i})|\mu,\sigma) \\ &\prod_{i=m+1}^{n} F(\log_{10}(R_{\rm MS,i}))|\mu,\sigma). \end{aligned} \tag{4.3}$$

As this likelihood function is too complex to maximise analytical, we adopt sampling package EMCEE (Foreman-Mackey et al., 2013) to derive posterior distributions for parameters μ and σ. We adopt uninformative flat prior distributions and note that changing the bounds of our flat prior does not affect the posterior distributions.

4.4 Results

4.4.1 R_{MS} distributions

In this study, we investigate how the R_{MS} distribution, under an assumed log-normal form, changes between a sample of low and high L_X AGNs. The parametric form of the distribution depends on the parameters μ (controlling the mode) and σ (controlling the width). Having constructed a posterior distribution for these (i.e., the product of the likelihood and prior distributions), we used sampling package EMCEE to derive posterior parameter distributions. These posterior distributions are shown in the contour plots in Figure 4.2. For comparison, in Figure 4.3 we also plot both the results from Mullaney et al. (2015), who investigated the R_{MS} distribution of a sample of higher redshift AGNs, and the results from Schreiber et al. (2015) who reported the R_{MS} distribution of main sequence galaxies. Our results for low L_X AGNs are consistent with those of Mullaney et al. (2015) whilst our high L_X sample appear to have R_{MS} distributions more similar to the star-forming main sequence. The median and 1σ uncertainties of the posterior distribution for each parameter are displayed in Table 4.1. The R_{MS} distribution, using the median parameter values from the posterior distribution, are displayed in Figure 4.3. Overall, the differences seen in the posterior distribution amount to a $\approx 2\sigma$ difference in the R_{MS} distribution of low and high L_X AGN samples.

These results suggest that the R_{MS} distributions for low and high L_X AGNs could be different, with high L_X AGNs having a slightly higher (as a result of having a larger

Sample	μ mean of $\ln(R_{MS})$	σ std. dev. of $\ln(R_{MS})$
This Work Low L_X AGNs	$-0.30^{\pm 0.06}$	$0.55^{\pm 0.05}$
This Work High L_X AGNs	$-0.10^{\pm 0.04}$	$0.40^{\pm 0.03}$
All AGNs ($z< 1.5$) (Mullaney et al., 2015)	$-0.38^{+0.07}_{-0.08}$	$0.6^{\pm 0.1}$
Main Sequence (Schreiber et al., 2015)	$-0.06^{\pm 0.02}$	$0.31^{\pm 0.02}$

Table 4.1: Median parameter posterior values for μ and σ with 1σ uncertainties.

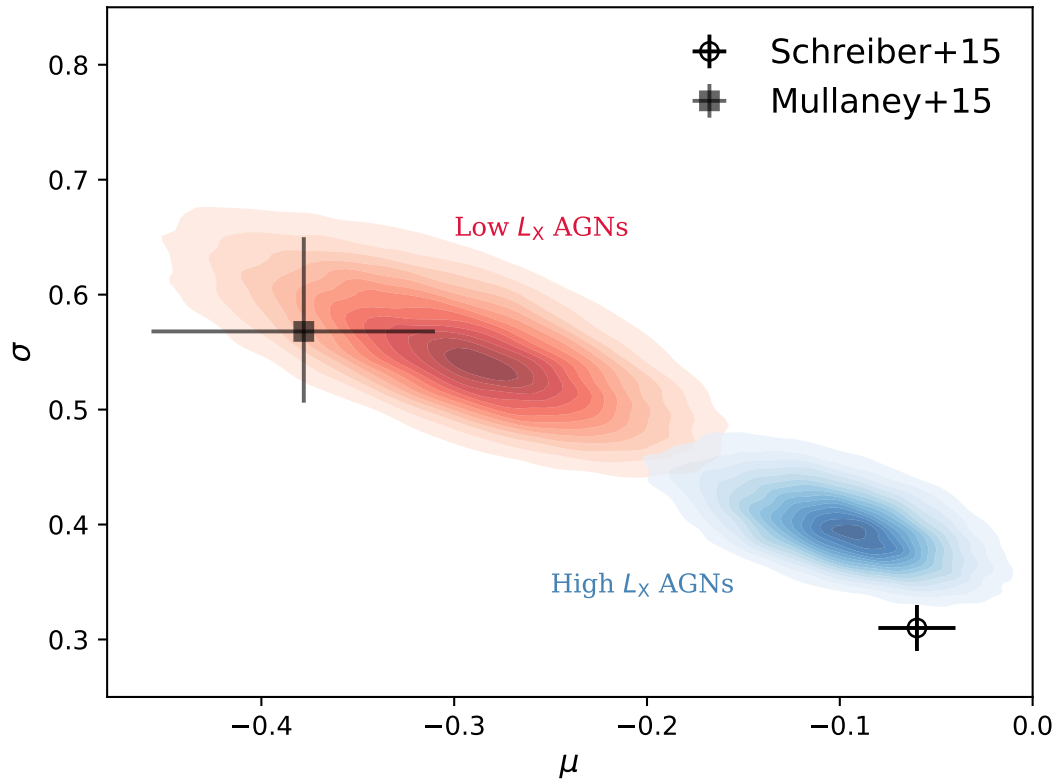


Figure 4.2: Contour plot for the posterior distribution of the R_{MS} model parameters μ and σ, which control the locus and width of the log-normal distribution respectively. The high L_X sample appears to have a higher μ and lower σ than the low L_X sample. Also, for comparison, the parameters of the main sequence from Schreiber et al. (2015) and the parameters for a higher redshift AGN sample from Mullaney et al. (2015) are plotted.

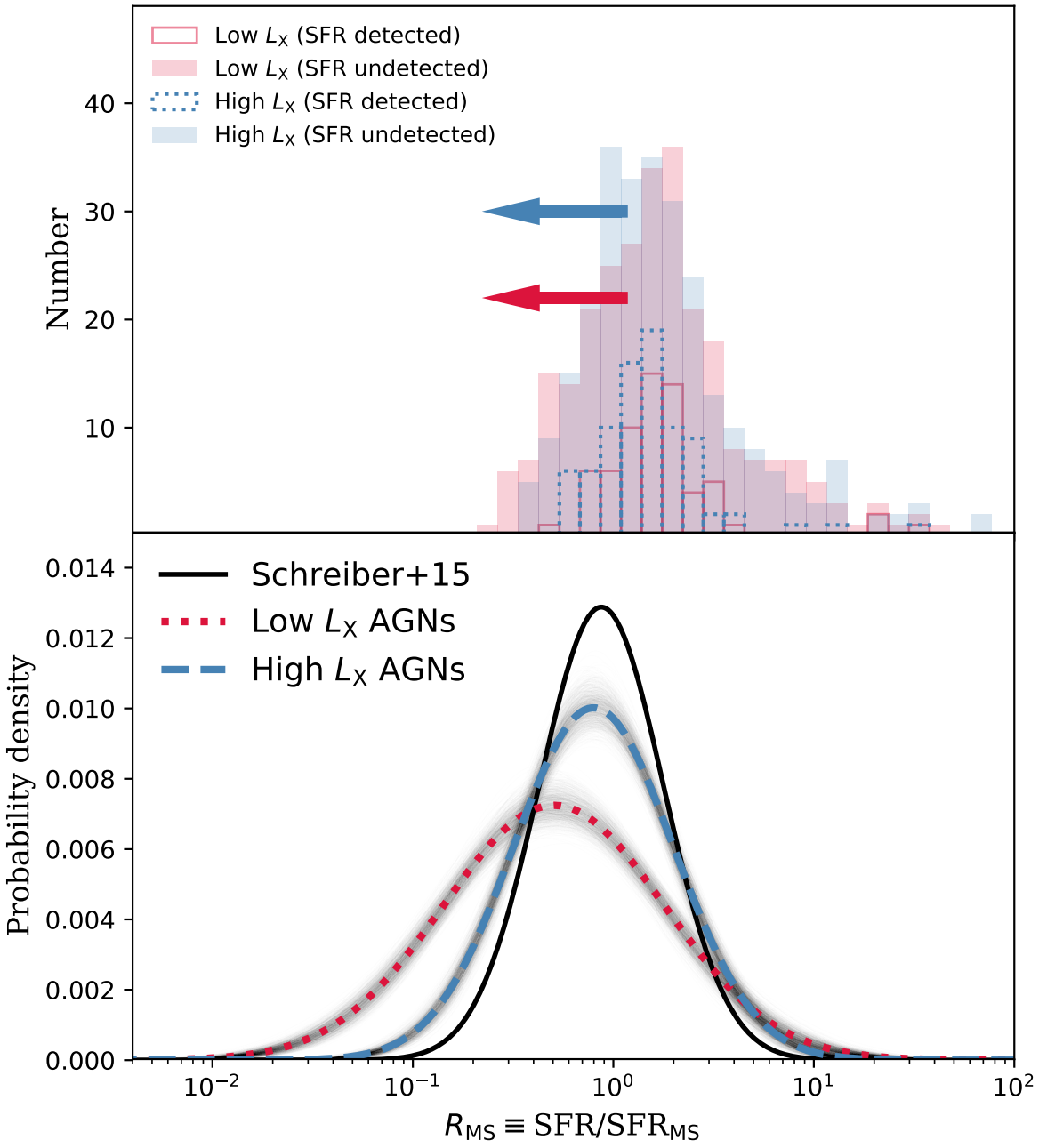


Figure 4.3: *Top*: The distribution of detected and upper limit R_{MS} values (empty and filled histograms respectively), split between the low and high L_X samples. *Bottom*: The inferred R_{MS} distribution from the median of the parameter posterior distributions. Also plotted is the R_{MS} distribution for main sequence galaxies from Schreiber et al. (2015)..

μ) and slightly narrower distribution (as a result of having a smaller σ) than the low L_X AGNs. If these results hold, the findings of this study indicate that not only do high L_X AGNs have star-forming properties that are more consistent with main sequence galaxies (Schreiber et al. 2015), they also have less diversity in their star-forming properties when compared to low L_X AGNs.

4.4.2 The relationship between SFR and L_X

Recall that studies investigating how the average SFR changes across bins of AGN power find little evidence of a connection (e.g., Harrison et al., 2012; Rosario et al., 2012; Stanley et al., 2015; Suh et al., 2017; Stanley et al., 2017; Ramasawmy et al., 2019). At first glance, given we have detected a subtle difference in the star-forming properties of AGNs depending upon their AGN power, the results of those studies appear to contradict the findings of this work. However, in Figure 4.4, we investigate how both the mean and the mode SFR, inferred from our R_{MS} distributions, change between the low and high L_X samples. Interestingly, the mean SFRs between our two samples are entirely consistent with the flat relationship seen in both Stanley et al. (2015) and Lanzuisi et al. (2017). Also, within their errors, the mean SFR of the low L_X sample and the mean SFR of the high L_X sample differ by less than 1σ, whereas the parameters of their respective R_{MS} distributions appear to differ by $\approx 2\sigma$. The mode SFR, however, being less influenced by bright outliers, shows a greater difference between the two samples, but not at the same significance level as the distribution-style analysis shows. However, the overall message remains that, the summary statistics tend to show less evidence of a connection between R_{MS} and L_X, even if the distributions provide more evidence of a difference.

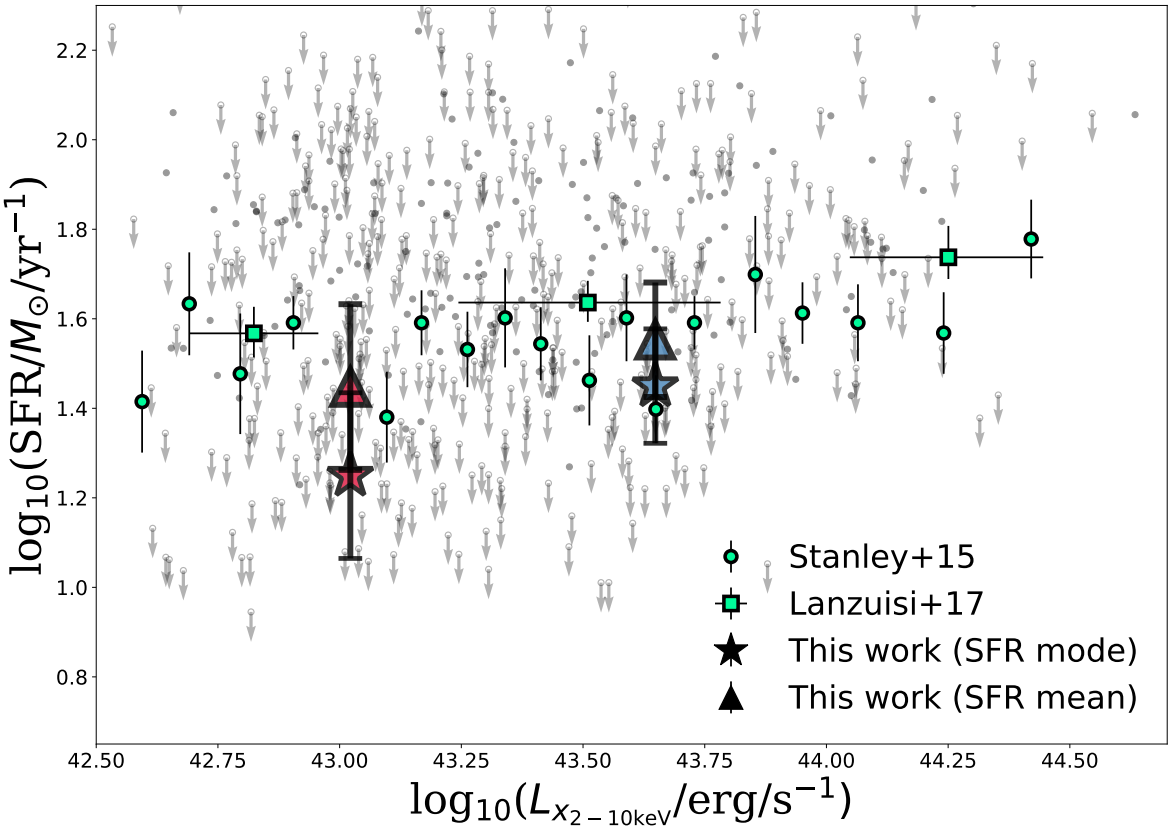


Figure 4.4: The mean (triangle) and mode (stars) SFR for the low (red) and high (blue) $L_{\rm X}$ samples derived from our $R_{\rm MS}$ posterior distributions. Also plotted are the flat relationships seen in Stanley et al. (2015) and Lanzuisi et al. (2017). Within uncertainties, there is very little evidence to suggest that the mean SFR changes between the two samples, whereas the mode, as a result of being less affected by outliers, show a greater difference. However, both summary statistics show less of a connection between $R_{\rm MS}$ and $L_{\rm X}$ than is suggested by our distribution-style analysis.

4.5 Discussion and Conclusions

In this chapter, we have investigated how the full (i.e., including upper limits) distribution of $R_{\rm MS}$ changes between two samples of AGNs grouped according to their $L_{\rm X}$ (tracing their accretion rate, see Section 2.2), under the assumption that $R_{\rm MS}$ is log-normally distributed. Our analysis provides tentative evidence that there is a difference between the star-forming properties of AGNs with low and high $L_{\rm X}$ values. More specifically, we find evidence that the high $L_{\rm X}$ sample (i.e., those with $L_{\rm X} > 2 \times 10^{43}$erg s^{-1}) have a narrower yet slightly higher $R_{\rm MS}$ distribution than the low $L_{\rm X}$ sample (i.e., those with $L_{\rm X} < 2 \times 10^{43}$erg s^{-1}). If this result holds true (e.g., with an increase sample size, or a more thorough analysis such as that used in the forthcoming chapter), this likely means that more luminous AGN reside in galaxies with a slightly higher, yet smaller range of star-forming environments, than lower luminosity AGNs.

We propose the results in this work are consistent with the idea that gas availability regulates both SMBH growth and galaxy growth. Being more easily triggered than their rapidly accreting counterparts (a natural corollary of the X-ray luminosity functions, see Aird et al. 2017), lower luminosity AGNs are more likely to reside in galaxies with varying gas abundances. Higher luminosity AGNs, however, require larger amounts of gas to be funnelled into the most central regions and we claim, therefore, more likely to require a higher abundance of gas in the host galaxy (although see e.g., Shlosman et al. 1989; Storchi-Bergmann et al. 2007; Audibert et al. 2019; Shimizu et al. 2019). Reinforced by the stronger link seen between more luminous AGN and star formation, than between lower luminosity AGNs, it is also likely the triggering of a lower luminosity AGN is less dependent of the gas content of the host galaxy, than the triggering of a high luminosity AGNs. Therefore, given higher luminosity AGNs have a stronger link with the gas content of the host, it is natural to suggest they have a stronger connection with gas-codependent star formation. In the broader picture, however, our results provide

reasonable evidence that the accretion rate of a SMBH is connected to the star-forming properties of the host galaxy, as is seen in other studies (e.g., Kauffmann & Heckman, 2009; Georgakakis et al., 2014; Wang et al., 2017; Aird et al., 2019; Bernhard et al., 2018) and additionally reaffirms the notion that quasars are often associated with high levels ongoing star formation (e.g., Rosario et al., 2013; Kalfountzou et al., 2014; Stanley et al., 2017).

Finally, whilst the results presented in this chapter provide tentative evidence of a relationship between star-forming properties and SMBH accretion rates, it fails to match the significance of the relationship found in Chapter 3 and the wider literature (i.e., when compared to galaxy selected samples, e.g., Azadi et al., 2015; Bernhard et al., 2016; Wang et al., 2017; Aird et al., 2017, 2018). In Chapter 3, the use of full distributions revealed a more detailed picture of how accretion rate properties change between starburst and non-starburst galaxies, than averages had done. Additionally, the use of full distributions here has revealed a more detailed (and more significant) picture of how star-forming properties change between low and high power AGNs. However, it may still be that uncertainties associated with binning (and specifically in the highly variable $L_{\rm X}$ axis) as we have here is weakening any true underlying connection. In the next chapter, we develop a technique that allows us to move away from binning in $L_{\rm X}$ such that we can investigate the change in $R_{\rm MS}$ distribution as a continuous function of $L_{\rm X}$.

Chapter 5

A binning-free method reveals a continuous relationship between galaxies' AGN power and offset from main sequence.

> The essence of the independent mind lies not in what it thinks, but in how it thinks.
>
> Christopher Hitchens

5.1 Introduction

A key means of investigating what galaxy-scale factors govern SMBH growth rates is by quantifying the properties of AGN-hosting galaxies and attempting to identify correlations between these host properties and AGN power. However, this is hampered by the fact that, compared to most other galactic processes (e.g., star-forming events, mergers), AGNs are extremely variable and short-lived. As demonstrated by Hickox et al. (2014),

this stochastic duty cycle tends to dilute the underlying connections between AGN power and other galactic properties, such that plots of mean galaxy star formation rate (SFR) vs. AGN power, for example, show a flat (i.e., independent) relationship (e.g., Harrison et al., 2012; Rosario et al., 2012; Mullaney et al., 2012a; Stanley et al., 2015, 2017; Suh et al., 2017; Ramasawmy et al., 2019). Recently, Scholtz et al. (2018) compared the distribution of specific SFR in two X-ray luminosity ($L_{\rm X}$) bins, but did not find any significant evidence of a difference between the two bins ($43 < \log_{10}(L_{\rm X}/{\rm ergs\ s^{-1}}) < 44$ and $44 < \log_{10}(L_{\rm X}/{\rm ergs\ s^{-1}}) < 45$). In Chapter 4, we compared the distribution of the $R_{\rm MS}$ statistic in bins of low $L_{\rm X}$ (i.e., $42.53 < \log_{10}(L_{\rm X}/{\rm erg\ s^{-1}}) < 43.3$) and high $L_{\rm X}$ (i.e., $43.3 < \log_{10}(L_{\rm X}/{\rm erg\ s^{-1}}) < 45.09$), and only found tentative evidence (i.e., $\approx 2\sigma$) of a dependency.

So whilst the use of distributions has allowed us to investigate the star-forming properties of AGNs in more detail than using simple averages, no study has demonstrated that the distribution of star-forming properties is dependent on $L_{\rm X}$[1]. Of course, this may be because no intrinsic connection exists. It could, however, be due to an often unaddressed limitation in the analysis: the use of arbitrarily-constructed bins of $L_{\rm X}$. As discussed in Chapter 1, binning can be somewhat arbitrary, weakly-motivated and can possibly impact results (Lanzuisi et al., 2017).

In this chapter, to investigate the implications of binning on our investigations of the relationship between star-forming properties and AGN power, we analyse the $R_{\rm MS}$ distribution as a *continuous* function of $L_{\rm X}$. To do this, we develop a comprehensive Bayesian hierarchical model which has two substantial benefits over binning. Firstly, it allows us to eliminate the possibility of binning-dependent results. Secondly, the model allows us to accurately account for all uncertainties (including, where necessary, upper limits) on the independent variable (i.e., in our case $L_{\rm X}$). Specifically, in this chapter

[1]Note, here we use "dependence" in the strict mathematical sense, rather than suggesting that SFR physically depends on AGN power.

we aim to quantify the dependence between the $R_{\rm MS}$ distribution and $L_{\rm X}$, without the need for binning or averaging. In doing so, we extract all available information from our data and find strong evidence of a relationship between the star-forming properties of AGN-hosting galaxies and $L_{\rm X}$.

The outline of this chapter is as follows. In Section 5.2 we briefly summarise how the dataset was constructed. In Section 5.3 we summarise the hierarchical Bayesian model, explain how we eliminate the need for binning and briefly introduce our MCMC model switching algorithm, which will test whether the $R_{\rm MS}$ distribution is dependent on $L_{\rm X}$. In Section 5.4 we present the output of the analysis and discuss the limitations and implications in Section 5.5. Where necessary, we adopt a WMAP-7 year cosmology Larson et al. (2011).

5.2 Data

So that we can compare the results of our new method with previously found results, we decide to reuse the same dataset as constructed previously in Chapter 4. This will ensure that any differences are the direct result of the analysis method, rather than from differences between two independent data sets. However, we provide a summary of the sample derivation here.

Briefly, we take the 541 X-ray detected sources with a redshift between $0.8 < z \leq 1.2$ from the C16 catalogue. This small redshift range ($\sim$ 75 per cent have spectroscopic redshifts) is chosen to minimise any potential redshift effects. These sources have rest-frame 2-10 keV, absorption-corrected X-ray luminosities spanning the range $42.53 < \log_{10}(L_{\rm X}/{\rm erg\ } s^{-1}) < 45.09$ (see Marchesi et al. 2016 for details on how they calculated $L_{\rm X}$, including how they corrected for absorption). We should note that in order to remain consistent with Chapter 4 for the aforementioned purposes, we do not include those sources with upper limits on $L_{\rm X}$ nor account for redshift variation, although it

would be straightforward to do so as explained in Section 5.3.2. Uncertainties on $L_{\rm X}$ values are derived by converting the percentage error on the flux measurement presented in Marchesi et al. (2016). On comparing these errors to the upper and lower $L_{\rm X}$ bounds in Marchesi et al. (2016), we find that our uncertainties are generally more conservative. The inclusion of these uncertainties is one benefit of the methodology presented in this chapter over that presented in Chapter 4. We then derive a SFR for each source using the DECOMPIR code (see Mullaney et al. 2011 for full details) on the super-deblended photometry presented in the catalogue of Jin et al. (2018) which used the deblending technique of Liu et al. (2018). The catalogue contains data from various sources such as *Spitzer* and *Herschel* and covers the 24-1200μm range.

In total, our sample contains 148 AGNs with measured SFRs, and 393 with upper limits on their SFRs. Stellar masses are calculated using the multi-wavelength spectral energy distribution fitting code CIGALE (Noll et al., 2009; Serra et al., 2011; Ciesla et al., 2015; Boquien et al., 2019) as described in Chapter 2. The stellar mass parameters were chosen to maximise the accuracy according to the testing presented in Ciesla et al. (2015). Next, we use the prescription of Schreiber et al. (2015), together with each galaxy's redshift and mass, to predict the SFR that it would have if it were on the star-forming main sequence (i.e., $\rm SFR_{MS}$). Finally, we calculate the starburstiness statistic, $R_{\rm MS}$, of each galaxy in our sample using the method outlined in Section 2.3.2.

5.3 The continuous model, model selection and MCMC algorithm

In this section we describe how we model the $R_{\rm MS}$ data, in such a way to remove the need for binning, which enables us to investigate whether (and, if so, how) the $R_{\rm MS}$ distribution changes as a continuous function of $L_{\rm X}$. In subsection 5.3.1, we introduce

the log-normal distribution we use to model the $R_{\rm MS}$ distribution and explain why we must use a "hierarchical" Bayesian approach to allow this to vary continuously with $L_{\rm X}$. Next, in subsection 5.3.2 we describe our Bayesian priors and how these provide a mechanism to include all uncertainties on each individual $L_{\rm X}$ value. Finally, in subsection 5.3.3, we introduce our bespoke MCMC sampler that explores the posterior parameter space in a way that allows us to test whether the $R_{\rm MS}$ distribution depends on $L_{\rm X}$.

5.3.1 $R_{\rm MS}$ distribution and likelihood function

In order to test the continuous relationship between the $R_{\rm MS}$ distribution and $L_{\rm X}$ we assume a functional parametric form for the $R_{\rm MS}$ distribution. In this chapter, we choose to model the $R_{\rm MS}$ distribution as a log-normal distribution (i.e., that $\log_{10}(R_{\rm MS})$ is normally distributed). A log-normal distribution is chosen to remain consistent with Chapter 4. Recall that our approach is to derive a posterior distribution for the parameters (i.e., the product of the data-driven likelihood and a prior distribution). Recall that in Chapter 4, we derived the likelihood function,

$$L(\log_{10}(R_{\rm MS})|\mu,\sigma) = \prod_{i=1}^{m} f(\log_{10}(R_{\rm MS,i})|\mu,\sigma) \prod_{i=m+1}^{n} F(\log_{10}(R_{\rm MS,i}))|\mu,\sigma). \tag{5.1}$$

where $f(\log_{10}(R_{\rm MS,i})|\mu,\sigma)$ is the PDF for detected $R_{\rm MS}$ values and $F(\log_{10}(R_{\rm MS,i}))|\mu,\sigma)$ is the CDF for $R_{\rm MS}$ upper limits. If we were going to assume no dependence of $R_{\rm MS}$ on $L_{\rm X}$, and no uncertainty on $L_{\rm X}$, then at this stage we could simply find the best-fitting values for μ and σ, as has been used previously in studies that bin in $L_{\rm X}$ (i.e., Chapter 4). Such studies derive the likelihood function in different bins, use parameter-maximisation techniques to find the best fitting value for μ and σ within each bin, and then compare

how parameters change between different bins (e.g., Mullaney et al., 2015; Scholtz et al., 2018). However, in order to analyse the $R_{\rm MS}$ distribution as a continuous function of $L_{\rm X}$, we *must* use a hierarchical model, since this allows the parameters that control the shape of the $R_{\rm MS}$ distribution (i.e., μ, σ) to vary as a function of $L_{\rm X}$. As the true relationship between the μ and σ parameters and the $L_{\rm X}$ values is unknown, the choice of relationship is arbitrarily specified. However, in order to test the case of no dependence (i.e., that $R_{\rm MS}$ and $L_{\rm X}$ are independent of one another), it is sufficient to show that a simple model that allows dependence is preferable to one that imposes independence. Therefore, we choose to use simple functions to relate the parameters of the $R_{\rm MS}$ distribution and the $L_{\rm X}$ values (hereafter referred to as the "functional relationships"), given by:

$$\mu_i = \theta_0 + \theta_1 \log_{10}\left(\frac{L_{\rm X,i}}{10^{40}}\right) \quad \text{and} \quad \sigma_i = e^{\theta_2 + \theta_3 \log_{10}\left(\frac{L_{\rm X,i}}{10^{40}}\right)}. \tag{5.2}$$

The rescaling of the $L_{\rm X}$ values ensures that $\boldsymbol{\theta} = \{\theta_0, \theta_1, \theta_2, \theta_3\}$ (hereafter, our hyper-parameters) are not orders of magnitude different, which could lead to problems in the analysis. Note that, throughout this chapter, we are only considering the effect of $L_{\rm X}$ on the $R_{\rm MS}$ distribution and hence our functional relationships only factor-in $L_{\rm X}$. If other parameters, such as redshift or stellar mass were also to be considered, they could be added to the functional relationships as described in Equation 5.2. Such an expansion of the model is presented in Chapter 6.

By introducing these functional relationships, we have essentially related the mode and width of the $R_{\rm MS}$ distribution to the $L_{\rm X}$ values. Additionally, we have changed the parameters of interest from μ and σ to the hyperparameters; this is what makes the approach "hierarchical". Note that we specify an exponential form for the functional relationship between σ_i and $L_{\rm X,i}$ as σ_i cannot be negative. The focus of this analysis is to now find the posterior distributions for $\boldsymbol{\theta}$. By considering these posteriors, the functional relationships allow us to test whether the $R_{\rm MS}$ distribution is dependent upon

$L_{\rm X}$. For example if $\theta_1 = \theta_3 = 0$, the functional relationships are no longer a function of $L_{\rm X}$ and therefore imply that the $R_{\rm MS}$ distributions are independent of $L_{\rm X}$. Additionally, relating the mode and width of the $R_{\rm MS}$ distribution to the $L_{\rm X}$ values has completely removed the need to bin the data in $L_{\rm X}$. The question of independence now becomes how likely is $\theta_1 = \theta_3 = 0$, given the data observed. More details of which are contained in Section 5.3.2.

As a result of adapting the mode and width of the distribution so that binning is not required, the likelihood function changes slightly and is now given by,

$$L(\boldsymbol{\theta}, L_{\rm X}|R_{\rm MS}) = \prod_{i=1}^{m} f(\log_{10}(R_{\rm MS,i})|\boldsymbol{\theta}, L_{\rm X,i}) \prod_{i=m+1}^{n} F(\log_{10}(R_{\rm MS,i})|\boldsymbol{\theta}, L_{\rm X,i}). \tag{5.3}$$

5.3.2 Prior and posterior distributions

Prior distribution on $L_{\rm X}$

We have now expressed the parameters as functions of the independent data (in this case, $L_{\rm X}$) and the hyperparameters, $\boldsymbol{\theta}$. The next step we must now consider is how to fully account for uncertainties on $L_{\rm X}$. In our hierarchical model, we are able to treat the $L_{\rm X}$ values as parameters, and can therefore place informative Bayesian priors on their values. The prior distribution on each $L_{\rm X,i}$ can be constrained by the measured value $L_{\rm X,i,meas}$ and uncertainty ξ_i and modelled as a log-normal (here, we are assuming that our errors are symmetric in log space). This means that the prior distribution on a specific $\log_{10}(L_{\rm X,i})$ is given by,

$$f(\log_{10}(L_{\rm X,i})|\log_{10}(L_{\rm X,i,meas}), \xi_i) = (2\pi\xi_i)^{-\frac{1}{2}} \exp\left(-\frac{(\log_{10}(L_{\rm X,i}) - \log_{10}(L_{\rm X,i,meas}))^2}{2\xi_i^2}\right). \tag{5.4}$$

where ξ_i is derived by converting the percentage error on the flux measurement presented in Marchesi et al. (2016). This can be thought of as the probability density of observing the true $L_{\rm X}$ given we have observed a measurement, $L_{\rm X,i,meas}$ and error ξ_i. It should be noted that in this chapter we are working with only detected X-ray luminosities to remain consistent with Chapter 4 and we assume all uncertainties are modelled with a log-normal. One could, however, replace this prior distribution with any probability distribution. Note that in this chapter, we have not accounted for the uncertainties on the $R_{\rm MS}$ values. This is largely to remain consistent with the modelling approach of Chapter 4. In future studies, uncertainties on the dependent variable (in our case, $R_{\rm MS}$) can be included using a similar method as the one applied to the s$L_{\rm X}$ distribution in Chapter 3. Whilst we do not believe that excluding these uncertainties has a major impact on our results, it is a limitation of the work in this chapter. However, it is not a limitation of the methodology.

At this stage, we have specified our likelihood function (Equation 5.3) and our priors on $L_{\rm X}$. The final terms we must consider are the prior distributions on the hyperparameters, which we discuss in the next subsection.

Prior distribution on hyperparameters

Because our primary scientific aim is to determine *whether* the $R_{\rm MS}$ distribution changes with $L_{\rm X}$, we are most interested in the (posterior) probability that the hyperparameters θ_1 and θ_3 are equal to 0 or whether they are non-zero (i.e., there is a dependence on $L_{\rm X}$). We therefore choose the prior distributions of these hyperparameters to be a "spike and slab distribution". This type of prior allows us to join two distributions; one defined in discrete space (the spike) and one in continuous space (the slab). This is necessary so that we can ensure that there is a defined prior probability that $\theta_1 = 0$ and $\theta_3 = 0$ (i.e., there is a prior probability of independence between $R_{\rm MS}$ and $L_{\rm X}$), as opposed to just a probability density. If we have a defined prior probability then we can calculate a

posterior probability, again as opposed to just to a probability density. [2]

Our spike and slab prior distributions take the form:

$$\begin{aligned} f(\theta_1|\omega) &= (1-\omega)\mathrm{N}(\theta_1; \mathrm{mean} = 0, \mathrm{S.D.} = 1) + \omega\delta_{\theta_1=0}, \\ f(\theta_3|\omega) &= (1-\omega)\mathrm{N}(\theta_3; \mathrm{mean} = 0, \mathrm{S.D.} = 1) + \omega\delta_{\theta_3=0}, \end{aligned} \tag{5.5}$$

where ω is the prior probability that $\theta_1, \theta_3 = 0$ and $\delta_{\theta_i=0}$ is the delta function. For our analysis, we choose $\omega = 0.5$ so that our prior probability favours neither the case of independence, $p(\theta_1 = 0) = p(\theta_3 = 0) = 0.5$, nor the case of dependence $p(\theta_1 \neq 0) = p(\theta_3 \neq 0) = 0.5$. As we are not interested in the posterior probabilities that $\theta_0, \theta_2 = 0$, the prior distributions on these parameters are Gaussian distributions with mean 0 and standard deviation 1.

By using spike and slab prior distributions we have constructed four potential models:

- Model 1: $\theta_1 = 0, \theta_3 = 0$, no dependence on L_X at all
- Model 2: $\theta_1 \neq 0, \theta_3 = 0$, mode depends on L_X, width does not
- Model 3: $\theta_1 = 0, \theta_3 \neq 0$, width depends on L_X, mode does not
- Model 4: $\theta_1 \neq 0, \theta_3 \neq 0$, both mode and width depend on L_X.

Note that as we have chosen $\omega = 0.5$ our prior distributions give no preferential weight to any of the model scenarios (according to the prior, they all have a probability of 0.25). Having now derived the likelihood function and all needed prior distributions we can construct the final posterior distribution:

[2] A probability density is a "relative" likelihood as opposed to an absolute one. For a distribution over a continuous space, the absolute probability of any one particular occurrence is 0, whilst the probability density can be non-zero. For a distribution over a discrete space, the probability mass function (the discrete equivalent of the density) is an absolute probability.

$$
\begin{aligned}
f(\boldsymbol{\theta}, \log_{\mathbf{10}}(\mathbf{L_X}) | \log_{\mathbf{10}}(\mathbf{R_{MS}}), \log_{\mathbf{10}}(\mathbf{L_{X,meas}})) = \\
L(\log_{\mathbf{10}}(\mathbf{R_{MS}}) | \boldsymbol{\theta}, \log_{\mathbf{10}}(\mathbf{L_X})) \\
\times f(\log_{\mathbf{10}}(\mathbf{L_X}) | \log_{\mathbf{10}}(\mathbf{L_{X,meas}}), \boldsymbol{\xi}) \\
\times f(\boldsymbol{\theta} | \omega)
\end{aligned}
\tag{5.6}
$$

5.3.3 MCMC algorithm and model switching

As our posterior distributions cannot be derived analytically, we have written a purpose-built MCMC sampler in order to sample from the posterior distributions of each given hyperparameter (i.e., θ_0, θ_1, θ_2, θ_3). However, in addition to sampling from the posterior distributions to find the most likely hyperparameter values, we also use our sampler to determine the posterior probability of each of our four models (i.e., for model comparison). The posterior probability of the models can be calculated analytically, however even advanced sampling methods (e.g., Nested Sampling, see Buchner et al. 2014) struggle to accurately calculate them due to the high dimensionality of our parameter space (i.e., up to 545 dimensions as a result of including the L_X values as parameters). Instead, we use "model switching" to compute the posterior model probabilities. In this subsection, we fully describe one full step of the MCMC sampler used to construct the posterior distributions presented in Section 5.4, which were then used to compare our various models. Interested readers should also refer to the study of Gottardo & Raftery (2008), from which our sampler is adapted.

For the most part, our MCMC sampler adopts a standard Metropolis-Hastings (MH) algorithm (Metropolis et al., 1953; Hastings, 1970) to explore the parameter space. On each iteration, the MH algorithm proposes a new set of parameter values, which are then accepted or rejected. For efficiency, we propose new values for two parameters at a time, and accept them based on their acceptance ratio α, where:

$$\alpha = \min\left(\frac{\pi(\theta')q(\theta',\theta)}{\pi(\theta)q(\theta,\theta')}, 1\right), \tag{5.7}$$

where θ' is the proposed parameter value, θ is the current parameter value, $\pi(\theta)$ is the full conditional of θ and $q(\theta,\theta')$ is the proposal density (i.e., the probability density of proposing θ' given the current θ). For our analysis, the parameter vector is given by $\boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2, \theta_3, \log_{10}(L_{\mathrm{X},1}), ..., \log_{10}(L_{\mathrm{X,n}}))$. We choose to sample θ_0, θ_1 together and θ_2, θ_3 together as the value of θ_0 is highly-dependent on the value of θ_1; similarly, the value of θ_2 is highly-dependent on θ_3. Proposing the dependent hyperparameters together can allow us to take into account the dependency and therefore propose more sensible values, which greatly improves the speed and efficiency of the sampler. If we were only considering one model, and simply wished to sample the posterior distributions, then we would simply iterate the above process. However, in our case we wish to compare the relative probability of four different models. As mentioned above, we do this using a technique known as "model switching", which we describe next.

A key component of our algorithm is that, when it proposes a switch between models, it proposes "reasonable" parameters within the proposed model. Otherwise, we run the risk of never switching models – not because the proposed model is necessarily worse, but because we always propose highly unlikely parameter values within that model. What we mean by "reasonable", therefore, is likely parameter values within each proposed model. As such, we need to have some knowledge of the posterior probability distributions of each model before we can start proposing switches between models. One way of achieving this would be to force Model 1, for example, to converge, then force a switch to Model 2, allow that to converge, and so on. Once all models have converged, we would then allow our sampler to switch between models by proposing reasonable parameter values (i.e., those close to the posterior mode). In our case, however, as we only have four models, we instead run a separate standard MCMC sampler for each model (i.e., without model

switching), which gives us an indication of the most suitable regions of the posterior parameter space for each model. Mathematically, these two approaches are exactly analogous.

With an estimate of the posterior parameter space for each model in-hand, we can propose reasonable regions of the parameter space when switching between models. In what follows, we describe how we switch between various models. For ease of explanation, we will only consider θ_0 and θ_1, but same process is applied when sampling θ_2 and θ_3. Recalling that we step through the parameters in pairs, we sample θ_0 and θ_1 at the same time. This leads to four possible cases, which are summarised in Table 5.1, and discussed in detail below.

Case A: Here, the sampler is currently in the state where $\theta_1 = 0$, and is proposing $\theta_1 = 0$ (i.e., it is in a μ-independent model [Models 1 or 3] and proposes to remain within a μ-independent model). However, because we progress through the vector pairwise, the sampler must still propose a θ_0 value. For this, we use a standard MH proposal – a value randomly selected from a Gaussian distribution centred on the current θ_0 value. Based on pilot runs, we choose a value for the width of the Gaussian distribution that results in good mixing (i.e., the acceptance rate is between 20–40 per cent). In this case, the $q(\theta, \theta')$ value is the product of the likelihood of choosing $\theta'_1 = 0$ (i.e., 0.5) and the proposed θ_0 value (i.e., $\theta'_0 = f(\theta'_0|\theta_0, s_1)$, where f is the Gaussian density function). This product is symmetrical on switching between θ and θ', meaning $q(\theta, \theta') = q(\theta', \theta)$, so the q terms cancel in Equation 5.7.

Case B: In this case, the sampler is currently in the state where $\theta_1 = 0$, and is proposing $\theta_1 \neq 0$ (i.e., it is in a μ-independent model [Models 1 or 3] and is proposing to switch to a μ-dependent model [Models 2 or 4]). As a result of proposing a switch to a μ-dependent model, we must propose values for both θ_0 and θ_1. To do this, we use a bivariate Gaussian distribution, centred on the "reasonable" values obtained using the process described above. Based on pilot runs, we choose a value for the widths of the

Case	Current θ_1	Proposed θ_1'	Model now	Model proposed	$q(\theta, \theta')$	$q(\theta', \theta)$
A	$\theta_1 = 0$	$\theta_1' = 0$	1 3	1 3	$0.5 \times f(\theta_0' \mid \theta_0, s_1)$	$0.5 \times f(\theta_0 \mid \theta_0', s_1)$
B	$\theta_1 = 0$	$\theta_1' \neq 0$	1 3	2 4	$0.5 \times f_2(\ [\theta_0', \theta_1'] \mid [\hat{\theta}_0, \hat{\theta}_1], \Sigma_1)$	$0.5 \times f(\theta_0 \mid \hat{\theta}_0, s_2)$
C	$\theta_1 \neq 0$	$\theta_1' = 0$	2 4	1 3	$0.5 \times f(\theta_0 \mid \hat{\theta}_0, s_2)$	$0.5 \times f_2(\ [\theta_0, \theta_1] \mid [\hat{\theta}_0, \hat{\theta}_1], \Sigma_1)$
D	$\theta_1 \neq 0$	$\theta_1' \neq 0$	2 4	2 4	$0.5 \times f_2(\ [\theta_0', \theta_1'] \mid [\theta_0, \theta_1], \Sigma_2)$	$0.5 \times f_2(\ [\theta_0, \theta_1] \mid [\theta_0', \theta_1'], \Sigma_2)$

Table 5.1: Summary of the possible model switches for 1 proposal of the μ-related hyperparameters, θ_0 and θ_1. There are four potential cases depending on whether the model is currently in a μ-dependent or a μ-independent state and whether we propose to move to a μ-dependent or μ-independent state. For the possible cases the value of the proposal density $q(\theta, \theta')$ and the inverse $q(\theta', \theta)$ are given. The univariate Gaussian density is given by f and the bivariate Gaussian density is given by f_2. The tuned proposal widths are given by s_1 and s_2, and the calculated covariance matrices by Σ_1 and Σ_2. When a model switch is proposed, the "reasonable" values must be used to sample a proposed parameter value and these are given by $\hat{\theta}_0$ and $\hat{\theta}_1$.

bivariate Gaussian distribution that results in good mixing (i.e., the acceptance rate is between 20–40 per cent). In addition to the widths, the bivariate Gaussian distribution accounts for the correlation between θ_0 and θ_1 by using the calculated covariance matrix. In this case, the $q(\theta, \theta')$ value is the product of the likelihood of choosing $\theta'_1 \neq 0$ (i.e., 0.5) and the proposed θ values (i.e., $\theta' = f_2(\,[\theta'_0, \theta'_1]\,|\,[\hat{\theta_0}, \hat{\theta_1}], \Sigma_1)$, where f_2 is the bivariate Gaussian density function, $\hat{\theta_0}, \hat{\theta_1}$ are the estimates of the posterior mode from the original chains and Σ_1 is the covariance matrix. This product is not symmetrical on switching between θ and θ', since the inverse process involves sampling from a univariate Gaussian. This means $q(\theta, \theta') \neq q(\theta', \theta)$, so they must be accounted for in the acceptance ratio.

Case C: Here, the sampler is currently in the state where $\theta_1 \neq 0$, and is proposing $\theta'_1 = 0$ (i.e., it is in a μ-dependent model [Models 2 or 4] and is proposing to switch to a μ-independent model [Models 1 or 3]). As a result of proposing a switch to a μ-independent model, we again must propose a "reasonable" value of θ_0 within the proposed model. To do this, we use a distribution, centred on the "reasonable" values obtained using the process described above. Based on pilot runs, we choose a value for the width of the Gaussian distribution that results in good mixing (i.e., the acceptance rate is between 20–40 per cent). In this case, the $q(\theta, \theta')$ value is the product of the likelihood of choosing $\theta'_1 = 0$ (i.e., 0.5) and the proposed θ_0 value (i.e., $\theta'_0 = f(\theta'_0|\hat{\theta_0}, s_2)$, where f is the Gaussian density function, $\hat{\theta_0}, \hat{\theta_1}$ are the estimates of the posterior mode from the original chains and Σ_1 is the covariance matrix). This product is not symmetrical on switching between θ and θ' for the same reason as in Case B (i.e., the inverse process involves sampling from a bivariate Gaussian distribution). This means $q(\theta, \theta') \neq q(\theta', \theta)$, so they must be accounted for in the acceptance ratio.

Case D: In this final case, the sampler is currently in the state where $\theta_1 \neq 0$, and is proposing $\theta'_1 \neq 0$ (i.e., it is in a μ-dependent model [Models 2 or 4] and is proposing to remain in a μ-dependent model). As a result we need to propose values for both θ_0 and θ_1. To do this, we use a bivariate Gaussian distribution, centred on the

current values. Based on pilot runs, we choose a value for the width of the Gaussian distribution that results in good mixing (i.e., the acceptance rate is between 20–40 per cent) and calculate the appropriate covariance matrix. In this case, the $q(\theta, \theta')$ value is the product of the likelihood of choosing $\theta'_1 \neq 0$ (i.e., 0.5) and the proposed θ value (i.e., $\theta' = f_2(\,[\theta'_0, \theta'_1] \,|\, [\theta_0, \theta_1], \Sigma_1)$, where f_2 is the bivariate Gaussian density function, and Σ_2 is the covariance matrix). This product is symmetrical on switching between θ and θ', meaning $q(\theta, \theta') = q(\theta', \theta)$ and so the terms cancel.

This process is then repeated for the next pair of hyperparameters (i.e., θ_2 and θ_3) followed by one sampling through the L_{X} values individually (i.e., not pair-wise), the latter of which is done by using a standard MH algorithm. In one iteration we sample through the full parameter vector and we run five chains in parallel for 25,000 iterations.[3] Each chain has the first 5000 iterations removed as a burn-in, then the remaining iterations from each chain are combined to form the final sample of 100,000 posterior draws for each parameter. The posterior probability of each of the four models presented in Section 5.3.2 is then straightforward to calculate from the combined chain: all we need to do is calculate the fraction of accepted samples from each model in the combined chain.

5.4 Results

Given that we now have 100,000 independent draws from the posterior distribution from each parameter, we can begin to investigate the relationship between the R_{MS} distribution and L_{X}. Recall that we modelled the R_{MS} distribution as a log-normal distribution and set a relationship between the mode and width, and the L_{X} values as outlined in Equation 5.2. We proposed values such that our sample was forced to consider $\theta_1 = 0$ and $\theta_3 = 0$ respectively, effectively allowing for the MCMC sampler to

[3]The choice of five chains for 25,000 iterations is arbitrary, but these values ensured that the combined chain contained a sufficiently high number of samples from the posterior.

switch between models of dependence or independence. In this Section, we present the posterior distributions of the hyperparameters and the posterior model probabilities.

5.4.1 Posterior distributions

Posterior model probabilities

As a result of implementing model switching in the MCMC algorithm we can easily calculate the posterior model probabilities by considering the fraction of samples of each chain within each model. The posterior model probabilities alongside the Bayes Factor comparison to the independent Model 1 are given in Table 5.2. The Bayes Factor, which can be accurately used to compare two models (Kass & Raftery, 1995), is given as the ratio of the posterior model probability of the more complex model to the posterior model probability more simple one. Naturally, the Bayes Factor includes a "penalty" for the number of parameters used. In our case, as a result of including L_{X} values as a parameters our models have vastly different numbers of parameters. Model 1, which ignores L_{X} values only has 2 parameters, whereas Models 2, 3 and 4 have 544, 544 and 545 parameters, respectively. This can help explain the very small posterior probabilities of Models 2 and 3, where the chain either prefers the simple Model 1, or for the sake of 1 extra parameter Model 4, which comprehensively outperforms them. The Bayes Factor comparing Model 4 to Model 1 gives us a value of 15.285, which can be seen as "strong" evidence in favour of Model 4 (Kass & Raftery, 1995). Using this model comparison model technique, the posterior model probability is not equal to the probability that the model is true, as the sum of all posterior model probabilities in the analysis must be equal to 1. It is therefore important to consider the Bayes Factor approach for comparing the models, rather than using the posterior model probabilities as they are.

Model	Value of μ	Value of σ	Posterior probability	Bayes Factor vs. Model 1
1	θ_0	e^{θ_2}	0.06102	-
2	$\theta_0 + \theta_1 \log_{10}\left(\frac{L_X}{10^{40}}\right)$	e^{θ_2}	0.00477	0.0781
3	θ_0	$e^{\theta_2+\theta_3 \log_{10}\left(\frac{L_X}{10^{40}}\right)}$	0.00148	0.02425
4	$\theta_0 + \theta_1 \log_{10}\left(\frac{L_X}{10^{40}}\right)$	$e^{\theta_2+\theta_3 \log_{10}\left(\frac{L_X}{10^{40}}\right)}$	0.93273	15.285

Table 5.2: The posterior model probabilities given for each model. These are calculated by considering the amount of time the MCMC chain spent in each of the models. Also shown is the Bayes Factor, which is used to judge, out of two models, the model considered to be the most likely.

Hyperparameters

In Figure 5.1 we present the posterior distributions for the hyperparameters as computed by the MCMC algorithm outlined in Section 5.3.3. The off-diagonal plots show the joint posterior distributions. As described in Section 5.4.1, we have strong evidence that a model of the $R_{\rm MS}$ distribution with a dependence on $L_{\rm X}$ is preferred to the independent model. The rest of this chapter therefore, works with the assumption that Model 4 is the most suitable model.

We present summary statistics for the posterior distributions of the hyperparameters in Table 5.3. The coefficients of $L_{\rm X}$ in the functional relationships (see Equation 5.2) are given by θ_1 and θ_3, which from Table 5.3 and Figure 5.1 are positive and negative respectively. This implies that as $L_{\rm X}$ increases, the mode and width of the $R_{\rm MS}$ distribution increase and decrease respectively. The relationship between the mode and width of the log-normal $R_{\rm MS}$ distribution and $L_{\rm X}$ can be seen in Figure 5.2, where the posterior distributions of the hyperparameters have been sampled 1000 times and combined with $L_{\rm X}$ to provide samples of μ and σ.

5.4.2 $R_{\rm MS}$ as a function of $L_{\rm X}$

In this chapter, we have used a hierarchical Bayesian framework to remove the need for binning and stacking when modelling the $R_{\rm MS}$ distribution of galaxies hosting AGN of different $L_{\rm X}$. In doing so, and in contrast to Chapter 4, we find *strong* evidence that there is relationship between the $R_{\rm MS}$ distribution and $L_{\rm X}$ (i.e., AGN power) as opposed to just tentative evidence.

In Figure 5.3 we show how the $R_{\rm MS}$ distribution, when modelled as a log-normal distribution, changes as a function of $L_{\rm X}$ in the range $42.53 \leq \log_{10}(L_{\rm X}/\text{ergs s}^{-1}) \leq 45.09$. As $L_{\rm X}$ increases, the mode of the $R_{\rm MS}$ distribution increases, whilst the width decreases. This is also shown in Figure 5.1, as θ_1 takes positive values (i.e., μ increases

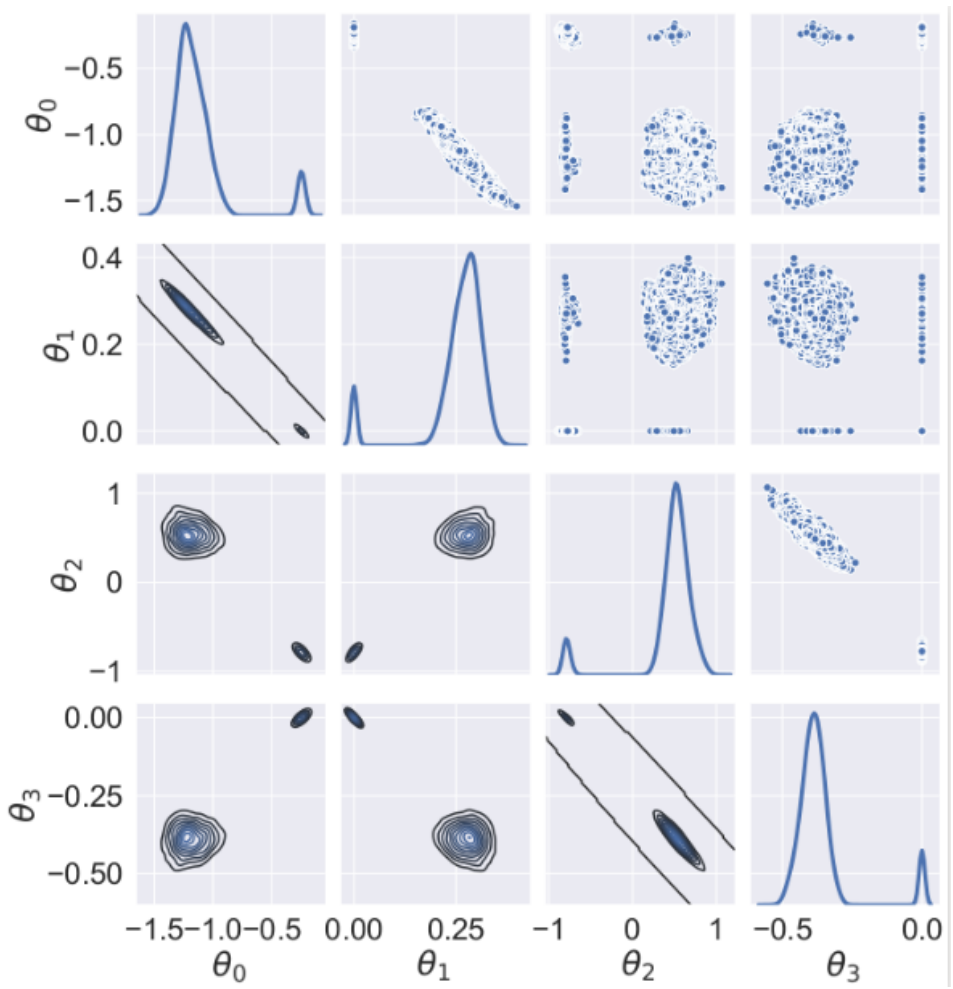


Figure 5.1: The output from our MCMC algorithm. The on-diagonal plots show the marginalised posterior distributions for each parameter, with the joint posterior distributions shown by the off-diagonal contour plots. The figures include results from the entire MCMC chain, which means that different peaks (on-diagonal) and contour regions (off-diagonal) illustrate when the chain is in a particular model. For example, in the plot in the second row, first column (from top left), the larger of the two contour regions corresponds to $\theta_1 \neq 0$, which is the case in both Model 2 and Model 4. From this posterior plot alone, one cannot distinguish whether the chain is in Model 2 or Model 4, as information about the other parameters is needed (i.e., a 4-dimensional plot would show four discrete model regions). Secondly, there is a smaller region in the lower-right corner that corresponds to the region where $\theta_1 = 0$, which is the case for both Model 1 and Model 3. Again, one cannot distinguish between these two models from this plot alone. However, given the negligible amount of time the chain spends in Model 2 and Model 3, it can be assumed without much loss of accuracy that the larger region represents the likelihood for Model 4 and the smaller region represents the likelihood for Model 1. This is analogous to the larger and smaller peaks in the on-diagonal plot for θ_1.

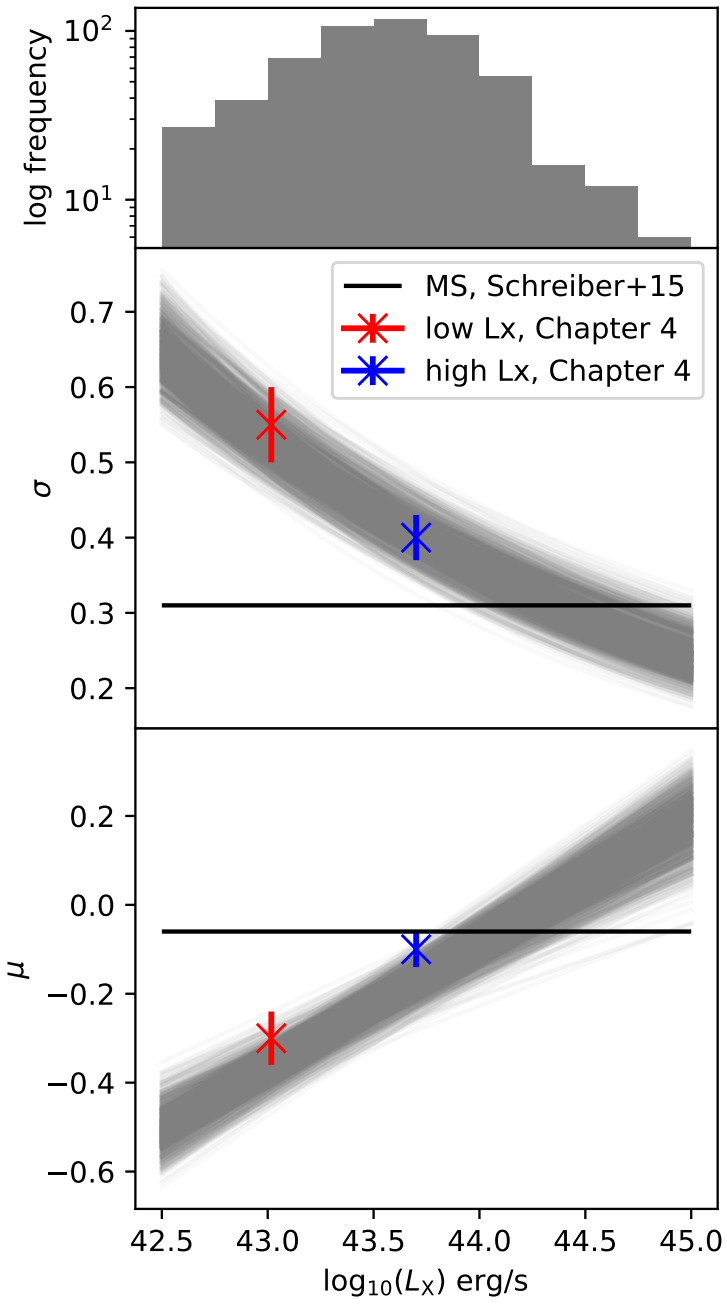


Figure 5.2: The evolution of the mode, μ, and width, σ, of the $R_{\rm MS}$ distribution as a function of $L_{\rm X}$ shown for 1000 bootstrapped samples from the posterior distributions of the hyperparameters, under the assumption of Model 4. Over-plotted are the results from Chapter 4, with 1-σ errors. Also plotted is the main sequence values from Schreiber et al. (2015) (solid black lines). The top plot is the histogram of $L_{\rm X}$ values of the sample for reference.

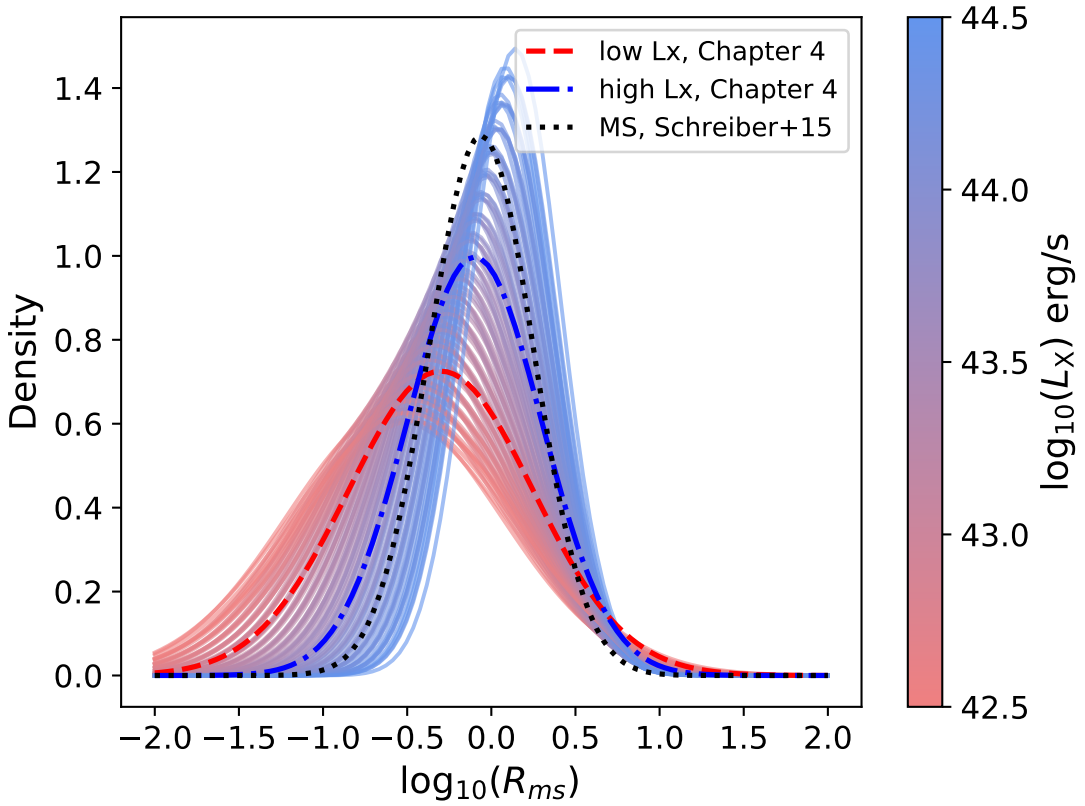


Figure 5.3: The evolution of the $R_{\rm MS}$ distributions as a continuous function of X-ray luminosity, plotted as thin curves. Over plotted are the results from Chapter 4 and the R_{ms} distribution for main sequence galaxies from Schreiber et al. (2015). As the X-ray luminosity of a galaxy increases, the probability density function for its $R_{\rm MS}$ shifts slightly to higher values and the distribution narrows, consistent with the findings in Chapter 4.

with increasing $L_{\rm X}$) and θ_3 takes negative values (i.e., σ decreases with increasing $L_{\rm X}$). These results, albeit with more evidence, are still consistent with the tentative findings of Chapter 4, which showed that more luminous X-ray AGNs have $R_{\rm MS}$ distributions closer to those of main sequence galaxies compared to lower $L_{\rm X}$ AGNs. This is also consistent with the findings of Schulze et al. (2019), who noticed no difference in the SFR distribution of 20 $z \sim 2$ quasars and the SFR distribution of main sequence galaxies.

With our new analysis showing stronger evidence of a dependence of $R_{\rm MS}$ on $L_{\rm X}$, it is natural to ask whether this is consistent with the observed flat relationship between SFR

and $L_{\rm X}$ reported by some other studies (e.g., Rosario et al., 2012; Stanley et al., 2015). We are able to explore this issue by generating synthetic SFRs using our $L_{\rm X}$-dependent $R_{\rm MS}$ model, together with the measured $L_{\rm X}$, redshifts, and stellar masses of our sample. To do this we:

1. randomly generate a sample from the joint posterior distribution of the hyper-parameters, $\theta_0^*, \theta_1^*, \theta_2^*, \theta_3^*$. This involves taking a random point from each of the off-diagonal plots in Figure 5.1 (and therefore respecting any correlations between parameters);

2. for each of the 541 sources in our sample we use their detected $L_{\rm X}$ values, alongside the aforementioned randomly sampled hyperparameters, to calculate the mode and width of the predicted $R_{\rm MS}$ distribution. Recall, we reuse the functional relationships we chose earlier so that we have a predicted mode, $\mu_{\rm pred}$ and predicted width, $\sigma_{\rm pred}$:

$$\begin{aligned} \mu_{\rm pred} &= \theta_0^* + \theta_1^* \log_{10}\left(\frac{L_{\rm X}}{10^{40}}\right) \quad \text{and} \\ \sigma_{\rm pred} &= e^{\theta_2^* + \theta_3^* \log_{10}\left(\frac{L_{\rm X}}{10^{40}}\right)}. \end{aligned} \tag{5.8}$$

3. we then sample an $R_{\rm MS}$ value from the log-normal distribution with the parameters $\mu_{\rm pred}$ and $\sigma_{\rm pred}$;

4. we then repeat steps 1-3 10,000 times so that we have, for each source in our sample, a set of 10,000 predicted $R_{\rm MS}$ values constrained by our hyperparameter posterior distributions and the assumption of our functional relationships;

5. we next multiply each of the sampled $R_{\rm MS}$ values by the corresponding main sequence SFR, calculated by using the stellar masses, redshifts and the prescription from Schreiber et al. (2015). This leaves us with a sample of 10,000 predicted

Table 5.3: Posterior mean and standard deviations for the hyperparameters for Model 4.

Parameter	Mode	Standard Deviation
θ_0	-1.191	0.119
θ_1	0.276	0.033
θ_2	0.54.	0.128
θ_3	-0.391	0.040

SFRs for each source calculated using our functional relationships and posterior distributions.

Figure 5.4 shows the relationship between SFR and $L_{\rm X}$ as predicted by our $L_{\rm X}$-dependent $R_{\rm MS}$ distribution. The red stars show the mean predicted SFR in bins of $L_{\rm X}$, using a bin width of 0.25 dex (with error bars indicating the 3σ standard error). Over-plotted are the observed mean SFRs (calculated using survival analysis), also in bins of $L_{\rm X}$, from Stanley et al. (2015). The yellow circles represent the SFRs of the 148 AGNs in our sample with measured fluxes, while the yellow triangles represent the upper-limits on SFRs for the remaining 393 AGNs. Despite our analysis providing strong evidence of a relationship between the $R_{\rm MS}$ distribution and $L_{\rm X}$, the projected relationship between the average predicted SFRs and $L_{\rm X}$ is comparable to the observed flat relationship of Stanley et al. (2015) (i.e., while the means are offset, they are well within the range of scatter given by the observed measurements). While the incorporation of mass and redshift information to convert our predicted $R_{\rm MS}$ values to SFR may contribute to some of the flattening, it is plausible that averaging over a log-normal distribution within a particular $L_{\rm X}$ bin could have significantly flattened the relationship also. This further demonstrates that even if a strong underlying relationship between star-forming properties and AGN power exists, it is extremely difficult to extract using average (or even individually-measured) SFRs in bins of $L_{\rm X}$.

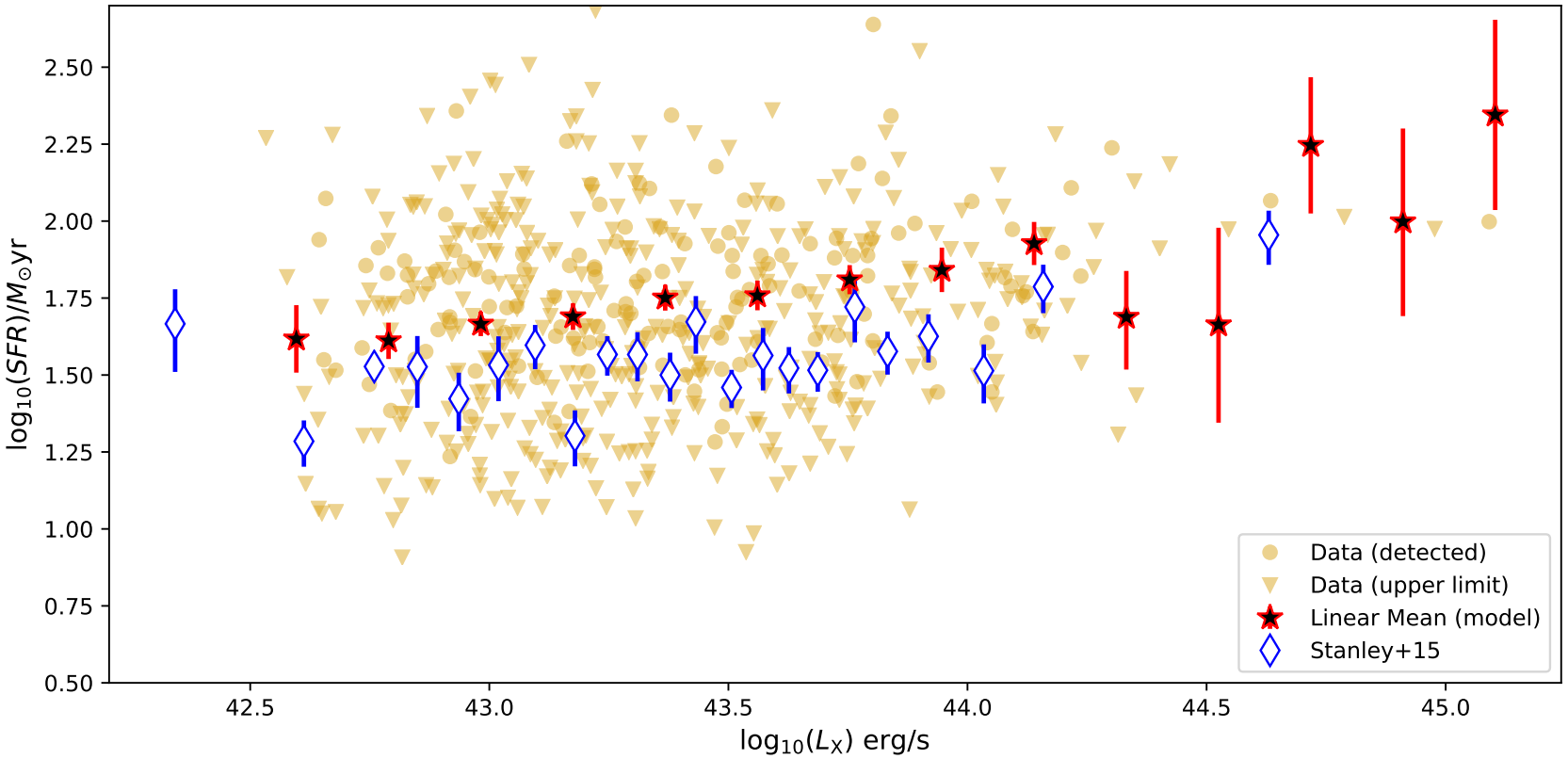


Figure 5.4: The predicted relationship between SFR and $L_{\rm X}$ using our functional relationships and hyperparameter posterior distributions. The red stars show the predicted linear mean SFRs in arbitrarily-chosen bins of $L_{\rm X}$, calculated using the functional relationships in Equation 5.2, the main sequence prescription of Schreiber et al. (2015) and the stellar mass and redshift of our sources. Also plotted in yellow (circles or triangles) are the SFRs from the raw data (detected and upper limits, respectively). The blue diamonds are the results from Stanley et al. (2015) for the redshift range $0.8 < z < 1.5$, which extends to slightly higher redshifts than our sample. While our results are systematically offset from those of Stanley et al. (2015), they are broadly consistent with their observed flat relationship. We include this plot purely to demonstrate that even after including a significant underlying connection between $R_{\rm MS}$ and $L_{\rm X}$, the we still obtain a flat relationship between average SFR in bins of $L_{\rm X}$.

5.5 Discussion

5.5.1 Limitations of our approach

Before discussing the implications of our results, in this section we aim to highlight limitations of our approach and discuss areas for potential improvement. Initially, as we reuse the same dataset as Chapter 4, we have adopted the same set of initial assumptions as that chapter - namely, the assumption about the parametric form of the $R_{\rm MS}$ distribution and the validity of the Schreiber et al. (2015) main sequence. However by removing the need for binning, we have relaxed the unstated assumption about sources in the same bins having similar properties. The remainder of this section, therefore aims to highlight additional limitations and assumptions with our methodology, as well as those of Chapter 4.

Firstly, the analysis is computationally expensive. This is mostly due to the large number of sampled parameters. In this case, there are four hyperparameters ($\theta_0, ...\theta_3$) plus, as described in Section 5.3.2, 541 $L_{\rm X}$ parameters with a well-defined (i.e., using by the measured value and its uncertainties) prior distribution. The parameters are sampled pair-wise throughout the MCMC algorithm, which reduces the time, but the algorithm is still computationally expensive. Despite having a large number of parameters, over-parameterisation is not a concern since the priors tightly constrain the $L_{\rm X}$ values.

Secondly, in this work, we have imposed simple relationships between the mode and width (μ, σ, respectively) of the $R_{\rm MS}$ distribution and $L_{\rm X}$. Whilst this relationship could be made more flexible, the aim of this chapter was to test the framework and to determine if there is any dependence on $L_{\rm X}$. We therefore chose simple relationships to assess whether we could rule-out the independent case. In future studies (as outlined further in Chapter 6), more flexible forms of the functional relationships could be tested and model comparison methods used to determine whether any other functional forms provide a better representation of the data. In addition to making the functional relationships

more flexible, other independent variables could be added (such as redshift and stellar mass). By doing so, and allowing for more models to be compared, future studies could use the techniques in this chapter to probe deeper into the connection between AGN power and host galaxy properties. As a result of this chapter only investigating how the $R_{\rm MS}$ distribution changes as a function of $L_{\rm X}$, we were cautious that, if there was a significant, systematic change of $L_{\rm X}$ with redshift, then a redshift evolution in both $L_{\rm X}$ and $R_{\rm MS}$ may introduce a spurious positive trend. However, we see no evidence of a strong systematic change of $L_{\rm X}$ with redshift. The median and standard deviation of $L_{\rm X}$ for the lowest and highest redshift quartiles were (43.23, 0.40) and (43.43, 0.44) respectively. Therefore we have no reason to believe that our results are being affected by an underlying redshift evolution in both $L_{\rm X}$ and $R_{\rm MS}$ across our redshift bin. With regards to redshift and stellar mass effects, it may be interesting to investigate whether assuming alternative models for the redshift and mass evolution of the main sequence (e.g., Speagle et al., 2014; Ilbert et al., 2015; Whitaker et al., 2015; Popesso et al., 2019a) has a large effect on the results.

Thirdly, posterior model probabilities can be dependent upon the choice of prior distribution chosen for individual parameters. As the marginal likelihood is the integral of the likelihood function over all the prior space (effectively a weighted average of the likelihood function), an analysis of this sort must make sure that the prior distributions are reflective of current up to date knowledge. Our prior distributions are influenced by the work of Chapter 4. By the construction of the marginal likelihood, however, overly vague prior distributions can excessively "penalise" more complex models. Likewise, prior distributions that are too constrained can favour more complex models. Therefore, prior distributions should be carefully chosen and justified.

Finally, we stress again that we have worked under the assumption that $R_{\rm MS}$ distribution is log-normal. This may not be entirely accurate. Indeed, it is known that some AGNs reside in quiescent and starburst galaxies whose combined $R_{\rm MS}$ values do not fol-

low a log-normal distribution (e.g., the main sequence/starburst population is believed to follow a bi-modal log-normal distribution in $R_{\rm MS}$). Having said that, our focus here is to assess whether, after eliminating the need for binning and averaging (and comparing to the same dataset in Chapter 4), the $R_{\rm MS}$ distribution could be $L_{\rm X}$-dependent. It is not immediately clear why a truly $L_{\rm X}$-independent $R_{\rm MS}$ distribution would be better modelled by a $L_{\rm X}$-dependent log-normal, as opposed to a $L_{\rm X}$-independent one. Therefore, we stress we are not suggesting that our model represents the true $R_{\rm MS}$ relationship, but instead that an $L_{\rm X}$-dependent model is strongly favoured when compared to an $L_{\rm X}$-independent one.

5.5.2 Implications of our analysis

The aim of this chapter was to introduce a Bayesian hierarchical framework that removes both the need to bin data (particularly in distribution-style analyses) and the need to use averaging techniques (or other summary statistics/parameters). To allow us to accurately demonstrate that any new results were driven by the methodology, we applied our hierarchical model on the same dataset as Chapter 4. The process involves assuming a distributional form for one variable (in this case the starburstiness of a galaxy) and setting a direct dependence between the parameters of this distribution and some independent variable (in this case, $L_{\rm X}$). Uncertainties on the independent variable are also fully considered by treating them as a parameter and applying an informative prior, which is derived from the measured values and their uncertainties.

Our results show that, under the assumption that $R_{\rm MS}$ is log-normally distributed, there is a strong evidence of a relationship between $R_{\rm MS}$ and $L_{\rm X}$ within the redshift range $0.8 < z < 1.2$. This reaffirms, to a stronger degree of significance, the result of Chapter 4, such that as $L_{\rm X}$ increases, the $R_{\rm MS}$ distribution is centred at a higher value and the diversity of $R_{\rm MS}$ values decreases. What this implies is that, within the

constraints of our model, an $L_{\rm X}$= 10^{44} erg s^{-1} AGN is 21 per cent more likely to reside in a galaxy with $R_{\rm MS}$> 2 than an $L_{\rm X}$= 10^{43} erg s^{-1} AGN. This is in agreement with other studies that suggested there is a *tighter* (i.e., more consistent) connection between more luminous AGNs and star formation than for lower-luminosity AGNs (e.g., Rosario et al., 2013; Stanley et al., 2017; Aird et al., 2017; Dai et al., 2018; Masoura et al., 2018; Aird et al., 2019): for example, it may be that any luminous AGN activity occurs close in time to the star formation activity while lower-luminosity AGN activity can occur when the galaxy is more quiescent (and hence the broader $R_{\rm MS}$ distribution) in addition to occurring during the periods of star-formation activity.

In this chapter, we have investigated the relationship between the $R_{\rm MS}$ distribution of AGN hosts and $L_{\rm X}$, and found strong evidence of a relationship between the two. Recently, a number of studies have approached this problem from the other direction; i.e., investigating how AGN power changes as a function of the star-forming properties of their hosts. For example, Chen et al. (2013) reported that, when binned in terms of SFR, the mean $L_{\rm X}$ of star-forming galaxies increases with average SFR (see also Delvecchio et al. 2015, who also accounted for the effects of galaxy stellar mass). Further, Rodighiero et al. (2015) found that, when binning according to stellar mass, the mean $L_{\rm X}$ of starburst galaxies is higher then that of main sequence galaxies which, in turn, is higher than that of quiescent galaxies. Both these results imply that average AGN power is higher in more actively star-forming systems. In Chapter 3 we reported that the distribution of specific $L_{\rm X}$ (i.e., = $L_{\rm X}/M_{*}$, a proxy for Eddington ratio $\lambda_{\rm Edd}$), changes as a function of the star-forming activity of their hosts, with a higher fraction of starbursts hosting AGNs with $\lambda_{\rm Edd} > 10\%$ than their main sequence counterparts (consistent with the work found in Aird et al. 2019). By exploring how the star-forming properties of galaxies change as a function of $L_{\rm X}$, this chapter (and Chapter 4) take the opposite approach. While there are significant differences between the properties being considered in each study (not least the exploration of Eddington ratio in Aird et al. 2019 and Chapter 3, whereas

we only consider $L_{\rm X}$ here) all appear to support the assertion that more powerful AGNs (whether expressed in terms of $L_{\rm X}$ or Eddington ratio) are preferentially found in more actively star-forming systems.

5.6 Conclusions

In this chapter, we have introduced a hierarchical Bayesian framework to assess whether the $R_{\rm MS}$ distribution of AGN-hosting galaxies changes as a *continuous* function of an X-ray luminosity ($L_{\rm X}$). Our approach removes the need for both binning and averaging and also allows for full consideration of the uncertainties on the independent variable.

By modelling the $R_{\rm MS}$ distribution as a log-normal, and proposing simple relationships between its parameters (i.e., mode and width) of that log-normal and X-ray luminosity, we found strong evidence that an $L_{\rm X}$-dependent model is preferred over an $L_{\rm X}$-independent one. By binning the same data, in Chapter 4 we reported the same overall trend, but without such strong evidence, thereby highlighting the importance of utilising all available information by removing the need for binning. By using the same dataset and pre-processing as Chapter 4, we ensured that any differences found in contrast to that chapter are a direct result of the new analysis technique.

Despite finding a strong relationship between the $R_{\rm MS}$ distribution and AGN power, when we convert our $L_{\rm X}$-dependent distributions back into the mean SFR - $L_{\rm X}$ plane, we find that the *dependent* model can reproduce results consistent with previously seen flat relationships (e.g., Stanley et al., 2015). This further highlights the difficulty in extracting underlying relationships between AGN power and host galaxy properties when averaging in bins of AGN power.

Chapter 6

Improvements to the binning-free methodology: multi-component distributions and functional relationships

If you're not improving, chances are you're not going to win.

Mike Shanahan

6.1 Introduction

The finding that the $R_{\rm MS}$ distribution changes significantly with $L_{\rm X}$ provides an alternative perspective on the connection between a SMBH and the host galaxy. More specifically, the results presented in Chapter 5 are consistent with the findings of studies that adopted the alternative approach (i.e., instead of an AGN-selected sample, using a galaxy-selected sample, e.g., Rafferty et al., 2011; Rosario et al., 2012; Chen et al.,

2013; Azadi et al., 2015; Delvecchio et al., 2015; Harris et al., 2016; Lanzuisi et al., 2017; Shimizu et al., 2017; Stemo et al., 2020, Chapter 3). All these results suggest that there is a stronger connection between higher-luminosity AGNs and the star-forming properties of the host galaxy than between lower-luminosity AGNs and the star-forming properties of the host galaxy.

Despite the binning-free methodology introduced in Chapter 5 being able to uncover a significant connection between the $R_{\rm MS}$ distribution and $L_{\rm X}$, there are still improvements that can be made. Firstly, there is considerable evidence that the $R_{\rm MS}$ distribution does not take the form of a single log-normal distribution. Indeed, it may be the case that starburst galaxies contribute to a bump in the high-$R_{\rm MS}$ tail of the distribution (Schreiber et al., 2015). Thus, in Section 6.2, we build a two-component model that could be used to describe the $R_{\rm MS}$ distribution in more detail. We test the two-component model on both a simulated dataset and then on the same AGN sample as was used in Chapters 4 and 5. A second possible improvement is to modify the functional relationships introduced in Equation 5.2 (the equations that link parameters of the $R_{\rm MS}$ distribution to $L_{\rm X}$) such that they include other independent variables. This could reveal how the $R_{\rm MS}$ distribution changes with a broad range of other independent data, for example, environment, galaxy colour, morphology (if somehow quantified). But perhaps the most relevant for the aims of this thesis would be to consider how the relationship between the $R_{\rm MS}$ distribution and $L_{\rm X}$ depends on stellar mass and redshift (which we discuss in Section 6.3). We finally summarise this chapter in Section 6.4

6.2 Adding a second component

One potential limitation of the analysis performed through Chapters 4 and 5 is the assumed parametric form of the $R_{\rm MS}$ distribution. Whilst some studies have found that $R_{\rm MS}$ is appropriately modelled by a singular log-normal distribution (e.g., Chang et al.,

2015; Mullaney et al., 2015; Caplar & Tacchella, 2019; Popesso et al., 2019a,b), others have used an additional component to account for the secondary bump attributable to starburst galaxies (e.g., Rodighiero et al., 2011; Sargent et al., 2012; Schreiber et al., 2015). An initial way to improve the modelling approach taken previously is, therefore, to move from a single component $R_{\rm MS}$ model to a multi-component one. In this section, we describe a possible implementation of a two-component model which would allow for a more appropriate modelling of the individual contributions from main sequence and starburst galaxies.

6.2.1 Density and likelihood function

To account for an additional starburst bump we use a mixture distribution (i.e., the combination of two separate distributions). For our model, we use the sum of two log-normal distributions: a main-sequence log-normal component controlled by $\mu_{\rm MS}$ and $\sigma_{\rm MS}$ and a starburst log-normal component controlled by $\mu_{\rm SB}$ and $\sigma_{\rm SB}$. In order to control the relative contribution from each component, we set the proportion of the contribution from the main-sequence component as ω and thus the starburst component contribution is given by $1-\omega$ (where ω is bound between 0 and 1 to ensure that the resulting model is still a probability distribution). The parameter ω can thus be used to estimate the fraction of starburst galaxies that are not explained by the singular main-sequence component. The PDF of the mixture distribution is given by the weighted (by ω) sum of the two log-normal PDFs:

$$\begin{aligned} f(\log_{10}(R_{\rm MS})|\omega,\mu_{\rm MS},\mu_{\rm SB},\sigma_{,}{\rm MS}\sigma_{\rm SB}) = \omega N(\log_{10}(R_{\rm MS})|\mu_{\rm MS},\sigma_{\rm MS}) \\ + (1-\omega)N(\log_{10}(R_{\rm MS})|\mu_{\rm SB},\sigma_{\rm SB}). \end{aligned} \tag{6.1}$$

where $N(\log_{10}(R_{\rm MS})|\mu_{\rm MS}, \sigma_{\rm MS})$ is the log-normal PDF from the main sequence component and $N(\log_{10}(R_{\rm MS})|\mu_{\rm SB}, \sigma_{\rm SB})$ is the log-normal PDF from the starburst component. This model has, compared to the singular log-normal model that was used in previous chapters, 3 additional parameters (i.e., the weight parameter ω and the parameters of the additional log-normal starburst component $\mu_{\rm SB}, \sigma_{\rm SB}$). An example of this model is shown in Figure 6.1, where we show the main sequence and starburst components separately. In a similar way, the CDF, which would be used for upper limits on $R_{\rm MS}$, can be expressed as the weighted sum of the two log-normal component CDFs:

$$\begin{aligned} F(\log_{10}(R_{\rm MS})|\omega, \mu_{\rm MS}, \mu_{\rm SB}, \sigma_{\rm MS}, \sigma_{\rm SB}) = \omega \int_{-\infty}^{R_{\rm MS}} N(\log_{10}(X)|\mu_{\rm MS}, \sigma_{\rm MS}) dX \\ + (1-\omega) \int_{-\infty}^{R_{\rm MS}} N(\log_{10}(X)|\mu_{\rm SB}, \sigma_{\rm SB}) dX. \end{aligned} \tag{6.2}$$

The likelihood function, L, is then, as usual, the product of the PDF for the detected sources multiplied by the product of the CDF for all upper limits:

$$\begin{aligned} L(\omega, \mu_{\rm MS}, \mu_{\rm SB}, \sigma_{,\rm MS}\sigma_{\rm SB}|\log_{10}(R_{\rm MS})) = \prod_{i=1}^{p} f(\log_{10}(R_{\rm MS})|\omega, \mu_{\rm MS}, \mu_{\rm SB}, \sigma_{,\rm MS}\sigma_{\rm SB}) \\ \prod_{p+1}^{n} F(\log_{10}(R_{\rm MS})|\omega, \mu_{\rm MS}, \mu_{\rm SB}, \sigma_{,\rm MS}\sigma_{\rm SB}), \end{aligned} \tag{6.3}$$

where $1, ..., .m$ correspond to detected sources and $p+1, ..., n$ are upper limits.

6.2.2 Functional relationships

The two-component model we have constructed would allow for more detailed modelling of the starburst galaxies' contribution to the $R_{\rm MS}$ distribution. However, as a result of now having five parameters, rather than just two, there are more parameters that could

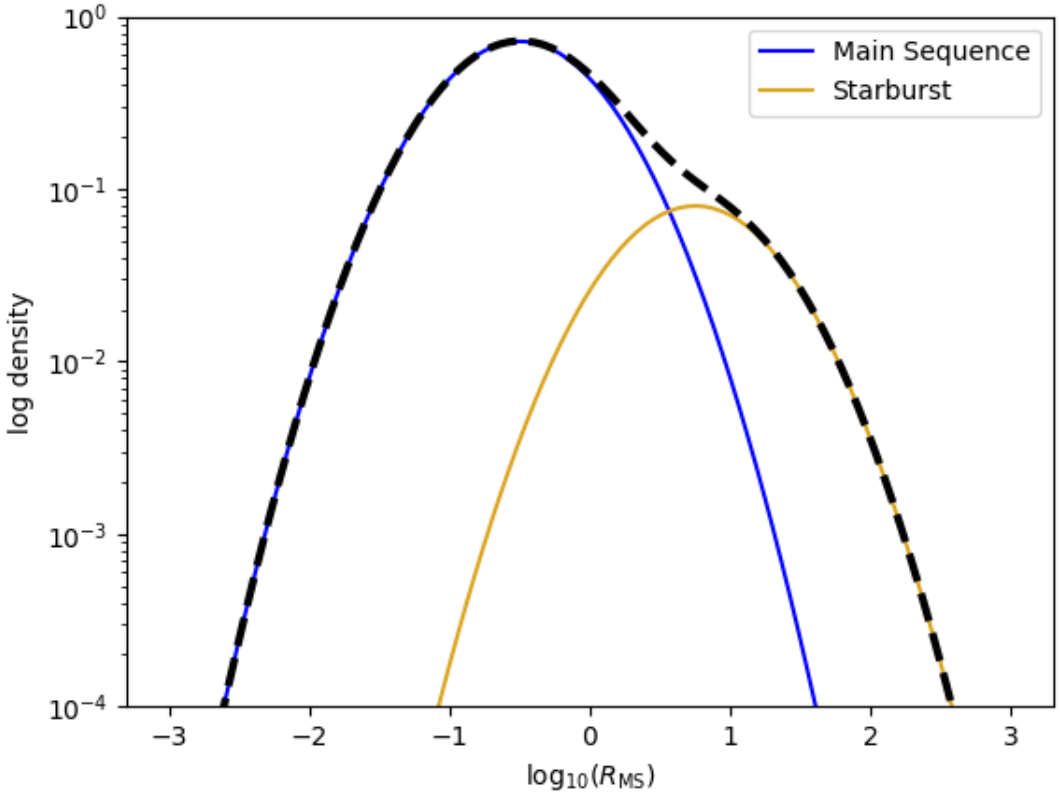


Figure 6.1: An example of the two-component log-normal model. The normalisations of the two components are controlled by the weight parameter ω, which in this case is set at 0.8. This corresponds to 80% of the density being accounted for by the main sequence component and 20% coming from the starburst component.

potentially change as a function of $L_{\rm X}$. As a starting point, given that we have already investigated the relationship between the parameters of the log-normal distribution and $L_{\rm X}$ in this thesis, we choose to investigate the relationship between ω and $L_{\rm X}$ and fix the mode and width of the two Gaussian components to the values presented in Schreiber et al. (2015, i.e., $\mu_{\rm MS} = -0.06, \mu_{\rm SB} = 0.72$ and $\sigma_{\rm MS} = \sigma_{\rm SB} = 0.31$). As ω represents the proportional contribution from the main sequence component, it is bound between 0 and 1. If ω fell outside this range, it would lead to negative PDF values for either the main sequence component (if $\omega < 0$) or the starburst component (if $\omega > 1$). Therefore, the functional relationship chosen must transpose $L_{\rm X}$ values to the [0,1] interval[1]. We therefore propose a functional relationship of the form:

$$\omega_i = \frac{1}{1 + e^{-(k \log_{10}\left(\frac{L_{\rm X,i}}{10^{40}}\right) + m)}}. \tag{6.4}$$

where k and m thus become parameters of interest (i.e., k describes how ω changes with $L_{\rm X}$and m is a coefficient term and we wish to find the best fitting k and m values). This function ensures that ω is bound between 0 and 1, for any values of $L_{\rm X}$, k or m. In Figure 6.2, we illustrate four different $R_{\rm MS}$ distributions for various values of k and m and how, according to Equation 6.4 they would change with different $L_{\rm X}$ values. It follows that if $k = 0$ then there is no evolution of ω with $L_{\rm X}$.

6.2.3 Findings from testing the two-component model

Simulated data

Thus far in this section we have outlined an extension of the one-component log-normal model that was used to describe the $R_{\rm MS}$ distribution in Chapters 4 and 5, such that a secondary component can be added, to account for additional contribution from star-

[1]It is possible to impose a bounded prior distribution of ω to account for this too. However, this would likely make the MCMC sampler – which given the complexity of the model and large number of parameters is already likely to be slow – incredibly inefficient.

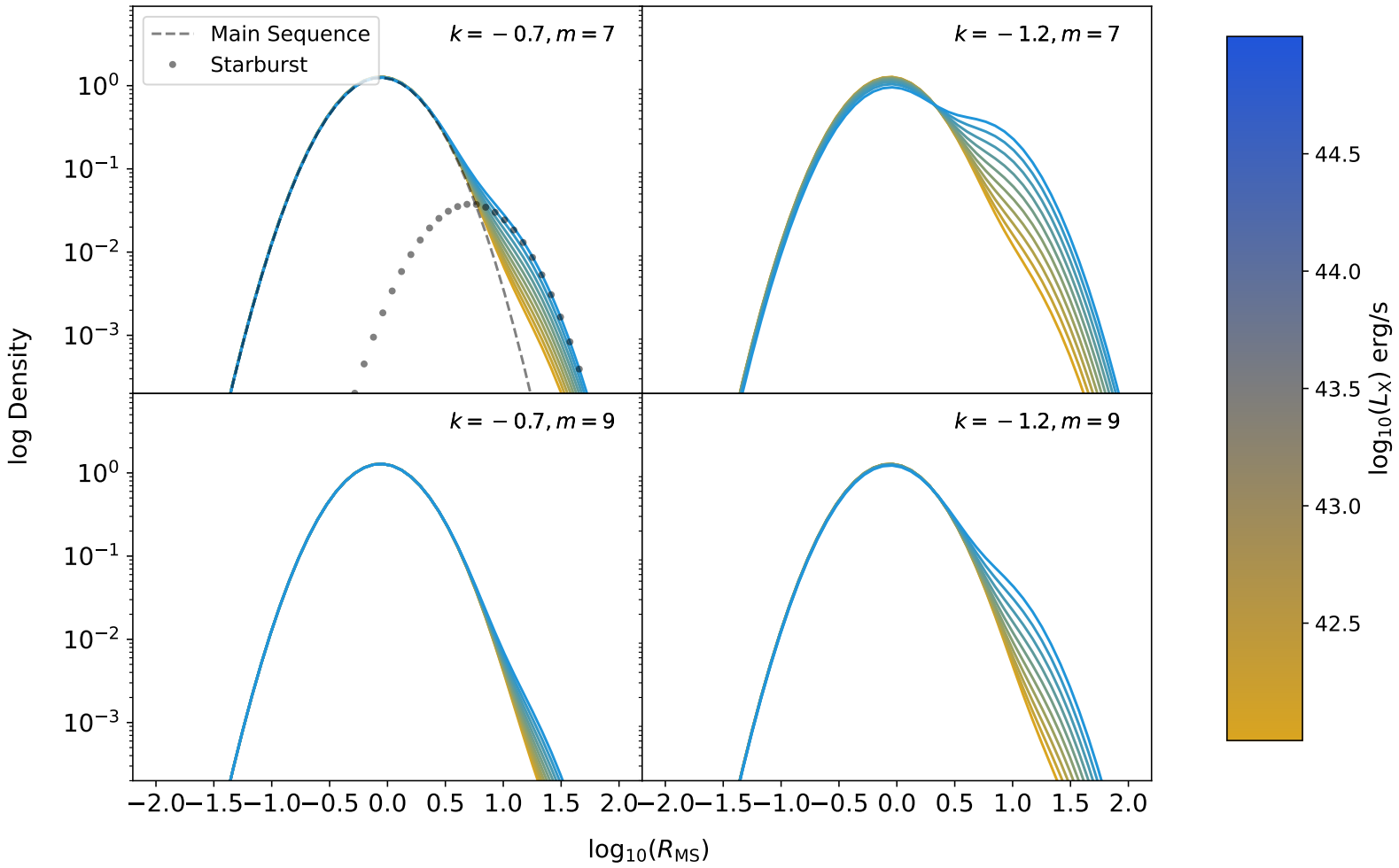


Figure 6.2: Four examples of possible distributions by assuming the functional relationship between the weight parameters ω_i and L_{X} as outlined in Equation 6.4, which depend on two further parameters k and m. The mode and width of the main sequence and starburst components are fixed according the values found in Schreiber et al. (2015, i.e., $\mu_{\mathrm{MS}} = -0.06, \mu_{\mathrm{SB}} = 0.72$ and $\sigma_{\mathrm{MS}} = \sigma_{\mathrm{SB}} = 0.31$).

burst galaxies. In order to test the capabilities of the two-component model, in this subsection, we test the model on a simulated dataset. The simulated data is generated using predetermined parameter values and, given the truth is known, is appropriate for initially testing the two-component model. Our simple data generation and testing algorithm is as follows:

1. We start by generating the independent data $L_{\rm X}$.

2. We then assume the form of the functional relationship between ω and the newly-generated $L_{\rm X}$ data. Next, we fix the values for the hyperparameters (k and m) to predetermined values (in this case we choose $k = -1.2$ and $m = 7$).

3. Given a particular $L_{\rm X}$ value and assumed functional relationships, we can accurately derive the parameters of the two-component $R_{\rm MS}$ distribution, from which we simulate an $R_{\rm MS}$ value.

4. Adopt a simple MCMC algorithm in order to attempt to recover the true values of k and m.

During our testing we generated $L_{\rm X}$ values from both a log-normal distribution (as seen in Kauffmann & Heckman, 2009) and a power law with exponential cut-offs (as seen in Aird et al., 2012). We find that the distribution used for generating $L_{\rm X}$ values has little impact on the ability of the MCMC sampler to recover the parameters of interest and thus choose to use the power law with exponential cut-offs. In order to investigate the performance of the MCMC sampler with sample size, we generated samples with size 100, 1000 and 10,000. The generated $R_{\rm MS}$ distributions and $L_{\rm X}$ distributions for the three samples are shown in Figures 6.3, 6.4 and 6.5.

After deriving the likelihood function from our simulated data, we use MCMC techniques in order to find the parameter posterior distributions. The traceplots of the MCMC algorithm for all three samples are shown in Figure 6.6, which highlights the

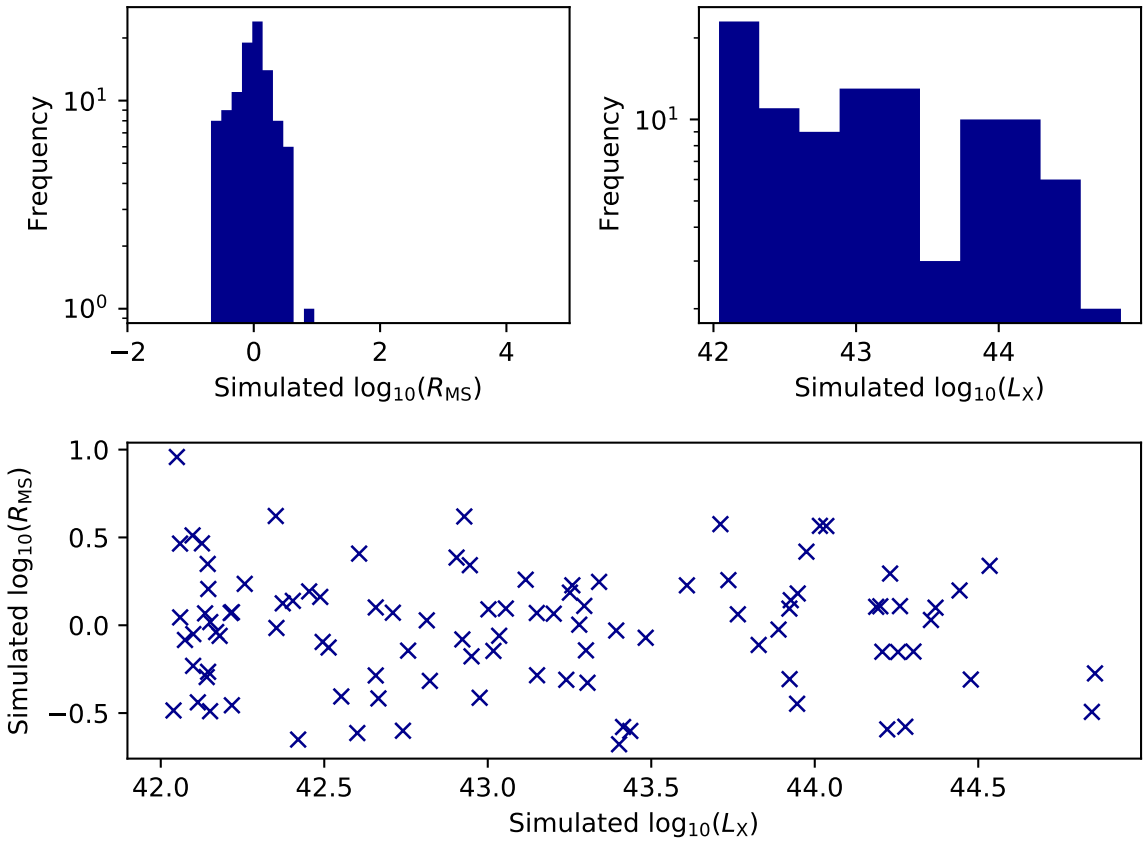


Figure 6.3: The simulated $R_{\rm MS}$ and $L_{\rm X}$ distributions for a sample size of 100. The $L_{\rm X}$ values are generated from a Schechter like function and the $R_{\rm MS}$ values generated using the functional relationships shown in Equation 6.4, alongside the $L_{\rm X}$ values. The true values for the hyperparameters are $k = -1.2$ and $m = 7$.

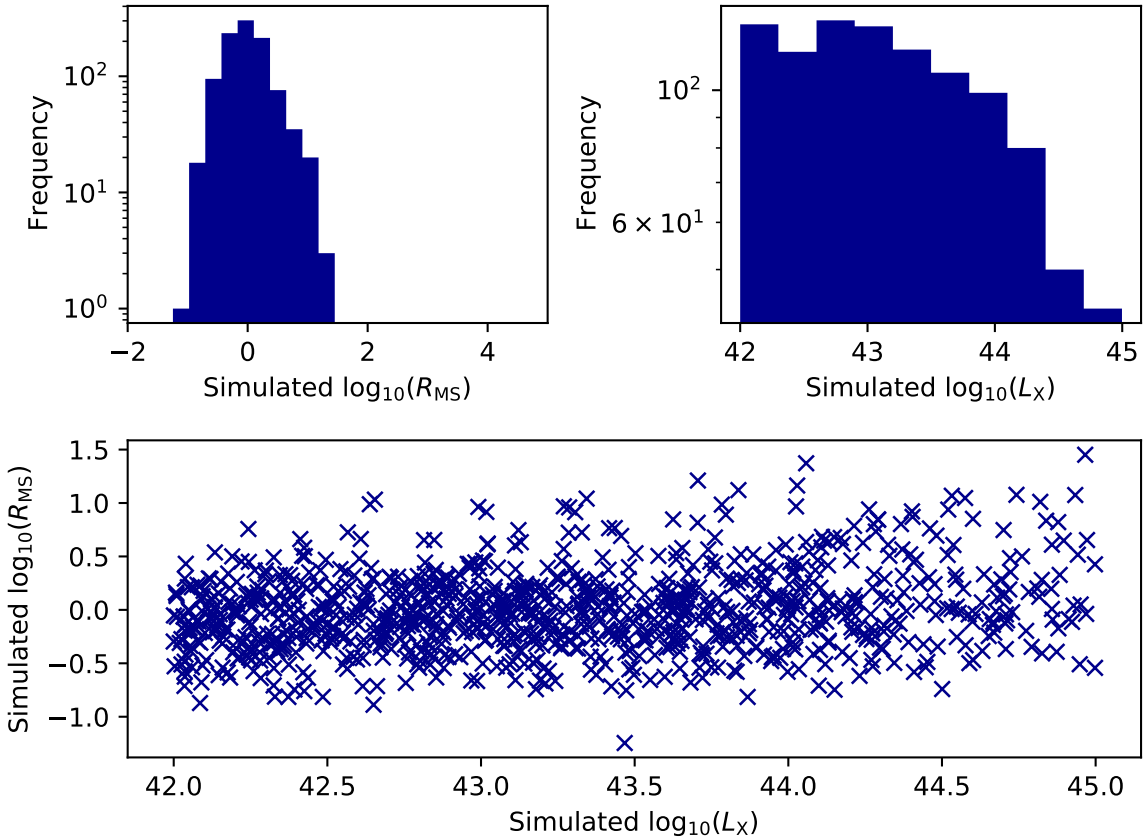


Figure 6.4: The same as Figure 6.3, but for a sample size of 1000.

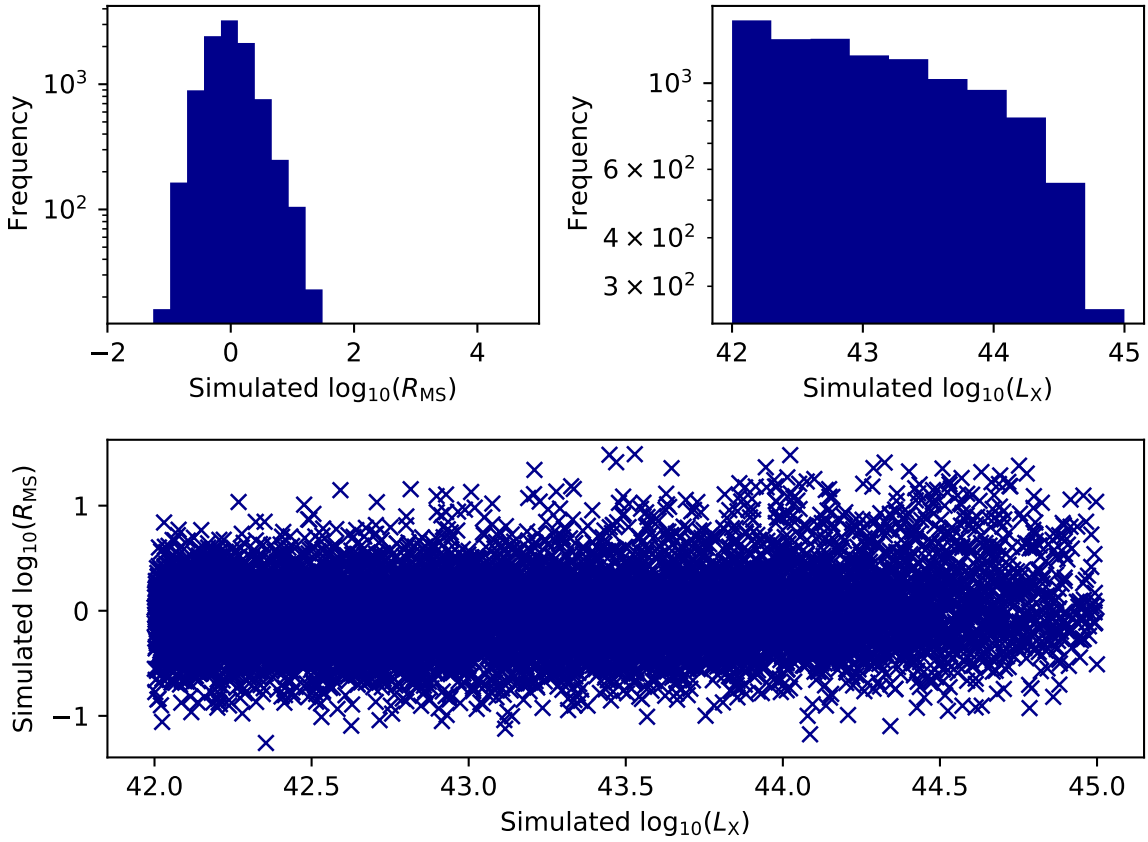


Figure 6.5: The same as Figure 6.3, but for a sample size of 10000.

performance of the sampler for the three different samples. For a sample size of only 100, the sampler recovers very little information about the true input parameter values for k and m (meaning very little information is coming from the data). However, the true input parameters values are better recovered for samples of 1000 and 10000. Indeed, very precise results are obtained from the simulated data using a sample size of 10000, demonstrating that the MCMC sampler can, with adequate data, perform well. However, it should be noted that, these samples are generated with 100% detections on $R_{\rm MS}$ (generating upper limits on $R_{\rm MS}$ is not feasible as the $R_{\rm MS}$ upper limit distribution is not known) and no uncertainty is included on $L_{\rm X}$. These simulated samples, therefore, represent the very best case scenarios. Should upper limits on $R_{\rm MS}$ and uncertainty on $L_{\rm X}$ be included in the analysis, resolving any true differences in the population is likely to require an even larger data set.

Real data

The natural progression from testing the two-component model on simulated data is to apply it to real data. The use of real data means we can include meaningful upper limits on $R_{\rm MS}$ and true uncertainties on $L_{\rm X}$. Therefore, in order to test the two-component model further, we reuse the sample of 541 AGNs as used in Chapters 4 and 5 and apply the functional relationships and the likelihood function that we previously outlined in this section. We are still only interested in how ω (i.e., the proportional contribution from the main sequence component to the $R_{\rm MS}$ distribution) changes with $L_{\rm X}$, thus we keep the other parameters fixed ($\mu_{\rm MS} = -0.06, \mu_{\rm SB} = 0.72$ and $\sigma_{\rm MS} = \sigma_{\rm SB} = 0.31$).

The posterior distributions for k and m are shown in the on-diagonal plots of Figure 6.7 and the joint distribution is shown in the off-diagonal plot. We find the mean posterior values are given by $k = 0.26 \pm 1.11$ and $m = -3.52 \pm 3.55$. As k describes the change in ω with $L_{\rm X}$ (see the functional relationship in Equation 6.4), this is arguable the most interesting parameter. By construction of the functional relationship, a positive

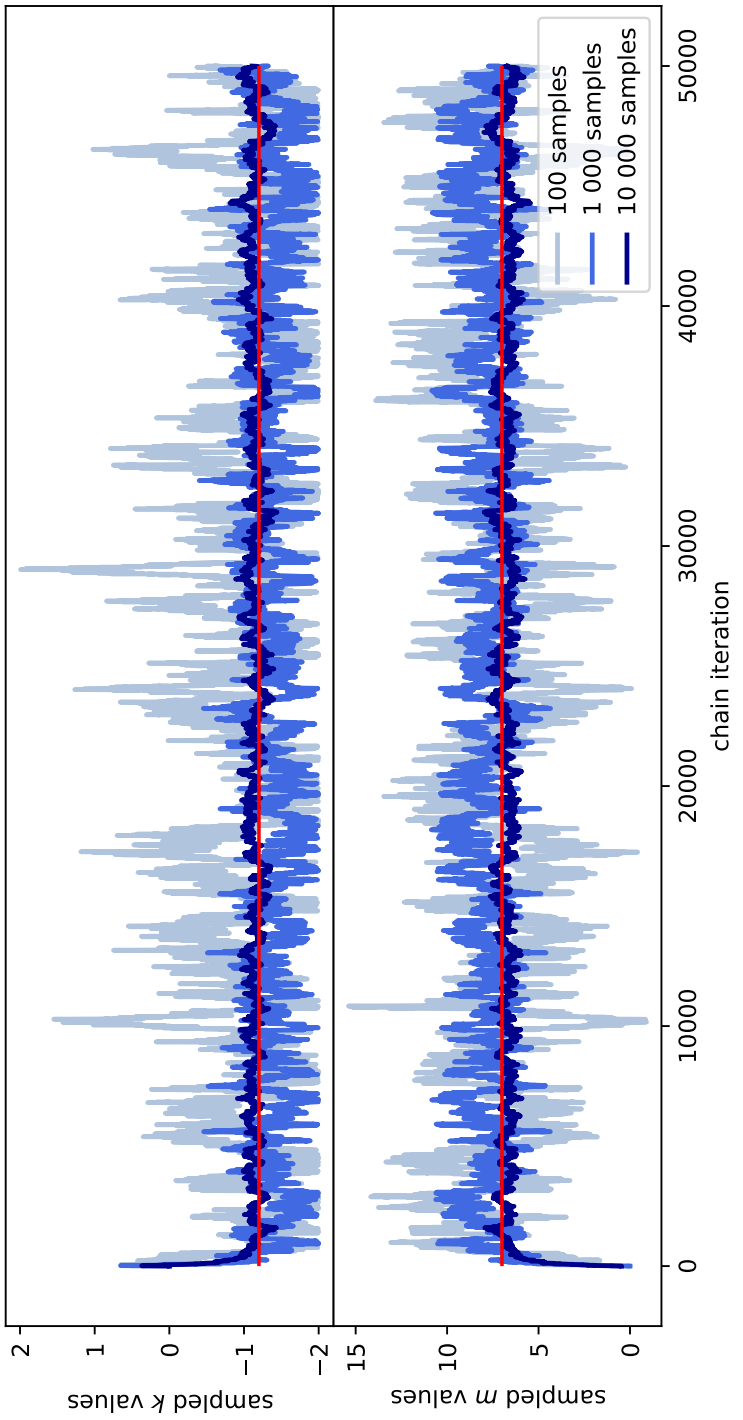


Figure 6.6: The chain output for the MCMC algorithm for both the hyperparameters k and m as a function of the sample size. The input parameter values are $k = -1.2$ and $m = 7$ and are shown as horizontal red lines. As is to be expected, the chain performs better for larger sample sizes, but note that the data is not informative enough to accurately describe the parameters for a sample size of only 100 detections. As this is simulated data, it does not include upper limits on R_{MS} nor does it include uncertainties on L_{X}, thus these are likely the best case scenarios for each sample size.

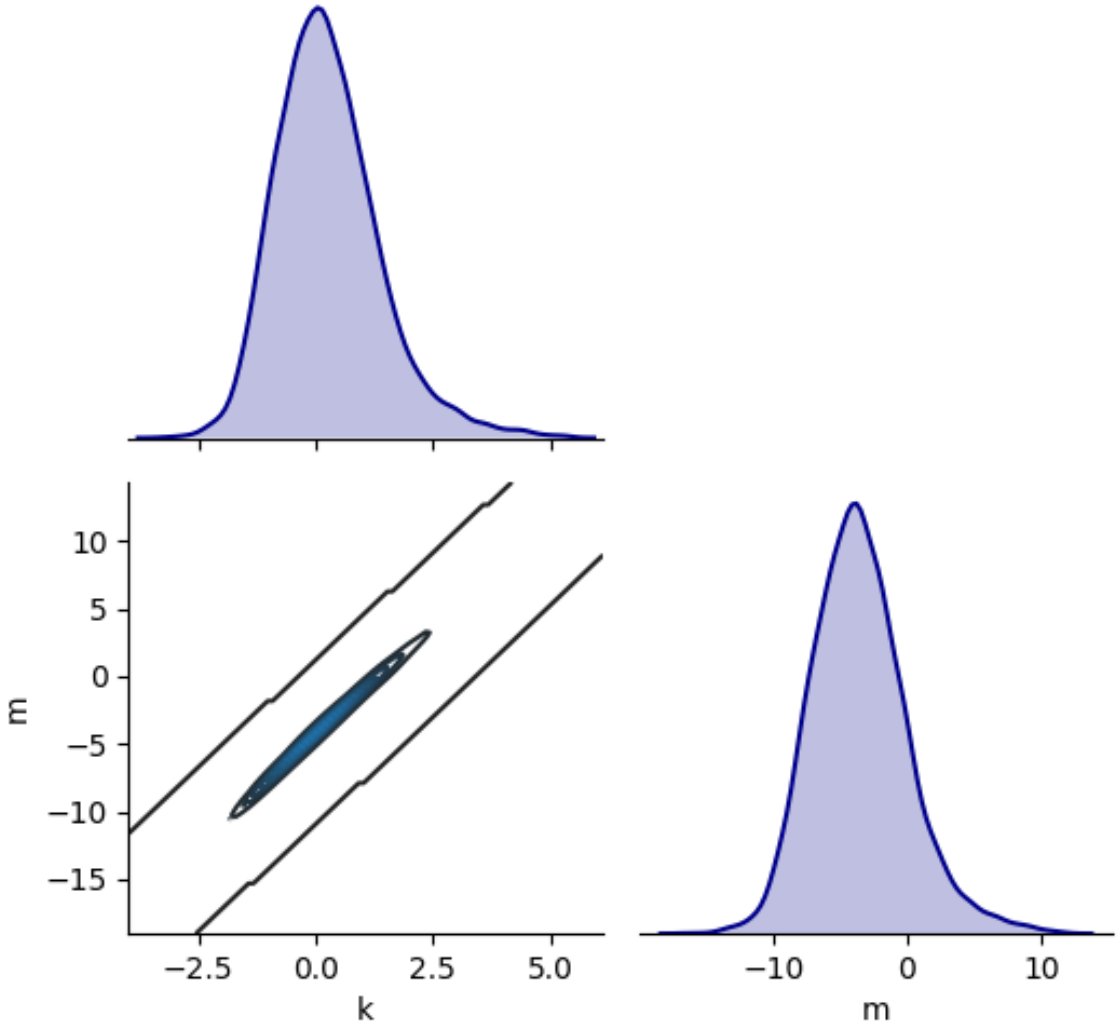


Figure 6.7: The posterior distributions for k and m as derived using the data from Chapters 4 and 5 and the functional relationship presented in Equation 6.4. The posterior distributions for k and m are highly correlated, which is likely a result of the very small range of ω values that accurately fit the data.

value of k implies that as $L_{\rm X}$ increases, the main sequence contribution increases (and thus the starburst contribution decreases). In Figure 6.8, we plot the distribution of ω for the whole sample assuming both the posterior mean and median values for both k and m. Our results show that, depending on $L_{\rm X}$, the range of fractional contribution of the main sequence component to the $R_{\rm MS}$ distribution is between 98.4% and 99.2%, thus meaning that the starburst component's contribution is between 0.8% and 1.6%. The fact that the posterior range of values for ω is quite small helps explain the tight correlation witness between the posterior distributions of k and m as seen in the off-diagonal plot of Figure 6.7 (i.e., assuming a fixed value for k, there is a very small range of m values that would give reasonable values for ω). Finally in Figure 6.9, we plot the two-component $R_{\rm MS}$ posterior distributions as a function of $L_{\rm X}$ for different values of our parameter posterior distribution (i.e., the mode, the median and the 25th and 75th percentiles). For comparison, we also plot the non-AGN two-component model results from Schreiber et al. (2015).

At first glance it may appear that our mean and median posterior $R_{\rm MS}$ distributions suggest two main results. Firstly, as $L_{\rm X}$ increases, the fractional contribution of the starburst component to the total $R_{\rm MS}$ distribution decreases. This would thus imply that the more powerful an AGN is the less likely its host galaxy is a starburst and the more likely it is to reside on the main sequence. Secondly, when compared to the distribution of non-AGN from Schreiber et al. (2015), it appears that there is a significantly larger contribution from the starburst component in non-AGN galaxies, when compared to our AGN sample. However, as aforementioned, the posterior distributions for k and m have large uncertainties (most probably due to the small sample size, consideration of $R_{\rm MS}$ upper limits and the inclusion of $L_{\rm X}$ uncertainties). As demonstrated in the bottom two plots of Figure 6.9, at the 25th percentile and 75th percentile of the posterior distribution the aforementioned results do not necessarily hold. At the 25th percentile, the value of k is negative, and thus the main sequence contribution decreases (thus

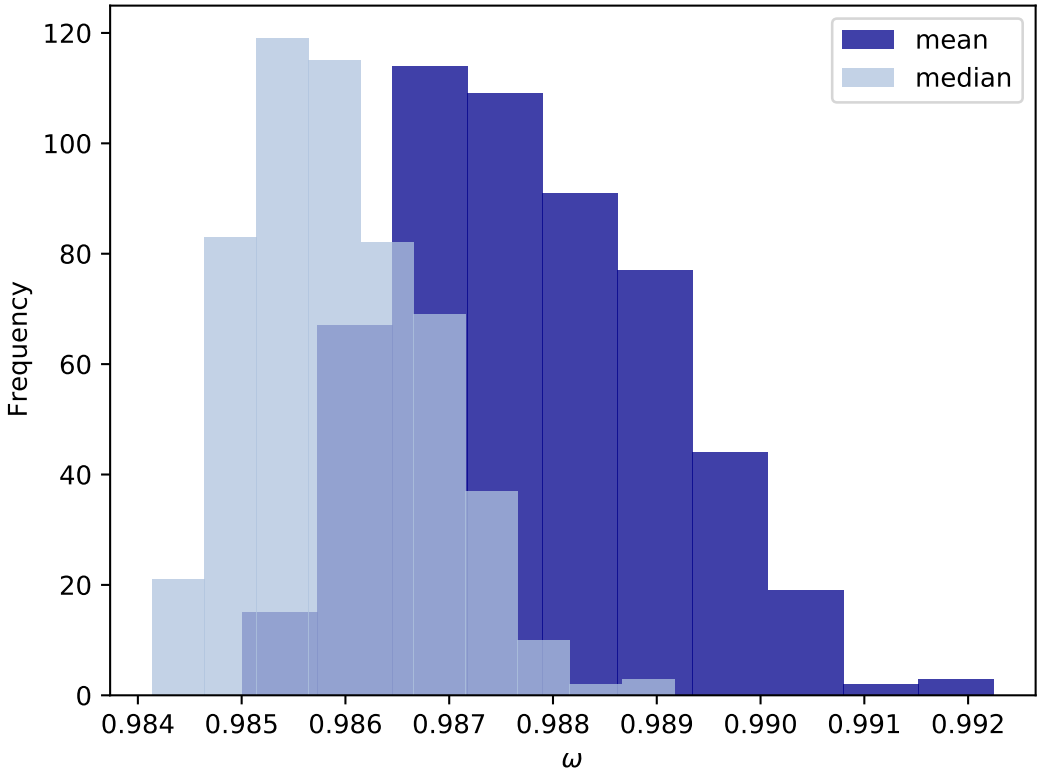


Figure 6.8: The distribution of ω (i.e., the fractional contribution to the R_{MS} distribution of the main sequence component) assuming the posterior mean and posterior median for k and m. The contribution of the main sequence to the total R_{MS} distribution varies (depending on L_{X}) between 98.4% and 99.2%.

the starburst component increases) as $L_{\rm X}$ increases. Additionally, at both the 25th and 75th percentiles, the fractional contribution of the starburst component is more aligned with the findings of Schreiber et al. (2015). So whilst the posterior mean (and median) suggest that higher-luminosity AGNs preferentially reside in main sequence galaxies when compared to lower-luminosity AGN and that AGNs in general seem to underrepresented in starburst galaxies, the uncertainties on our posterior distributions mean that neither of these claims can be made with strong evidence.

Instead, this example with real data reaffirms the findings of the work that used simulated data. In order to reveal the true evolution of the two-component $R_{\rm MS}$ distribution with $L_{\rm X}$, much larger samples are required. A sufficiently increased sample would likely reveal the true aforementioned connection but also may allow further investigation of the other parameters, which we have fixed here (i.e., $\mu_{\rm MS}, \mu_{\rm SB}, \sigma_{\rm MS}, \sigma_{\rm SB}$). There is, however, a further discussion to be held about whether the additional complexity in this model (i.e., adding a second component) is statistically appropriate (or indeed beneficial), given that the most likely starburst contribution found from this sample is between $1 - 1.5\%$. Therefore, any immediate study progressing with the two-component model should perform similar model testing as was demonstrated in Chapter 5 to compare whether the two-component form of the model is significantly advantageous over the previously used one-component form.

In addition to the limitations of sample size and whether the additional starburst component is necessary, the fixing of the mode and width parameters of the main sequence and starburst components also provide an additional limitation of this two-component model. Unlike in the two previous chapters, the fixing of the component parameters restricts the model such that it can no longer accurately account for the contribution to the $R_{\rm MS}$ distribution from AGN in quiescent galaxies. In the previous two works, the flexibility of the width parameter allowed the model to account for quiescent galaxies by extending the tails of the distribution down to lower $R_{\rm MS}$ values. However, in this

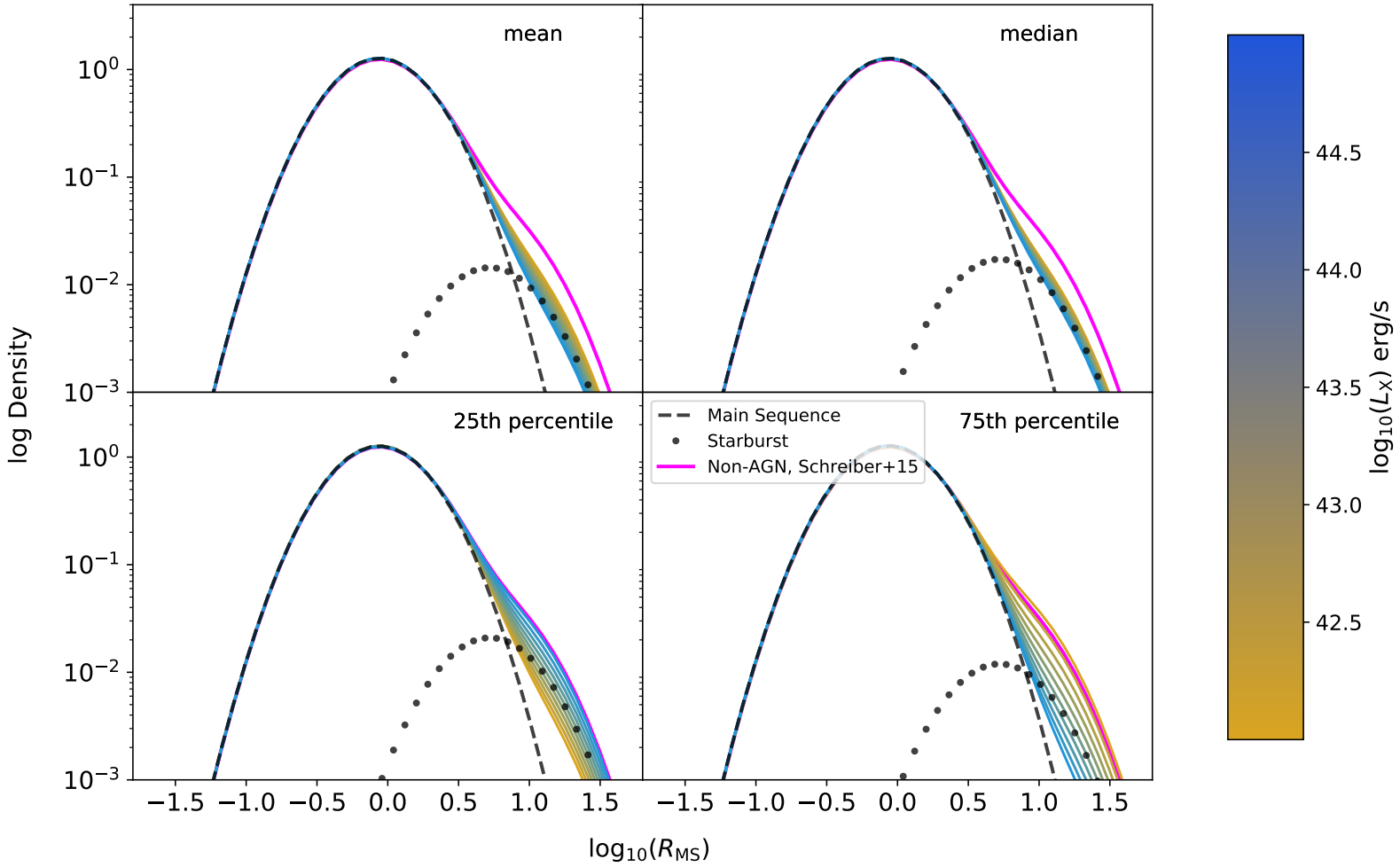


Figure 6.9: Posterior $R_{\rm MS}$ distributions derived using the k and m values at the posterior mean (top left), posterior median (top right), 25th percentile (bottom left) and 75th percentile (bottom right) using the sample of 541 AGNs used in Chapters 4 and 5.

example, the widths are kept fixed to specifically investigate the contribution to the $R_{\rm MS}$ distribution of the starburst galaxies. As AGNs are known to reside in quiescent galaxies, it is important to account for their contribution. Therefore, allowing the mode and width parameters to vary should be a priority in extending this model. However, an alternative approach (motivated by the fact the starburst contribution in our example is restricted to $1 - 1.5\%$), would be to use the two-component model to instead model the quiescent galaxies' contribution to the $R_{\rm MS}$ distribution. This asks a slightly different science question, which is beyond the remit of this thesis, but would be a more appropriate use of the two-component model (at least one with fixed mode and width).

6.3 Upgrading the functional relationships

The use of functional relationships to connect the parameters of the $R_{\rm MS}$ distribution to independent data allows us to investigate the continuous relationship between star-forming properties of host galaxies and AGN power. As well as also removing the need to construct $L_{\rm X}$ bins, they allow us to account for the uncertainty on $L_{\rm X}$, by considering $L_{\rm X}$ as a parameter with a prior (or proposal) distribution that is constrained by the measured X-ray luminosity and uncertainty. This method effectively samples a range of possible $L_{\rm X}$ values, throughout the lifetime of the MCMC sampler, thereby carrying through the measured uncertainty. The form of the functional relationships chosen thus far in this thesis (and in particularly for the one-component model) are, however, arbitrarily chosen and have little motivation other than their simplicity (i.e., for the mode, μ and width, σ, a simple linear dependence was chosen between μ on $L_{\rm X}$ and an exponential dependence was chosen between σ and $L_{\rm X}$). In this section, we discuss the most insightful improvements that can be made to these functional relationships.

6.3.1 Adding stellar mass and redshift

Whilst model comparison techniques could be used to investigate the best-fitting functional relationships, we believe the most insightful change that could be made to the functional relationships would be the inclusion of other independent data such as redshift (z) and stellar mass (M_*). By including other such independent data, the ultimate goal is to investigate whether the connection identified in Chapter 5 holds across a variety of redshift ranges, and down to lower stellar masses. Reverting back to a single log-normal distribution (for simplicity) with its peak and width controlled by μ and σ as before, a potential new pair of functional relationships (which we refer to as Set A) could be of the form:

$$\begin{aligned} \mu_i &= \theta_0 + \theta_1 \log_{10}\left(\frac{L_{\mathrm{X,i}}}{10^{40}}\right) + \theta_2 z + \theta_3 \log_{10}\left(\frac{M_*}{10^9}\right), \\ \sigma_i &= \exp\left\{\theta_4 + \theta_5 \log_{10}\left(\frac{L_{\mathrm{X,i}}}{10^{40}}\right) + \theta_6 z + \theta_7 \log_{10}\left(\frac{M_*}{10^9}\right)\right\}. \end{aligned} \tag{6.5}$$

These two equations attempt to model how the parameters of the R_{MS} distribution change with L_{X}, z and M_*. However, whilst functional relationships of this form would allow us to investigate how the R_{MS} distribution changes as a function of L_{X}, z or M_* independently, it would not allow for a comparison of how the *relationship between* R_{MS} *and* L_{X} changes with z and M_*. For example in this thesis, we are less interested in how R_{MS} changes with redshift, but more interested in how the connection between R_{MS} and L_{X} changes with redshift (or put candidly, is the strength of the connection we witness in Chapter 5 specific to our mass or redshift choices).

To investigate how the relationship between R_{MS} and L_{X} changes with redshift, we need to introduce a term into the functional relationship that is codependent on both L_{X} and either z or M_*. Instead of Set A, therefore, for the questions posed in this thesis functional relationships of the following form are more appropriate (we will refer to these

as Set B):

$$\mu_i = \theta_0 + \theta_1 \log_{10}\left(\frac{L_{\rm X,i}}{10^{40}}\right) + \theta_2 z \log_{10}\left(\frac{L_{\rm X,i}}{10^{40}}\right) + \theta_3 \log_{10}\left(\frac{M_*}{10^9}\right) \log_{10}\left(\frac{L_{\rm X,i}}{10^{40}}\right)$$

$$\sigma_i = \exp\left\{\theta_4 + \theta_5 \log_{10}\left(\frac{L_{\rm X,i}}{10^{40}}\right) + \theta_6 z \log_{10}\left(\frac{L_{\rm X,i}}{10^{40}}\right) + \theta_7 \log_{10}\left(\frac{M_*}{10^9}\right) \log_{10}\left(\frac{L_{\rm X,i}}{10^{40}}\right)\right\}. \tag{6.6}$$

The difference between Set A and Set B is that in Set B, the data controlled by the parameters $\theta_2, \theta_3, \theta_6$ and θ_7 changes to have an additional $L_{\rm X}$ term. In Set A, θ_2 and θ_3 describe the change of μ with redshift and stellar mass respectively, and θ_6 and θ_7 describe the change of σ with redshift and stellar mass respectively. Whereas in Set B, θ_2 describes how μ changes with both redshift and $L_{\rm X}$ codependently (i.e., how the relationship between μ and $L_{\rm X}$ changes with redshift) and θ_3 describes how μ changes with both stellar mass and $L_{\rm X}$ codependently (i.e., how the relationship between σ and $L_{\rm X}$ changes with stellar mass). Similarly, θ_6 describes how σ changes with both redshift and $L_{\rm X}$ codependently (i.e., how the relationship between μ and $L_{\rm X}$ changes with redshift) and θ_7 describes how σ changes with both stellar mass and $L_{\rm X}$ codependently (i.e., how the relationship between σ and $L_{\rm X}$ changes with stellar mass). In Set A, if θ_2 is positive, then μ increases with redshift, regardless of $L_{\rm X}$. In Set B, if θ_2 is positive, then μ increases with $L_{\rm X}$ more at higher redshifts (and vice versa). This means that, if θ_2 is non-zero, there is a change in the relationship between $R_{\rm MS}$ and $L_{\rm X}$ as a function of redshift. This applies to θ_3 and stellar mass, and similarly to θ_6, θ_7 with σ. It should also be noted that here we have chosen simple dependencies (i.e., linear for μ and exponential for σ) to relate our parameters of the $R_{\rm MS}$ distribution to all independent variables. Again, similar to the motivation in Chapter 5, these relationships are unlikely to represent the true scenario (i.e., they have little physical motivation). Instead, they

are suitable for testing whether the relationship between $R_{\rm MS}$ and $L_{\rm X}$ changes with z and M_*, or not. It remains the case that it is not immediately clear why if compared to the null case (i.e., the relationship between $R_{\rm MS}$ and $L_{\rm X}$ does not change with z or M_*) a dependent model (i.e., one using Set B), would be preferred over an independent one (i.e., the functional relationships used in Chapter 5, where we do not account for z or M_*).

6.3.2 Including upper limits on $L_{\rm X}$

The use of functional relationships provides us with two significant benefits. Firstly, the use of arbitrarily constructed bins is no longer required and secondly, uncertainties on independent data can be accurately considered throughout the analysis (as was seen in Chapter 5). Removing the need to bin our data means that we no longer need an AGN-selected sample or a star-forming galaxy-selected sample, as we do not need a binning axis (which is a requirement of the binning-and-averaging approach seen in the wider literature, e.g., Rafferty et al., 2011; Harrison et al., 2012; Rosario et al., 2012; Chen et al., 2013). However, if we want to investigate how the $R_{\rm MS}$ distribution changes with SMBH accretion rate for the wider galaxy population, we need to be able to include upper limits on $L_{\rm X}$. This would allow us to probe the lower X-ray luminosity regime. Recall that in Section 3 upper limits on s$L_{\rm X}$ were included as we were investigating the s$L_{\rm X}$ distribution, but given in Chapter 5 we investigated the $R_{\rm MS}$ distribution as a continuous function of $L_{\rm X}$, the question raised here is: how can we include $L_{\rm X}$ upper limits when it is being used as an independent variable?

In Chapter 5, we appropriately consider uncertainties on $L_{\rm X}$ by sampling it through the MCMC algorithm with a sampling distribution described by the measured value and uncertainty. This means, throughout the lifetime of the MCMC chain, for an individual source, various $L_{\rm X}$ values are considered, representing the range of possible true $L_{\rm X}$

values. This approach is flexible and can thus be changed for those $L_{\rm X}$ values for which we do not have a detection and instead have only an upper limit. The crux is to ensure that $L_{\rm X}$ values proposed throughout the chain are plausible and realistic values. Therefore, one can either propose random values (that are constrained by the prior) or directly influence the proposal distributions, such that only realistic $L_{\rm X}$ values are proposed. An immediately obvious choice for a sampling (or prior) distribution is to adopt the AGN X-ray luminosity function below the upper limit - although this does correspond to the uninformative prior case (i.e., there is no source-specific information and instead we resort to the population's characteristics) and would require an assumed lower turnover. More accurate distributions could be derived based on corrections from other wavelengths (such as the UV-X-ray luminosity ratio, Lusso et al., 2010), which would allow some source-specific knowledge to be introduced. Throughout the course of the MCMC chain, possible $L_{\rm X}$ values will be proposed and therefore the upper limit can be included in the functional relationships and thus the analysis.

6.4 Discussion and Conclusions

In this chapter we have provided a series of potential extensions to the modelling approach taken in Chapters 5. We introduce two beneficial improvements: the movement from a one-component model to a two-component one and the inclusion of other independent data (such as stellar mass and redshift). Firstly, we modified the parametric form used to model the $R_{\rm MS}$ distribution such that an additional component could be included to account more accurately for any potential excess contribution from starburst galaxies. In doing so, we tested the two-component model on a simulated dataset with known true parameters and identified that, even in the best case scenario, large datasets were required to begin to resolve the true underlying parameters. Extending this, we retested the two-component model on the real sample of the 541 AGNs previously used

in the analyses in Chapters 4 and 5. In doing so, we forego knowledge of the underlying true parameters and instead have a more realistic dataset (i.e., one with upper limits on $R_{\rm MS}$ and meaningful uncertainties on $L_{\rm X}$). Our findings, however, corroborate with the findings from the simulated dataset, as no significant results were identified, largely as the uncertainties on the parameter posterior distributions were far too large. Thus, should this line of work be continued in the future, substantially larger datasets will be required. We also noted that, any study wishing to use the two-component model should, at first, determine if its performance is significantly better than the one-component case described earlier in this thesis. One possible way of doing this is to adopt the model comparison techniques demonstrated in Chapter 5.

The second highlighted improvement we made in this chapter was the upgrading of the functional relationships (particularly in the one-component case) such that the relationship between $R_{\rm MS}$ and $L_{\rm X}$ can be investigated across various redshift and stellar mass ranges. By removing the need to construct bins of $L_{\rm X}$, the functional relationships provided the framework for removing redshift and stellar mass binning too. In Section 6.3, we therefore proposed a set of functional relationships (Set B) that could be used to investigate how the strength of the relation witness in Chapter 5 changes as a function of redshift and stellar mass, without repeating the analysis in different redshift bins.

Chapter 7

Discussion and Conclusions

> It's not about how hard you can hit. It's about how hard you can get hit and keep moving forward. How much you can take and keep moving forward. That's how winning is done.
>
> Rocky Balboa

7.1 Introduction

The discovery that the mass of SMBHs correlate with both their host galaxy's bulge stellar mass and stellar velocity dispersion (e.g., Magorrian et al., 1998; Ferrarese & Merritt, 2000; Gebhardt et al., 2000; Merritt & Ferrarese, 2001; Tremaine et al., 2002; Marconi & Hunt, 2003; Häring & Rix, 2004; Wyithe, 2006; Hu, 2008; Gültekin et al., 2009; McConnell & Ma, 2013; de Nicola et al., 2019; Ding et al., 2020a,b) provides evidence of an evolutionary link between the two. Further evidence of a connection between SMBH and host galaxy is also provided by the discovery that, as discussed in the Introduction (Chapter 1), the volume average SMBH accretion rate and volume averaged star formation have followed similar evolutionary tracks.

In an effort to further reveal more details about any potential connection, a large

number of studies have sought to identify whether the rate at which the SMBH has grown is correlated with the rate at which the host galaxy has grown via star formation. To do this, some studies have investigated how the average SMBH growth (calculated using stacking to include upper limits) changes as a function of SFR (e.g Rafferty et al., 2011; Chen et al., 2013; Azadi et al., 2015; Delvecchio et al., 2015; Harris et al., 2016; Lanzuisi et al., 2017; Shimizu et al., 2017; Stemo et al., 2020). These studies tend to find evidence of a positive correlation, implying that as SFR increases, so does average SMBH growth. However, averages are summary statistics and various differences in the underlying properties of SMBH growth can cause a similar increase in the average. For example, does the average SMBH growth increase because of a handful of extreme outliers? Or is a more widespread, yet less pronounced, systematic increase responsible? Therefore, in order to better understand the relationship between SMBH growth and star formation, in Chapter 3 we investigated how the full distribution of host galaxy stellar mass specific X-ray luminosity (sL_X, tracing mass-normalised accretion rate) changes between galaxies classified as starburst (i.e., a factor of three above the main sequence) and non-starburst. We find that, amongst starburst galaxies, there is an significantly increased fraction of SMBHs accreting at higher rates (i.e., greater than 10% of their Eddington limit, similar to the results seen in Aird et al., 2018, who used SFRs from UV-NIR SEDs). This means that a SMBH residing within a starburst galaxy has a greater probability of having a higher accretion rate and can therefore explain the increased average witnessed in correlation-based studies.

In addition to those aforementioned studies, an alternative approach to investigating any statistical connection between the growth rates of SMBHs and their host galaxies is to derive an AGN sample and instead investigate how the average SFR changes as a function of AGN luminosity (e.g., Harrison et al., 2012; Rosario et al., 2012; Stanley et al., 2015; Lanzuisi et al., 2017; Stanley et al., 2017; Suh et al., 2017; Ramasawmy et al., 2019). However, contrasting the results using the aforementioned approach, these studies tend to

find little evidence of a correlation. Similarly to how various differences in the underlying population can cause the same apparent increase in the average (as mentioned before), no difference in the average SFR does not necessarily imply that there is no difference in the underlying star-forming properties. Therefore, in Chapter 4, we investigated the full distribution of $R_{\rm MS}$ (i.e., the deviation from the star forming main sequence) between low luminosity AGNs (i.e., those with $L_{\rm X} < 2 \times 10^{43}$ erg s^{-1}) and high luminosity AGNs ($L_{\rm X} > 2 \times 10^{43}$ erg s^{-1}). In that chapter, we report tentative evidence of a difference in the $R_{\rm MS}$ distribution between the two samples, suggesting that higher luminosity AGNs tend to have a slightly higher peaked, yet considerably narrower $R_{\rm MS}$ distribution than lower luminosity AGNs. Interestingly, however, although we found a subtle difference in the underlying $R_{\rm MS}$ distributions, we reported that there was no obvious difference in the average of the two, which may reaffirm the importance of investigating full distributions. That said, we still only found tentative evidence of a connection between star-forming properties of the host galaxy and AGN luminosity, which does not fully unify these results with those seen when adopting the inverse approach.

Hickox et al. (2014) demonstrated that one potential explanation for the contradictory results seen depending on whether an AGN-derived sample or a galaxy-derived sample is used, could be the problems associated with binning in the highly variable AGN luminosity axis. As a result of being highly variable, it is possible that the binning process could wash-out potentially underlying correlations. However, binning data in general has, as discussed throughout this thesis, other limitations, such as considering uncertainties and implicit assumptions about sources within a bin having similar (if not the same) properties. Motivated by these problems, in Chapter 5 we presented a binning-free methodology to repeat the study performed in Chapter 4. By constructing functional relationships that directly related the parameters of the $R_{\rm MS}$ distribution to $L_{\rm X}$, we removed the need to construct AGN luminosity bins and were able to more accurately include uncertainties on – and arguably more relevantly the information from

– every data point. In that chapter, we report strong evidence (rather than just tentative) that the $R_{\rm MS}$ distribution changes as a function of $L_{\rm X}$. More specifically, we reaffirmed to a stronger degree of significance the results of Chapter 4 that the peak of the $R_{\rm MS}$ increases very slightly with $L_{\rm X}$, whilst the width decreases. Going a step further than we did in Chapter 4, we also demonstrated that, even if we assume the relationship found between $R_{\rm MS}$ and $L_{\rm X}$ to be true, the correlation observed between average SFR and $L_{\rm X}$ falls well within the scatter of the flat relationship found in the literature (e.g., Stanley et al., 2015). Whilst these results provide evidence that star-forming properties of AGNs are connected to the central SMBH accretion rate, arguably a more important finding was that these results tend to agree with studies deriving galaxy-selected samples – i.e., it appears that higher levels of SMBH growth are associated with more star-forming activity.

Throughout Chapters 4 and 5 we assumed that the parametric form of the $R_{\rm MS}$ distribution is adequately represented by a singular log-normal distribution. However, it may be the case that the two-component model can more accurately include an excess contribution from starburst galaxies (e.g., Sargent et al., 2012; Schreiber et al., 2015). Therefore, in Chapter 6 we constructed a binning-free two-component model, in which we investigated the fractional contribution from the starburst component to the total $R_{\rm MS}$ distribution. In short, we find that the model is likely to only work for large datasets, and reusing the sample of 541 AGNs from Chapters 4 and 5, we were unable to deduce any significant correlations, as the uncertainties on our posterior distributions were too large. Additionally in that chapter, we discussed the possibility of an alternative line of research, by which instead of improving the parametric form of the model (i.e., switching from one component to two), we modified the functional relationships of the singular component, such that other independent data could be included. Specifically, we highlighted a series of functional relationships that could be used to describe how the relationship between the $R_{\rm MS}$ distribution and $L_{\rm X}$ changes with either redshift or stellar

mass. By modifying the analysis in this way, future studies will be able to see whether the results seen throughout this thesis hold at earlier or later times in the Universe and down to smaller stellar masses. This is a crucial step in advancing our knowledge of how the relationship between SMBH and host galaxy applied to the entire galaxy population.

7.2 Comparison to literature

As mentioned in the introductory chapter (Chapter 1), the statistical approach is only one way to use observational data to investigate the relationship between SMBH accretion rate and star-forming properties of the host galaxy. The alternative approach is to look at fewer numbers of galaxies, but in a lot more individual detail, hoping to witness directly any connecting process between SMBH and host galaxy rather than extract it from characteristics of the population. In addition to observational approaches, there is also the theoretical approach which has, up to now, been largely ignored in this thesis. In this section we highlight two proposed connecting mechanisms that have been observed in observational studies of smaller samples or proposed in studies that take a more theoretical approach and discuss where our results agree (or disagree) with them. The two proposed connecting mechanisms we discuss are AGN feedback (i.e., a direct connection) and the co-regulation of SMBH growth and star formation by the availability of gas in the galaxy (i.e., a less direct, more codependent connection).

7.2.1 AGN Feedback

Since the 1980s, it has been suspected that the vast amount of energy produced during a SMBH accretion event could propagate through the host galaxy (e.g., Sanders et al., 1988). Many theoretical studies have predicted that this vast amount of energy would heat or expel gas en route (Magorrian et al. 1998; Silk & Rees 1998 and see the reviews by Alexander & Hickox 2012; Fabian 2012; Harrison 2017 and references therein). This

energy, termed "AGN feedback", has been thought to negatively impact star formation to such an extent that it could potentially stagnate further galaxy growth (at least in the most massive galaxies). More succinctly, by introducing the effects of AGN feedback (alongside supernova feedback for the lower stellar mass-regime) into their models, theoretical studies have managed to explain fundamental properties of the galaxy population (e.g., the observed galaxy mass function, see Figure 7.1, Silk & Rees, 1998; Bell et al., 2003; Benson et al., 2003; Di Matteo et al., 2005; Springel et al., 2005; Croton et al., 2006; Bower et al., 2006; Cattaneo et al., 2006; Hopkins et al., 2006; Panter et al., 2007; Hopkins et al., 2008; Somerville et al., 2008; Booth & Schaye, 2009; Silk & Mamon, 2012; Dubois et al., 2013; Vogelsberger et al., 2013; Schaye et al., 2015; Dubois et al., 2016; Pillepich et al., 2018). According to this scenario, the relationship between SMBH accretion and star-formation should be a negatively-correlated one, where more rapidly accreting SMBHs (or more luminous AGNs) should be suppressing star formation more than less rapidly accreting SMBHs. If this paradigm holds, AGN feedback has had a severe impact on the ability of the host galaxy to grow and has thus played large part in the way galaxies have evolved. Some studies have thus explored AGN feedback in the context of galaxy evolution and have suggested that AGN feedback could be responsible for transforming star-forming galaxies into more quiescent ones (e.g., Sanders et al., 1988; Springel et al., 2005; Sijacki et al., 2007; Booth & Schaye, 2009; Kormendy et al., 2009; Kormendy & Ho, 2013, although Jackson et al. 2019 suggested that AGN feedback is not a requirement of this transition).

Observationally, AGN feedback – mediated by AGN driven outflows – has been widely reported (e.g., Greene et al., 2012; Brusa et al., 2018). Whilst some studies do find that the presence of AGN driven outflows appears to suppress star formation (e.g., Cicone et al., 2014; Cresci et al., 2015b; Carniani et al., 2016; Cresci & Maiolino, 2018), some studies still identify strong star formation in (or certainly close to) an AGN driven outflow (e.g., Cresci et al., 2015a; Maiolino et al., 2017; Cresci & Maiolino, 2018; Gallagher et al.,

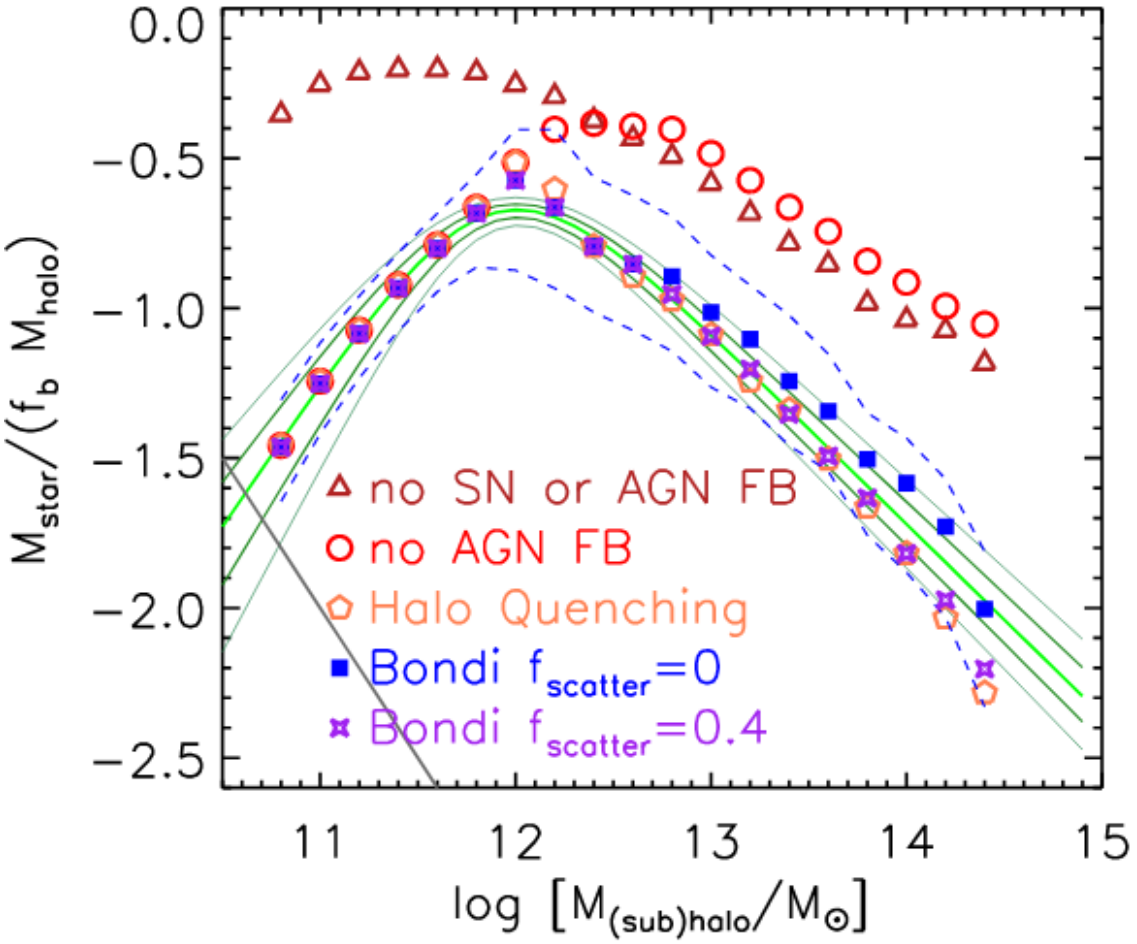


Figure 7.1: The stellar mass to halo mass ratio as a function of halo mass. Plotted in green are the empirical results found in Moster et al. (2010). All other plots show the results of different theoretical models with differences in the quenching mechanisms. The red open circles show the model without AGN feedback, where there is an excess of massive galaxies compared to the empirical values. This Figure was originally presented in Somerville et al. (2008).

2019), leaving the true effects of AGN feedback still uncertain. Additionally, if AGN feedback is to negatively impact star formation on galaxy-wide scales, the extent of the outflow must be galaxy wide. However, there is also considerable debate as to the extent at which feedback can propagate into the host galaxy (see Tadhunter et al., 2018, for a discussion). This implies that if AGN feedback is quenching star formation, it is either not instantaneous, not consistent or not on galaxy-wide scales.

Given the studies presented in this thesis compare the galaxy-wide SFR with the power of the AGN, it may initially make sense to compare our results in the context of AGN feedback. However, because any study investigating the instantaneous SMBH accretion rate with recently time-averaged star formation fails to account for the inevitable time-delay between the onset of SMBH activity and the time required to quench (or enhance) any star formation, our results provide little clarity to the AGN feedback paradigm. More specifically, it would be inaccurate to suggest our results provide evidence of positive feedback just because we observe a connection between higher levels of recent (yet historic) star formation and greater (instantaneous) SMBH activity. Therefore we stress that the results of this thesis provide little help in solving the AGN feedback paradigm. Instead, our results are more helpful at looking at the triggering and fuelling of AGNs and star-formation.

7.2.2 Gas availability

Both star formation and SMBH accretion rely on gas supply (e.g Maiolino et al., 1997). However, both of these processes require specific conditions. Firstly, stars are formed from the gravitational collapse of *cold, dense, molecular clouds* of gas and secondly, an AGN may only be triggered if gas is driven down to the innermost regions of galaxies (i.e., the vicinity of the SMBH). It is therefore logical to suggest that just because a particular galaxy hosts a large cold molecular gas reservoir, it does not necessarily

mean a guaranteed increase in star forming activity or AGN activity. Instead, a more logical proposition is that the probability of SMBH accretion and SFR is regulated by gas supply. To this extent, a tight correlation between SFR and dense molecular gas has been long identified (e.g., Schmidt, 1959; Kennicutt, 1998; Gao & Solomon, 2004; Shangguan et al., 2020a), indicating a strong dependency of star formation on the cold molecular gas content of the host galaxy. However, studies investigating the host galaxy gas fraction of AGNs have found mixed results, with some studies suggesting that AGNs reside in systems with enhanced gas fractions (e.g., Bertram et al., 2007; König et al., 2009; Yesuf et al., 2017; Shangguan et al., 2018, 2019, 2020b; Yesuf & Ho, 2020), in systems with normal levels of gas (i.e., comparable to non-AGN galaxies, e.g., Fabello et al., 2011; Xia et al., 2012; Krips et al., 2012; Villar-Martín et al., 2013; Husemann et al., 2017; Rosario et al., 2018; Ellison et al., 2019), and some in gas-poor systems (e.g., Haan et al., 2008; Brusa et al., 2015; Kakkad et al., 2017; Perna et al., 2018). However, in a very recent study that attempted to understand the interplay between gas and AGN activity and star formation, Yesuf & Ho (2020) analysed the molecular gas content of a sample of type 2 AGNs and provided an evolutionary scenario by which both star formation and SMBH accretion are "mediated" by the host galaxy gas content. Those authors suggest that in a gas-rich system (the creation of which may be due to a gas-rich merger, see Sanders et al. 1988; Di Matteo et al. 2005; Hopkins et al. 2006), vigorous star formation takes place initially, after-which stochastic SMBH accretion is more likely to trigger AGNs. Star formation would then likely be self-regulated by the impact of supernovae, which could lower the SFR. After consumption of the majority of its fuel, the galaxy then moves to a stage of gas-poor, star-forming quiescence, with only small episodes of AGN activity occurring with any remaining gas supply and very little star formation. This situation would explain why AGNs are seen in galaxies with varying levels of gas. This paradigm is also supported by the work of Delvecchio et al. (2019), who claimed in the early stages of BH-galaxy growth, the SMBH grows slower than the

host galaxy, before undergoing rapid growth later.

The results of the studies in this thesis are consistent with the prediction that gas content is ultimately regulating the *probability* of SMBH accretion and, in a more direct way, the probability of star-formation, with the observed connections being the result of codependence rather than direct dependence. More specifically, the results of Chapter 3 (suggesting a change in the probability distribution of s$L_{\rm X}$ for starbursting galaxies) support the idea that that, whilst starburst and non-starburst galaxies had similar maximum accretion rates, there is a larger fraction of SMBHs in starburst galaxies with accretion rates above 10% of their Eddington limit (approximately 1 in 10 at $z \approx 2$) when compared to their non-starburst counterparts (1 in $60 - 70$ at $z \approx 2$). Referring back to the previously mentioned evolution scenario, AGNs can be triggered throughout, but there is a higher probability of excess accretion shortly after the initial star-forming burst when the host galaxy gas fraction remains high. This message is reinforced by the results of Chapters 4 and 5. AGNs with higher $L_{\rm X}$ tend to have more consistent and slightly higher rates of star-forming activity. Firstly, the ability of lower $L_{\rm X}$ AGNs to reside in a wider range of star-forming galaxies is consistent with the idea that the vast majority of galaxies have enough gas to increase the probability of triggering a lower luminosity AGN (even during the gas-poor quiescent scenario as mentioned previously), whereas those systems with higher levels of star formation (and hence gas) have enough gas to trigger a higher $L_{\rm X}$ AGN. This scenario is only strengthened by the consideration that quasars tend to reside in the most rapidly star-forming systems (e.g., Rosario et al., 2013; Kalfountzou et al., 2014; Stanley et al., 2017; Jarvis et al., 2020), which would be consistent with an extrapolation of our results in Chapters 4 and 5.

7.3 Concluding remarks

Revealing the full extent as to which a SMBH and its host galaxy are connected remains a complex and uncertain process. In this thesis, we specifically investigated the relationship between the rate of growth of a SMBH (witnessed as AGN) and host galaxy (witnessed through star formation). Given the complex and mixed results seen in the literature, we set out to pursue a more detailed statistical analysis and have presented studies utilising some of the most revealing statistical analyses to date. In summary, we have identified some new and reinforced some previous findings of some crucial characteristics of galaxies and SMBHs. More specifically, the results of this thesis tend to suggest that rapid SMBH accretion is more closely connected to higher rates of star formation and we propose this is consistent with the probability of both processes being regulated by cold, dense, molecular gas supply in the host galaxy.

Bibliography

Aird J., et al., 2010, MNRAS, 401, 2531

Aird J., et al., 2012, ApJ, 746, 90

Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, MNRAS, 451, 1892

Aird J., Coil A. L., Georgakakis A., 2017, MNRAS, 465, 3390

Aird J., Coil A. L., Georgakakis A., 2018, MNRAS, 474, 1225

Aird J., Coil A. L., Georgakakis A., 2019, MNRAS, 484, 4360

Alexander D. M., Hickox R. C., 2012, New Astronomy Reviews, 56, 93

Alonso-Herrero A., et al., 2006, ApJ, 640, 167

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

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

Armus L., Heckman T., Miley G., 1987, AJ, 94, 831

Audibert A., et al., 2019, A&A, 632, A33

Avni Y., Tananbaum H., 1986, ApJ, 305, 83

Azadi M., et al., 2015, ApJ, 806, 187

Barvainis R., 1987, ApJ, 320, 537

Bassani L., Dadina M., Maiolino R., Salvati M., Risaliti G., Della Ceca R., Matt G., Zamorani G., 1999, ApJS, 121, 473

Bastian N., Covey K. R., Meyer M. R., 2010, ARAA, 48, 339

Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJS, 149, 289

Bender R., Kormendy J., Dehnen W., 1996, ApJL, 464, L123

Bennert V. N., Auger M. W., Treu T., Woo J.-H., Malkan M. A., 2011, ApJ, 742, 107

Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38

Bernhard E., Mullaney J. R., Daddi E., Ciesla L., Schreiber C., 2016, MNRAS, 460, 902

Bernhard E., Mullaney J. R., Aird J., Hickox R. C., Jones M. L., Stanley F., Grimmett L. P., Daddi E., 2018, MNRAS,

Bernhard E., Grimmett L. P., Mullaney J. R., Daddi E., Tadhunter C., Jin S., 2019, MNRAS, 483, L52

Berti E., Volonteri M., 2008, ApJ, 684, 822

Bertram T., Eckart A., Fischer S., Zuther J., Straubmeier C., Wisotzki L., Krips M., 2007, A&A, 470, 571

Best P. N., Kauffmann G., Heckman T. M., Brinchmann J., Charlot S., Ivezić Ž., White S. D. M., 2005, MNRAS, 362, 25

Bisigello L., Caputi K. I., Grogin N., Koekemoer A., 2018, A&A, 609, A82

Blaes O., 2007, Accretion Disks in AGNs. p. 75

Blandford R. D., Znajek R. L., 1977, MNRAS, 179, 433

Bongiorno A., et al., 2012, MNRAS, 427, 3103

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

Boogaard L. A., et al., 2018, A&A, 619, A27

Booth C. M., Schaye J., 2009, MNRAS, 398, 53

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

Bottrill A. L., Haigh M. E., Hole M. R. A., Theakston S. C. M., Allen R. B., Grimmett L. P., Parker R. J., 2020, ApJ, 895, 141

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645

Boyle B. J., Terlevich R. J., 1998, MNRAS, 293, L49

Brandt W. N., Alexander D. M., 2015, Astronomy and Astrophysics Review, 23, 1

Brandt W. N., Laor A., Wills B. J., 2000, ApJ, 528, 637

Brightman M., Nandra K., Salvato M., Hsu L.-T., Aird J., Rangel C., 2014, MNRAS, 443, 1999

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151

Brusa M., et al., 2015, MNRAS, 446, 2394

Brusa M., et al., 2018, A&A, 612, A29

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Buat V., Xu C., 1996, A&A, 306, 61

Buat V., et al., 2005, ApJL, 619, L51

Buchner J., et al., 2014, A&A, 564, A125

Burbidge G., 1967, Quasi-stellar objects

Burlon D., Ajello M., Greiner J., Comastri A., Merloni A., Gehrels N., 2011, ApJ, 728, 58

Calzetti D., 2001, PASP, 113, 1449

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

Calzetti D., Harris J., Gallagher John S. I., Smith D. A., Conselice C. J., Homeier N., Kewley L., 2004, AJ, 127, 1405

Caplar N., Tacchella S., 2019, MNRAS, 487, 3845

Carniani S., et al., 2016, A&A, 591, A28

Cattaneo A., Dekel A., Devriendt J., Guiderdoni B., Blaizot J., 2006, MNRAS, 370, 1651

Chabrier G., 2003, PASP, 115, 763

Chang Y.-Y., van der Wel A., da Cunha E., Rix H.-W., 2015, ApJS, 219, 8

Chen C.-T. J., et al., 2013, ApJ, 773, 3

Cicone C., et al., 2014, A&A, 562, A21

Ciesla L., et al., 2015, A&A, 576, A10

Civano F., et al., 2016, VizieR Online Data Catalog, 181

Comerford J. M., Pooley D., Barrows R. S., Greene J. E., Zakamska N. L., Madejski G. M., Cooper M. C., 2015, ApJ, 806, 219

Conroy C., White M., 2013, ApJ, 762, 70

Courvoisier T. J. L., 1998, Astronomy and Astrophysics Review, 9, 1

Cresci G., Maiolino R., 2018, Nature Astronomy, 2, 179

Cresci G., et al., 2015a, A&A, 582, A63

Cresci G., et al., 2015b, ApJ, 799, 82

Croton D. J., et al., 2006, MNRAS, 365, 11

Daddi E., et al., 2007, ApJ, 670, 156

Daddi E., et al., 2017, ApJL, 846, L31

Dai Y. S., Wilkes B. J., Bergeron J., Kuraszkiewicz J., Omont A., Atanas A., Teplitz H. I., 2018, MNRAS, 478, 4238

Dale D. A., Helou G., Magdis G. E., Armus L., Díaz-Santos T., Shi Y., 2014, ApJ, 784, 83

Davies J. J., Crain R. A., McCarthy I. G., Oppenheimer B. D., Schaye J., Schaller M., McAlpine S., 2019, MNRAS, 485, 3783

Delvecchio I., et al., 2014, MNRAS, 439, 2736

Delvecchio I., et al., 2015, MNRAS, 449, 373

Delvecchio I., et al., 2019, ApJL, 885, L36

Delvecchio I., et al., 2020, arXiv e-prints, p. arXiv:2002.08965

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604

Dietrich J., et al., 2018, MNRAS, 480, 3562

Ding X., et al., 2020a, arXiv e-prints, p. arXiv:2005.13550

Ding X., et al., 2020b, ApJ, 888, 37

Donas J., Deharveng J. M., 1984, A&A, 140, 325

Done C., Diaz Trigo M., 2010, MNRAS, 407, 2287

Donley J. L., et al., 2012, ApJ, 748, 142

Dovčiak M., Done C., 2016, Astronomische Nachrichten, 337, 441

Dressler A., Richstone D. O., 1988, ApJ, 324, 701

Dubois Y., Gavazzi R., Peirani S., Silk J., 2013, MNRAS, 433, 3297

Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, MNRAS, 463, 3948

Eckart A., Genzel R., 1996, Nature, 383, 415

Eckart A., Genzel R., 1997, MNRAS, 284, 576

Eisenhardt P. R. M., et al., 2012, ApJ, 755, 173

Elbaz D., et al., 2007, A&A, 468, 33

Elitzur M., Shlosman I., 2006, ApJL, 648, L101

Ellison S. L., Brown T., Catinella B., Cortese L., 2019, MNRAS, 482, 5694

Elvis M., et al., 1994, ApJS, 95, 1

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

Fabello S., Kauffmann G., Catinella B., Giovanelli R., Haynes M. P., Heckman T. M., Schiminovich D., 2011, MNRAS, 416, 1739

Fabian A. C., 2012, ARAA, 50, 455

Fabian A. C., Iwasawa K., 1999, MNRAS, 303, L34

Fabian A. C., Sanders J. S., Taylor G. B., Allen S. W., Crawford C. S., Johnstone R. M., Iwasawa K., 2006, MNRAS, 366, 417

Fabian A. C., Lohfink A., Kara E., Parker M. L., Vasudevan R., Reynolds C. S., 2015, MNRAS, 451, 4375

Ferrarese L., Merritt D., 2000, ApJL, 539, L9

Fiore F., et al., 2008, ApJ, 672, 94

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Fritz J., Franceschini A., Hatziminaoglou E., 2006, MNRAS, 366, 767

Galeev A. A., Rosner R., Vaiana G. S., 1979, ApJ, 229, 318

Gallagher J. S., Bushouse H., Hunter D. A., 1989, AJ, 97, 700

Gallagher R., Maiolino R., Belfiore F., Drory N., Riffel R., Riffel R. A., 2019, MNRAS, 485, 3409

Gandhi P., Horst H., Smette A., Hönig S., Comastri A., Gilli R., Vignali C., Duschl W., 2009, A&A, 502, 457

Gao Y., Solomon P. M., 2004, ApJ, 606, 271

Gebhardt K., et al., 2000, ApJL, 539, L13

Georgakakis A., et al., 2014, MNRAS, 440, 339

Georgakakis A., Aird J., Schulze A., Dwelly T., Salvato M., Nandra K., Merloni A., Schneider D. P., 2017, MNRAS, 471, 1976

Georgantopoulos I., Georgakakis A., Rowan-Robinson M., Rovilos E., 2008, A&A, 484, 671

Ghez A. M., Klein B. L., Morris M., Becklin E. E., 1998, ApJ, 509, 678

Gibson R. R., Brandt W. N., Schneider D. P., 2008, ApJ, 685, 773

Gilfanov M., Merloni A., 2014, Space Science Reviews, 183, 121

Girichidis P., et al., 2020, Space Science Reviews, 216, 68

Glikman E., Simmons B., Mailly M., Schawinski K., Urry C. M., Lacy M., 2015, ApJ, 806, 218

Goldader J. D., Meurer G., Heckman T. M., Seibert M., Sanders D. B., Calzetti D., Steidel C. C., 2002, ApJ, 568, 651

Gottardo R., Raftery A. E., 2008, Journal of Computational and Graphical Statistics, 17, 949

Greene J. E., Zakamska N. L., Smith P. S., 2012, ApJ, 746, 86

Grimmett L. P., Mullaney J. R., Jin S., Bernhard E., Daddi E., Walters K., 2019, MNRAS, 487, 4071

Grimmett L. P., Mullaney J. R., Bernhard E. P., Harrison C. M., Alexander D. M., Stanley F., Masoura V. A., Walters K., 2020, Monthly Notices of the Royal Astronomical Society

Gültekin K., et al., 2009, ApJ, 698, 198

Guo Y., et al., 2013, ApJS, 207, 24

Haan S., Schinnerer E., Mundell C. G., García-Burillo S., Combes F., 2008, AJ, 135, 232

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

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

Haardt F., Maraschi L., Ghisellini G., 1994, ApJL, 432, L95

Hao H., et al., 2010, ApJL, 724, L59

Hao H., Elvis M., Civano F., Lawrence A., 2011, ApJ, 733, 108

Häring N., Rix H.-W., 2004, ApJL, 604, L89

Harris K., et al., 2016, MNRAS, 457, 4179

Harrison C. M., 2017, Nature Astronomy, 1, 0165

Harrison C. M., et al., 2012, ApJL, 760, L15

Hastings W. K., 1970, Biometrika, 57, 97

Hatziminaoglou E., et al., 2010, A&A, 518, L33

Hawkins M. R. S., 2007, A&A, 462, 581

Hayward C. C., Smith D. J. B., 2015, MNRAS, 446, 1512

Heckman T. M., Best P. N., 2014, ARAA, 52, 589

Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., Tremonti C., White S. D. M., 2004, ApJ, 613, 109

Heckman T. M., Ptak A., Hornschemeier A., Kauffmann G., 2005, ApJ, 634, 161

Hickox R. C., et al., 2007, ApJ, 671, 1365

Hickox R. C., Mullaney J. R., Alexander D. M., Chen C.-T. J., Civano F. M., Goulding A. D., Hainline K. N., 2014, ApJ, 782, 9

Hickox R. C., Myers A. D., Greene J. E., Hainline K. N., Zakamska N. L., DiPompeo M. A., 2017, ApJ, 849, 53

Hinshaw G., et al., 2013, ApJS, 208, 19

Hopkins A. M., 2018, PASA, 35, 39

Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, ApJS, 163, 1

Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, ApJS, 175, 356

Hopkins P. F., Hickox R., Quataert E., Hernquist L., 2009, MNRAS, 398, 333

Horst H., Gandhi P., Smette A., Duschl W. J., 2008, A&A, 479, 389

Hu J., 2008, MNRAS, 386, 2242

Husemann B., Davis T. A., Jahnke K., Dannerbauer H., Urrutia T., Hodge J., 2017, MNRAS, 470, 1570

Ilbert O., et al., 2015, A&A, 579, A2

Jackson R. A., Martin G., Kaviraj S., Laigle C., Devriendt J. E. G., Dubois Y., Pichon C., 2019, MNRAS, 489, 4679

Jarvis M. E., et al., 2020, arXiv e-prints, p. arXiv:2007.10351

Jin S., et al., 2018, ApJ, 864, 56

Jones M. L., Hickox R. C., Black C. S., Hainline K. N., DiPompeo M. A., Goulding A. D., 2016, ApJ, 826, 12

Kakkad D., et al., 2017, MNRAS, 468, 4205

Kalfountzou E., et al., 2014, MNRAS, 442, 1181

Karim A., et al., 2011, ApJ, 730, 61

Kashino D., et al., 2013, ApJL, 777, L8

Kass R. E., Raftery A. E., 1995, Journal of the American Statistical Association, 90, 773

Kauffmann G., 2018, MNRAS, 480, 3201

Kauffmann G., Heckman T. M., 2009, MNRAS, 397, 135

Kauffmann G., et al., 2003, MNRAS, 346, 1055

Kaviraj S., Martin G., Silk J., 2019, MNRAS, 489, L12

Kennicutt Jr. R. C., 1998, ARAA, 36, 189

Kennicutt R. C., Evans N. J., 2012, ARAA, 50, 531

Kennicutt Robert C. J., et al., 2009, ApJ, 703, 1672

King A. R., Pringle J. E., 2006, MNRAS, 373, L90

König S., Eckart A., García-Marín M., Huchtmeier W. K., 2009, A&A, 507, 757

Kormendy J., Ho L. C., 2013, ARAA, 51, 511

Kormendy J., Richstone D., 1995, ARAA, 33, 581

Kormendy J., et al., 1997, ApJL, 482, L139

Kormendy J., Fisher D. B., Cornell M. E., Bender R., 2009, ApJS, 182, 216

Krips M., Neri R., Cox P., 2012, ApJ, 753, 135

Krolik J. H., Begelman M. C., 1988, ApJ, 329, 702

Kroupa P., 2001, MNRAS, 322, 231

Kroupa P., Weidner C., Pflamm-Altenburg J., Thies I., Dabringhausen J., Marks M., Maschberger T., 2013, The Stellar and Sub-Stellar Initial Mass Function of Simple and Composite Populations. p. 115, doi:10.1007/978-94-007-5612-0˙4

Krumholz M. R., 2014, Physics Reports, 539, 49

Krumholz M. R., Thompson T. A., 2007, ApJ, 669, 289

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

LaMassa S. M., Heckman T. M., Ptak A., Martins L., Wild V., Sonnentrucker P., 2010, ApJ, 720, 786

Lacy M., et al., 2004, ApJS, 154, 166

Lacy M., Ridgway S. E., Sajina A., Petric A. O., Gates E. L., Urrutia T., Storrie-Lombardi L. J., 2015, ApJ, 802, 102

Laigle C., et al., 2016, ApJS, 224, 24

Laigle C., et al., 2019, MNRAS, 486, 5104

Lamastra A., Bianchi S., Matt G., Perola G. C., Barcons X., Carrera F. J., 2009, A&A, 504, 73

Lansbury G. B., et al., 2017, ApJ, 846, 20

Lanzuisi G., et al., 2017, A&A, 602, A123

Larson R. B., 2003, Reports on Progress in Physics, 66, 1651

Larson R. B., 2010, Reports on Progress in Physics, 73, 014901

Larson D., et al., 2011, ApJS, 192, 16

Laurent O., Mirabel I. F., Charmandaris V., Gallais P., Madden S. C., Sauvage M., Vigroux L., Cesarsky C., 2000, A&A, 359, 887

Lequeux J., Maucherat-Joubert M., Deharveng J. M., Kunth D., 1981, A&A, 103, 305

Liang E. P. T., Price R. H., 1977, ApJ, 218, 247

Lira P., Videla L., Wu Y., Alonso-Herrero A., Alexander D. M., Ward M., 2013, ApJ, 764, 159

Liu D., et al., 2018, ApJ, 853, 172

Lo K. Y., Shen Z.-Q., Zhao J.-H., Ho P. T. P., 1998, ApJL, 508, L61

Lusso E., et al., 2010, A&A, 512, A34

Lutz D., et al., 2011, A&A, 532, A90

Lynden-Bell D., 1969, Nature, 223, 690

Lynden-Bell D., 1978, Physica Scripta, 17, 185

Lynden-Bell D., Rees M. J., 1971, MNRAS, 152, 461

Madau P., Dickinson M., 2014, ARAA, 52, 415

Madau P., Pozzetti L., Dickinson M., 1998, ApJ, 498, 106

Magorrian J., et al., 1998, AJ, 115, 2285

Maiolino R., Ruiz M., Rieke G. H., Papadopoulos P., 1997, ApJ, 485, 552

Maiolino R., et al., 2017, Nature, 544, 202

Malkan M. A., Sargent W. L. W., 1982, ApJ, 254, 22

Marchesi S., et al., 2016, ApJ, 817, 34

Marconi A., Hunt L. K., 2003, ApJL, 589, L21

Martínez-Sansigre A., Rawlings S., Lacy M., Fadda D., Jarvis M. J., Marleau F. R., Simpson C., Willott C. J., 2006, MNRAS, 370, 1479

Masoura V. A., Mountrichas G., Georgantopoulos I., Ruiz A., Magdis G., Plionis M., 2018, A&A, 618, A31

Mateos S., Alonso-Herrero A., Carrera F. J., Blain A., Severgnini P., Caccianiga A., Ruiz A., 2013, MNRAS, 434, 941

McAlpine S., Bower R. G., Harrison C. M., Crain R. A., Schaller M., Schaye J., Theuns T., 2017, MNRAS, 468, 3395

McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184

McKee C. F., Ostriker E. C., 2007, ARAA, 45, 565

Merloni A., Rudnick G., Di Matteo T., 2004, MNRAS, 354, L37

Merritt D., Ferrarese L., 2001, ApJ, 547, 140

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, Journal of Chemical Physics, 21, 1087

Miller J. S., Goodrich B. F., 1986, in Bulletin of the American Astronomical Society. p. 1001

Mobasher B., et al., 2015, ApJ, 808, 101

Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903

Mullaney J. R., Alexander D. M., Goulding A. D., Hickox R. C., 2011, MNRAS, 414, 1082

Mullaney J. R., et al., 2012a, MNRAS, 419, 95

Mullaney J. R., et al., 2012b, ApJL, 753, L30

Mullaney J. R., et al., 2015, MNRAS, 453, L83

Mushotzky R. F., Done C., Pounds K. A., 1993, ARAA, 31, 717

Netzer H., 2015, ARAA, 53, 365

Netzer H., et al., 2007, ApJ, 666, 806

Noeske K. G., et al., 2007, ApJL, 660, L47

Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, A&A, 507, 1793

Offner S. S. R., Clark P. C., Hennebelle P., Bastian N., Bate M. R., Hopkins P. F., Moraux E., Whitworth A. P., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. p. 53 (arXiv:1312.5326), doi:10.2458/azu˙uapress˙9780816531240-ch003

Oteo I., et al., 2013, A&A, 554, L3

Padovani P., et al., 2017, Astronomy and Astrophysics Review, 25, 2

Panessa F., Bassani L., Cappi M., Dadina M., Barcons X., Carrera F. J., Ho L. C., Iwasawa K., 2006, A&A, 455, 173

Panter B., Jimenez R., Heavens A. F., Charlot S., 2007, MNRAS, 378, 1550

Pawlik M. M., et al., 2018, MNRAS, 477, 1708

Penner K., et al., 2012, ApJ, 759, 28

Perna M., et al., 2018, A&A, 619, A90

Petrucci P. O., Ursini F., De Rosa A., Bianchi S., Cappi M., Matt G., Dadina M., Malzac J., 2018, A&A, 611, A59

Pillepich A., et al., 2018, MNRAS, 473, 4077

Pope A., et al., 2019, Bulletin of the American Astronomical Society, 51, 330

Popesso P., et al., 2019a, MNRAS, 483, 3213

Popesso P., et al., 2019b, MNRAS, 490, 5285

Pozo Nuñez F., et al., 2015, A&A, 576, A73

Price S. H., Kriek M., Feldmann R., Quataert E., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., Barro G., 2017, ApJL, 844, L6

Pringle J. E., Rees M. J., Pacholczyk A. G., 1973, A&A, 29, 179

Rafferty D. A., Brandt W. N., Alexander D. M., Xue Y. Q., Bauer F. E., Lehmer B. D., Luo B., Papovich C., 2011, ApJ, 742, 3

Raimundo S. I., Fabian A. C., 2009, MNRAS, 396, 1217

Raimundo S. I., Fabian A. C., Vasudevan R. V., Gand hi P., Wu J., 2012, MNRAS, 419, 2529

Ramasawmy J., Stevens J., Martin G., Geach J. E., 2019, MNRAS, 486, 4320

Ramos Almeida C., et al., 2011, ApJ, 731, 92

Reddy N. A., Erb D. K., Pettini M., Steidel C. C., Shapley A. E., 2010, ApJ, 712, 1070

Rees M. J., 1984, ARAA, 22, 471

Rees M. J., 1998, in Wald R. M., ed., Black Holes and Relativistic Stars. p. 79 (arXiv:astro-ph/9701161)

Reynolds C. S., Nowak M. A., 2003, Physics Reports, 377, 389

Ricci C., et al., 2017, MNRAS, 468, 1273

Richards G. T., et al., 2006, ApJS, 166, 470

Riguccini L., et al., 2011, A&A, 534, A81

Rodighiero G., et al., 2011, ApJL, 739, L40

Rodighiero G., et al., 2015, ApJL, 800, L10

Roehlly Y., Burgarella D., Buat V., Boquien M., Ciesla L., Heinis S., 2014, in Manset N., Forshay P., eds, Astronomical Society of the Pacific Conference Series Vol. 485, Astronomical Data Analysis Software and Systems XXIII. p. 347 (arXiv:1309.6366)

Rosario D. J., et al., 2012, A&A, 545, A45

Rosario D. J., et al., 2013, ApJ, 771, 63

Rosario D. J., et al., 2018, MNRAS, 473, 5658

Rowan-Robinson M., et al., 2005, AJ, 129, 1183

Ruffini R., Wilson J. R., 1975, Physical Review D, 12, 2959

Salim S., et al., 2007, ApJS, 173, 267

Salpeter E. E., 1955, ApJ, 121, 161

Salpeter E. E., 1964, ApJ, 140, 796

Sandage A., 1965, ApJ, 141, 1560

Sanders D. B., Mirabel I. F., 1996, ARAA, 34, 749

Sanders D. B., Soifer B. T., Elias J. H., Madore B. F., Matthews K., Neugebauer G., Scoville N. Z., 1988, ApJ, 325, 74

Santini P., et al., 2009, A&A, 504, 751

Santini P., et al., 2017, ApJ, 847, 76

Sargent W. L. W., Young P. J., Boksenberg A., Shortridge K., Lynds C. R., Hartwick F. D. A., 1978, ApJ, 221, 731

Sargent M. T., Béthermin M., Daddi E., Elbaz D., 2012, ApJ, 747, L31

Scalo J. M., 1986, Fundamentals of Cosmic Physics, 11, 1

Schaye J., et al., 2015, MNRAS, 446, 521

Schmidt M., 1959, ApJ, 129, 243

Schmidt M., 1963, Nature, 197, 1040

Schmidt M., 1969, ARAA, 7, 527

Schoettler C., Parker R. J., Arnold B., Grimmett L. P., de Bruijne J., Wright N. J., 2019, MNRAS, 487, 4615

Scholtz J., et al., 2018, MNRAS, 475, 1288

Schreiber C., et al., 2015, A&A, 575, A74

Schulze A., et al., 2019, MNRAS, 488, 1180

Schweitzer M., et al., 2006, ApJ, 649, 79

Scoville N., et al., 2007, ApJS, 172, 1

Serra P., Amblard A., Temi P., Burgarella D., Giovannoli E., Buat V., Noll S., Im S., 2011, ApJ, 740, 22

Seyfert C. K., 1943, ApJ, 97, 28

Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337

Shangguan J., Ho L. C., Xie Y., 2018, ApJ, 854, 158

Shangguan J., Ho L. C., Li R., Zhuang M.-Y., Xie Y., Li Z., 2019, ApJ, 870, 104

Shangguan J., Ho L. C., Bauer F. E., Wang R., Treister E., 2020a, arXiv e-prints, p. arXiv:2007.11286

Shangguan J., Ho L. C., Bauer F. E., Wang R., Treister E., 2020b, ApJS, 247, 15

Shao L., et al., 2010, A&A, 518, L26

Shields G. A., 1978, Nature, 272, 706

Shimizu T. T., Mushotzky R. F., Meléndez M., Koss M. J., Barger A. J., Cowie L. L., 2017, MNRAS, 466, 3161

Shimizu T. T., et al., 2019, MNRAS, 490, 5860

Shlosman I., Frank J., Begelman M. C., 1989, Nature, 338, 45

Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877

Silk J., Mamon G. A., 2012, Research in Astronomy and Astrophysics, 12, 917

Silk J., Rees M. J., 1998, A&A, 331, L1

Silverman J. D., et al., 2008, ApJ, 675, 1025

Smolčić V., et al., 2017, A&A, 602, A1

Sobolewska M. A., Siemiginowska A., Życki P. T., 2004, ApJ, 617, 102

Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481

Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJS, 214, 15

Springel V., et al., 2005, Nature, 435, 629

Stalevski M., Ricci C., Ueda Y., Lira P., Fritz J., Baes M., 2016, MNRAS, 458, 2288

Stanley F., Harrison C. M., Alexander D. M., Swinbank A. M., Aird J. A., Del Moro A., Hickox R. C., Mullaney J. R., 2015, MNRAS, 453, 591

Stanley F., et al., 2017, MNRAS, 472, 2221

Steinhardt C. L., et al., 2014, ApJL, 791, L25

Stemo A., Comerford J. M., Barrows R. S., Stern D., Assef R. J., Griffith R. L., 2020, ApJ, 888, 78

Stern B. E., Poutanen J., Svensson R., Sikora M., Begelman M. C., 1995, ApJL, 449, L13

Stern D., et al., 2005, ApJ, 631, 163

Storchi-Bergmann T., Dors Oli L. J., Riffel R. A., Fathi K., Axon D. J., Robinson A., Marconi A., Östlin G., 2007, ApJ, 670, 959

Suganuma M., et al., 2006, ApJ, 639, 46

Suh H., et al., 2017, ApJ, 841, 102

Suh H., et al., 2019, ApJ, 872, 168

Tadhunter C., et al., 2018, MNRAS, 478, 1558

Tomczak A. R., et al., 2016, ApJ, 817, 118

Tonry J. L., 1984, ApJL, 283, L27

Tonry J. L., 1987, ApJ, 322, 632

Torrey P., et al., 2015, MNRAS, 447, 2753

Tremaine S., et al., 2002, ApJ, 574, 740

Turner T. J., Miller L., 2009, Astronomy and Astrophysics Review, 17, 47

Ulrich M. H., et al., 1980, MNRAS, 192, 561

Ulrich M.-H., Maraschi L., Urry C. M., 1997, ARAA, 35, 445

Vaiana G. S., Rosner R., 1978, ARAA, 16, 393

Villar-Martín M., et al., 2013, MNRAS, 434, 978

Vito F., et al., 2018, MNRAS, 474, 4528

Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031

Volonteri M., Madau P., Quataert E., Rees M. J., 2005, ApJ, 620, 69

Walcher J., Groves B., Budavári T., Dale D., 2011, Ap&SS, 331, 1

Wang T., et al., 2017, A&A, 601, A63

Whitaker K. E., Kriek M., van Dokkum P. G., Bezanson R., Brammer G., Franx M., Labbé I., 2012, ApJ, 745, 179

Whitaker K. E., et al., 2015, ApJL, 811, L12

Wong T., 2009, ApJ, 705, 650

Wu J., Evans Neal J. I., Gao Y., Solomon P. M., Shirley Y. L., Vanden Bout P. A., 2005, ApJL, 635, L173

Wuyts S., Franx M., Cox T. J., Hernquist L., Hopkins P. F., Robertson B. E., van Dokkum P. G., 2009, ApJ, 696, 348

Wuyts S., et al., 2011, ApJ, 742, 96

Wyithe J. S. B., 2006, MNRAS, 365, 1082

Xia X. Y., et al., 2012, ApJ, 750, 92

Yesuf H. M., Ho L. C., 2020, arXiv e-prints, p. arXiv:2007.12026

Yesuf H. M., French K. D., Faber S. M., Koo D. C., 2017, MNRAS, 469, 3015

Zel'dovich Y. B., 1964, Soviet Physics Doklady, 9, 195

de Nicola S., Marconi A., Longo G., 2019, MNRAS, 490, 600