



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## The Incidence of X-ray selected AGN in Nearby Galaxies

**Citation for published version:**

Birchall, KL, Watson, MG, Aird, J & Starling, RLC 2022, 'The Incidence of X-ray selected AGN in Nearby Galaxies', *Monthly Notices of the Royal Astronomical Society*, vol. 510, no. 3, pp. 4556-4572.  
<https://doi.org/10.1093/mnras/stab3573>

**Digital Object Identifier (DOI):**

[10.1093/mnras/stab3573](https://doi.org/10.1093/mnras/stab3573)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Monthly Notices of the Royal Astronomical Society

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# The Incidence of X-ray selected AGN in Nearby Galaxies

Keir L. Birchall<sup>1</sup>, <sup>\*</sup> M. G. Watson<sup>1</sup>, J. Aird<sup>2,1</sup>, and R. L. C. Starling<sup>1</sup>

<sup>1</sup>*School of Physics & Astronomy, University of Leicester, University Road, Leicester LE1 7RH, UK*

<sup>2</sup>*Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK*

Accepted 2021 December 3. Received 2021 October 20; in original form 2021 January 29

## ABSTRACT

We present the identification and analysis of an unbiased sample of AGN that lie within the local galaxy population. Using the MPA-JHU catalogue (based on SDSS DR8) and 3XMM DR7 we define a parent sample of 25,949 local galaxies ( $z \leq 0.33$ ). After confirming that there was strictly no AGN light contaminating stellar mass and star-formation rate calculations, we identified 917 galaxies with central, excess X-ray emission likely originating from an AGN. We analysed their optical emission lines using the BPT diagnostic and confirmed that such techniques are more effective at reliably identifying sources as AGN in higher mass galaxies: rising from 30% agreement in the lowest mass bin to 93% in the highest. We then calculated the growth rates of the black holes powering these AGN in terms of their specific accretion rates ( $\propto L_X/M_*$ ). Our sample exhibits a wide range of accretion rates, with the majority accreting at rates  $\leq 0.5\%$  of their Eddington luminosity. Finally, we used our sample to calculate the incidence of AGN as a function of stellar mass and redshift. After correcting for the varying sensitivity of 3XMM, we split the galaxy sample by stellar mass and redshift and investigated the AGN fraction as a function of X-ray luminosity and specific black hole accretion rate. From this we found the fraction of galaxies hosting AGN above a fixed specific accretion rate limit of  $10^{-3.5}$  is constant (at  $\approx 1\%$ ) over stellar masses of  $8 < \log M_*/M_\odot < 12$  and increases (from  $\approx 1\%$  to  $10\%$ ) with redshift.

**Key words:** galaxies:active – galaxies:evolution – black hole physics – X-rays:galaxies

## 1 INTRODUCTION

The degree to which a supermassive black hole (SMBH) and its host galaxy interact and affect each other’s evolution is an important and contested question. Among the well-studied and constrained examples of co-evolution are the relations between an SMBH’s mass and the velocity dispersion, mass and luminosity of the host galaxy’s classical bulge. (Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000). Such relationships imply that the SMBH and bulge co-evolve by regulating each others growth (Kormendy & Ho 2013). SMBH growth can also be identified through a range of electromagnetic signatures released when they undergo periods of accretion. Tracing the features of active galactic nuclei (AGN) activity across cosmic time has highlighted more evidence of co-evolution. For example, total AGN accretion rate density is highly correlated with changes in star-formation rate (SFR) density (Madau & Dickinson 2014; Delvecchio et al. 2014; Aird et al. 2015); the effect of AGN on their host galaxies in the form of feedback is associated with the quenching of star formation (Di Matteo et al. 2005; Fabian 2012; Greene et al. 2020); and feedback could also explain the observed bi-modality in colour-magnitude and colour-mass space (Baldry et al. 2004;

Martin et al. 2007; Schawinski et al. 2014).

It is still unclear what mechanism is the main driver behind getting gas and dust down into the centre (Alexander & Hickox 2012). As discussed in a number of studies, taking large samples of galaxies is a useful way of constraining the AGN fuelling mechanism as it allows us to smooth out stochastic differences and try to identify trends (Aird et al. 2013; Hickox et al. 2014; Delvecchio et al. 2020). The most complete AGN samples are typically found by using X-ray surveys as they can identify AGN across large redshift ranges and down to relatively low luminosities where non-AGN emission may dominate (Brandt & Alexander 2015). Studies that have adopted this approach suggest that AGN are much more likely to be found in higher mass galaxies, and in galaxies at higher redshifts (Xue et al. 2010; Haggard et al. 2010; Mendez et al. 2013; Hernán-Caballero et al. 2014; Williams & Röttgering 2015).

Whilst these surveys employ large samples of AGN, they are subject to several observational biases. Firstly, most X-ray detection techniques define AGN as objects that exceed absolute luminosity thresholds, or more generally dominate galaxy emission (Brandt & Alexander 2015), causing them to miss the lower luminosity, less efficiently accreting black holes typically found in dwarf galaxies. Any low mass black holes that are detected are the much rarer,

\* E-mail: klb69@leicester.ac.uk

more actively accreting ones. However, over the past decade there has been a huge increase in the number of AGN detected in dwarf galaxies aided, in part, by adapting these techniques. X-ray selected AGN in dwarf galaxies are identified through isolating a centrally located excess emission that cannot be explained by other sources, including X-ray binaries and hot gas (Lemons et al. 2015; Paggi et al. 2016; Pardo et al. 2016; Baldassare et al. 2017; Mezcuca et al. 2018; Birchall et al. 2020). Such a change in detection method allows us to overcome the bias towards actively accreting, higher mass black holes associated with the absolute X-ray luminosity thresholds and helps challenge the assumption that AGN are less likely in lower mass galaxies.

Secondly, a lot of these studies make little attempt to understand and correct for the limitations of the survey data used. Aird et al. (2012) was among the first AGN studies to attempt to overcome these observational limitations and apply completeness corrections to their sample. By applying these corrections to a sample of galaxies out to  $z \approx 1$ , from the Prism Multi-object Survey (Coil et al. 2011; Cool et al. 2013), they were able to determine the probability of finding an AGN as a function of various host galaxy properties. Aird et al. (2012) showed that the probability of hosting an AGN can be described as a power law of X-ray luminosity and specific black hole accretion rate (sBHAR). Both X-ray luminosity and sBHAR distributions were found to be consistent over a wide range of stellar masses but with normalisations that dropped rapidly with decreasing redshift. Similar approaches to Aird et al. (2012) have been adopted and adapted by subsequent AGN population studies (Bongiorno et al. 2012; Georgakakis et al. 2017; Aird et al. 2017, 2018; Birchall et al. 2020).

By combining the large galaxy sample available in the SDSS and the X-ray data from XMM-Newton Serendipitous Sky Survey, Birchall et al. (2020) found power law distributions describing the AGN incidence in dwarf galaxies as a function of X-ray luminosity and sBHAR. From this we identified a redshift-independent fraction out to  $z \sim 0.7$ . We also found a weak increase in AGN fraction with stellar mass, however it was also consistent with being flat for the range probed. In this paper, we take the methods used in Birchall et al. (2020), remove the mass threshold and investigate the X-ray luminosity and sBHAR distributions for this expanded sample of local AGN. Georgakakis et al. (2011) have previously calculated the space density of X-ray AGN with  $\log_{10} L_X > 41$  at  $0.03 < z < 0.2$  based on a sample constructed using both XMM and SDSS observations (Georgakakis & Nandra 2011). However, there have been thousands of new XMM observations since this work was produced. Thus, with this expanded survey and our new methods to account for sample incompleteness we aim to identify whether there is any connection between AGN incidence and stellar mass or redshift.

The paper is structured as follows. First, we describe the construction of our AGN sample (§2), making sure to verify that large non-stellar nuclear emission does not affect our host galaxy measurements. Then, we assess the extent to which our X-ray selected sample overlaps with BPT selection (§3). Next, we investigate the rate at which the black holes powering our AGN are accreting material (§4). With this information we then create our X-ray luminosity and sBHAR distributions (§5) and use them to calculate how the AGN fraction varies with stellar mass and redshift (§6). Throughout, we assume Friedmann-Robertson-Walker cosmology:

$$\Omega = 0.3, \Lambda = 0.7 \text{ and } H_0 = 70 \text{ km s}^{-1} \text{Mpc}^{-1}.$$

## 2 DATA & SAMPLE SELECTION

As this paper aims to build upon the work started in Birchall et al. (2020), we continue to use the catalogues described therein. In this section, we will provide a brief summary of these catalogues - please refer to the original paper for a more detailed description. In addition, we will outline the new, more rigorous position-matching process needed to efficiently match our expanded sample.

Optical photometry and spectroscopy covering about a quarter of the sky ( $\approx 9274 \text{ deg}^2$ ) can be found in the Sloan Digital Sky Survey Data Release 8 (SDSS DR8). Any SDSS DR8 spectra classified by the pipeline as a galaxy was then considered for further processing by the MPA-JHU value-added catalogue<sup>1</sup>. It provides estimates of galaxy properties such as stellar mass, star formation rate (SFR) and emission line fluxes, for 1,472,583 objects in this release. From this catalogue we isolate the 835,861 objects with a "good" photometric reliability flag. This indicates valid results from MPA-JHU's photometric fits which indicates useable mass and star-formation rate (SFR) measurements.

Stellar masses are calculated using the *ugriz* photometry from the full extent of the galaxy, and UV data from GALEX (Martin et al. 2005). The total stellar mass is calculated by fitting to model magnitudes based on the method described in Kauffmann et al. (2003b). MPA-JHU stellar masses were found to be largely consistent with those from the Galex-SDSS-WISE Legacy Catalogue (Salim et al. 2016). The typical difference ranged from 0.03 – 0.13 dex. SFR is similarly measured using fits to the galaxy photometry taken from the SDSS and GALEX. SFRs are calculated using the method outlined in Salim et al. (2007) with dust attenuation calculated using the Charlot & Fall (2000) model. Salim et al. (2016) found that the MPA-JHU SFRs showed no evidence of bias when compared with two independent SFR measures, including mid-IR SFRs.

Whilst the MPA-JHU catalogue is formally deprecated by the SDSS we found the three alternative catalogues (Wisconsin, Portsmouth and Granada<sup>2</sup>) were insufficient for our purposes. The Wisconsin catalogue did not calculate any SFR values, nor did they present any spectral quantity from which one could be derived. The Portsmouth catalogue does calculate the appropriate quantities but 64% of their SFRs have a zero value across a wide range of magnitudes. In addition, the non-zero SFRs are gridded, implying a lack precision at lower values. The Granada catalogue has the most promising quantities, having a greater proportion of non-zero SFRs. However, there is also evidence of gridding at lower values. Furthermore, there appears to be a strict limit applied to the specific SFR. It is unclear why this limit has been applied and it has the effect of removing the galactic main sequence of star formation. MPA-JHU is not subject to any of the above issues. It is a widely used catalogue with robust stellar mass and SFR values. Thus, despite its deprecated status, MPA-JHU is the best catalogue

<sup>1</sup> Available at <http://www.mpa-garching.mpg.de/SDSS/DR7/>

<sup>2</sup> For more information about these catalogues, visit the [SDSS galaxy properties page](#)

for this work and its quantities are used throughout.

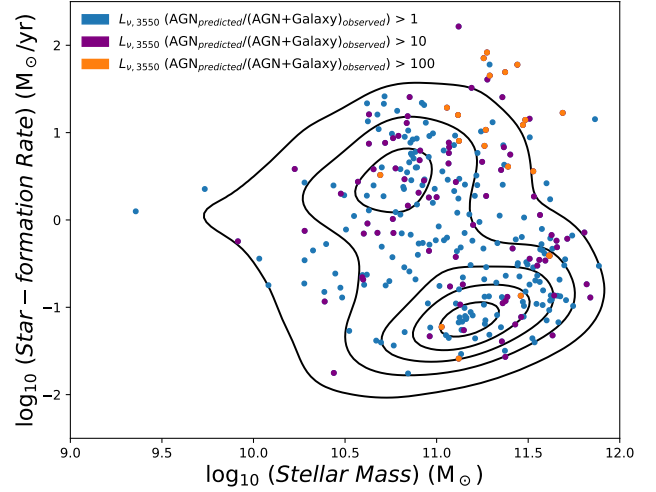
The X-ray data come from the 3XMM DR7 catalogue (Rosen et al. 2016). It is based on 9,710 pointed observations with the XMM-Newton EPIC cameras which have a field of view  $\approx 30'$  and cover the energy range  $\sim 0.2 - 12$  keV. DR7 contains  $\sim 400,000$  unique X-ray sources based on 727,790 individual detections. For our study we use the unique source list rather than the individual detections. Our results are thus averaged over several individual observations for a significant number of sources. We summed fluxes in the 2 - 4.5 keV and 4.5 - 12 keV bands and converted them to luminosities in the 2 - 12 keV energy range using the MPA-JHU redshifts, assuming  $\Gamma = 1.7$ . We continue to use this catalogue, instead of 4XMM, because our analysis requires the use of comprehensive upper limits data accessible through Flix<sup>3</sup> (Carrera et al. 2007). 3XMM DR7 is the most recent version of the serendipitous sky survey available with this infrastructure.

## 2.1 Position Matching

In Birchall et al. (2020) the dwarf galaxies in the MPA-JHU catalogue were matched to 3XMM with a simple but robust statistical technique. The small sample size allowed for individual inspection of degenerate matches i.e. galaxies matched to multiple X-ray signals, and vice versa. By removing the mass threshold, we are massively expanding our potential sample. This opened up the possibility of thousands of potential degenerate matches between MPA-JHU and 3XMM. An automated method that assessed the strength of these degenerate matches would be needed to feasibly process such a large catalogue.

To overcome this challenge we turned to the ARCHES cross-correlation tool, `xmatch` (Pineau et al. 2017). It is an astronomical matching tool able to identify the counterparts of one catalogue to multiple others, whilst also computing probabilities of associations using background sources and positional errors. In consultation with its creator, we wrote a script to match MPA-JHU and 3XMM DR7. It uses a standard approach adopted when working with XMM, where the data and matching process is broken down into individual XMM fields. Every 3XMM and SDSS object within  $15'$  of the field's central co-ordinates was found. Once complete, a level of quality control was applied, to isolate the most secure X-ray detections - those entries whose absolute position error was no greater than  $4''$  and whose physical extent was less than  $10''$ . Crucially, `xmatch` calculated the Bayesian probability of association and non-association for each remaining pair of X-ray and galaxy co-ordinates within the current 3XMM field. After repeating this process for all 3XMM fields, `xmatch` identified 110,274 potential matches.

Unfortunately, by focusing on individual fields the matching process did account not for any overlap that might occur between them. Before we assessed the match strengths, we removed repeating galaxy or X-ray matches produced from overlapping fields. Once done, we isolated the SDSS objects with an entry in MPA-JHU. All the SDSS co-ordinates were matched to MPA-JHU co-ordinates within  $1''$ , leaving 3,357 galaxies with mass and SFR estimates. Next, we identified the most likely X-ray & galaxy pairs. We chose to set



**Figure 1.** Comparing the stellar mass and SFR of galaxies with a high, predicted optical AGN contribution ( $L_{\nu,3550\text{\AA}}(\text{AGN}_{\text{predicted}}/(\text{AGN} + \text{Galaxy})_{\text{observed}})$ ; coloured points) against the the full X-ray emitting galaxy sample (black contours). The contours encompass 90%, 70%, 50%, 30% and 10% of the galaxy sample. See section 2.2 for more information on how  $L_{\nu,3550\text{\AA}}(\text{AGN}_{\text{predicted}}/(\text{AGN} + \text{Galaxy})_{\text{observed}})$  was calculated.

a 90% probability of association as the matching threshold which produced a well-matched sample of 1,559 X-ray emitting galaxies.

## 2.2 Clarifying the Effect of High Optical Emission on the SFRs & Stellar Masses

MPA-JHU stellar masses were calculated by fitting each galaxy's spectral energy distribution (SED) (Kauffmann et al. 2003b) using optical data from the SDSS; SFRs use the optical SEDs and add UV data (Salim et al. 2007) from GALEX (Martin et al. 2005) as well. As outlined above, we were confident that these values are robust and well-suited to our study. However, we expected to find high luminosity AGN in this sample. Thus we wanted to check that strong central optical emission, likely originating from an AGN, was not biasing these quantities within our sample of candidate AGN-hosting galaxies. To determine what proportion of the total optical emission originated from the AGN, we drew upon techniques previously used in Birchall et al. (2020). This technique aimed to calculate and compare the luminosity densities at  $3550\text{\AA}$  originating from the AGN and compare to it the overall optical emission ( $L_{\nu,3550\text{\AA}}(\text{AGN}_{\text{predicted}}/(\text{AGN} + \text{Galaxy})_{\text{observed}})$ ).

For this sake of this calculation we assumed that every galaxy in this sample hosted an AGN. To predict the AGN's contribution to the total observed optical light we used the X-ray to optical spectral index published in Lusso & Risaliti (2016). Using the X-ray emission ensured there was little contamination from stellar processes. The Lusso & Risaliti (2016) relation has the form,

$$\log_{10}(L_{\nu,2500\text{\AA}}) = \frac{1}{0.6}(\log_{10}(L_{\nu,2\text{keV}}) - 7) \quad (1)$$

It relates the luminosity density at 2 keV ( $L_{\nu,2\text{keV}}$ ) to that at  $2500\text{\AA}$  ( $L_{\nu,2500\text{\AA}}$ ) from a sample of SDSS quasars. Assuming our galaxies follow the same relation, it can provide a useful starting point in estimating the effects of the AGN on optical observations.

<sup>3</sup> Found at [https://www.ledas.ac.uk/flix/flix\\_dr7.html](https://www.ledas.ac.uk/flix/flix_dr7.html)

To find the luminosity density at 2 keV, we calculated the geometric means of 3XMM bands 2 & 3 and bands 4 & 5, giving the luminosity densities at 1 keV and 5 keV respectively. Using linear interpolation between these values we calculated  $L_{\nu, 2 \text{ keV}}$  and used it in equation (1) to find  $L_{\nu, 2500\text{\AA}}$ . Translating  $L_{\nu, 2500\text{\AA}}$  to the required emission at 3550\AA required the composite UV-optical quasar spectrum from [Vanden Berk et al. \(2001\)](#). They model this region as a power law spectrum,  $f_{\nu} = c\nu^{-\alpha}$ , where  $\alpha = 0.44$ . From this an average spectrum was constructed and the flux density at 3550\AA extracted.

The total observed optical emission was calculated by converting the SDSS U-band magnitude to a flux density at 3550\AA. We then divided the predicted AGN emission at 3550\AA by the total optical emission to obtain the  $L_{\nu, 3550\text{\AA}}(\text{AGN}_{\text{predicted}}/(\text{AGN} + \text{Galaxy})_{\text{observed}})$  ratio. We plotted the stellar mass and SFR distribution of the 1,559 X-ray emitting galaxies as black contours in figure 1. Overlaid on this figure are coloured points representing galaxies with high predicted AGN contributions.

The [Lusso & Risaliti \(2016\)](#) relation assumes the AGN are unobscured. Since MPA-JHU selects only narrow-line objects this predicted AGN luminosity will be an upper limit on its contribution. Thus some of these AGN have predicted optical contributions that are greater than the actual observed values. However, these sources constitute less than 20% of the galaxy sample. This fraction drops to  $\sim 7\%$  at  $L_{\nu, 3550\text{\AA}}(\text{AGN}_{\text{predicted}}/(\text{AGN} + \text{Galaxy})_{\text{observed}}) > 10$  and  $\sim 1\%$  at  $L_{\nu, 3550\text{\AA}}(\text{AGN}_{\text{predicted}}/(\text{AGN} + \text{Galaxy})_{\text{observed}}) > 100$ . Figure 1 clearly shows that even for these sources, with a strong predicted AGN contribution to the optical light, the stellar mass and SFR values calculated are consistent with the wider X-ray emitting galaxy distribution. Thus we continued to confidently use these values.

### 2.3 Identifying AGN

To identify emission from an AGN, we modelled the combined emission coming from other X-ray emitting sources - X-ray binary stars and hot gas emission - and compared it to the observed X-ray luminosity, following the process used in [Birchall et al. \(2020\)](#).

X-ray binary star contributions,  $L_{\text{XRB}}$ , were modelled using the relationship provided in [Lehmer et al. \(2016\)](#) (see also [Aird et al. \(2017\)](#)),

$$L_{\text{XRB}} = \alpha_0(1+z)^{\gamma} M_*(M_{\odot}) + \beta_0(1+z)^{\delta} \text{SFR}(M_{\odot}\text{yr}^{-1}) \quad (2)$$

where  $\log_{10}(\alpha_0) = 29.37 \pm 0.15$ ,  $\gamma = 2.03 \pm 0.60$ ,  $\log_{10}(\beta_0) = 39.28 \pm 0.03$  and  $\delta = 1.31 \pm 0.13$  for 2 - 10 keV. The low-mass X-ray binary (LMXB) contribution is correlated to the stellar mass ( $M_*$ ), the high mass X-ray binary (HMXB) contribution to the SFR and the redshift ( $z$ ) dependence accounts for changes in metallicity and evolution of the XRB population.

Hot gas in the interstellar medium could also account for some portion of the observed X-ray emission. Its contribution,  $L_{\text{Gas}}$ , can be estimated using the [Mineo et al. \(2012b\)](#) relation,

$$L_{\text{Gas}} = (8.3 \pm 0.1) \times 10^{38} \text{SFR}(M_{\odot}\text{yr}^{-1}) \quad (3)$$

We calculated the expected emission from the hot gas using [Mineo et al. \(2012b\)](#) and added this to their  $L_{\text{XRB}}$ . Despite the relatively

low magnitude of this relationship, it was still important to calculate as all significant alternative X-ray sources needed to be considered.

Once both these contributions had been calculated, we summed them and compared this quantity to the observed X-ray luminosity,  $L_{\text{X,Obs}}$ . We considered any object which met or exceeded the following criterion to be an AGN,

$$\frac{L_{\text{X,Obs}}}{L_{\text{XRB}} + L_{\text{Gas}}} \geq 3 \quad (4)$$

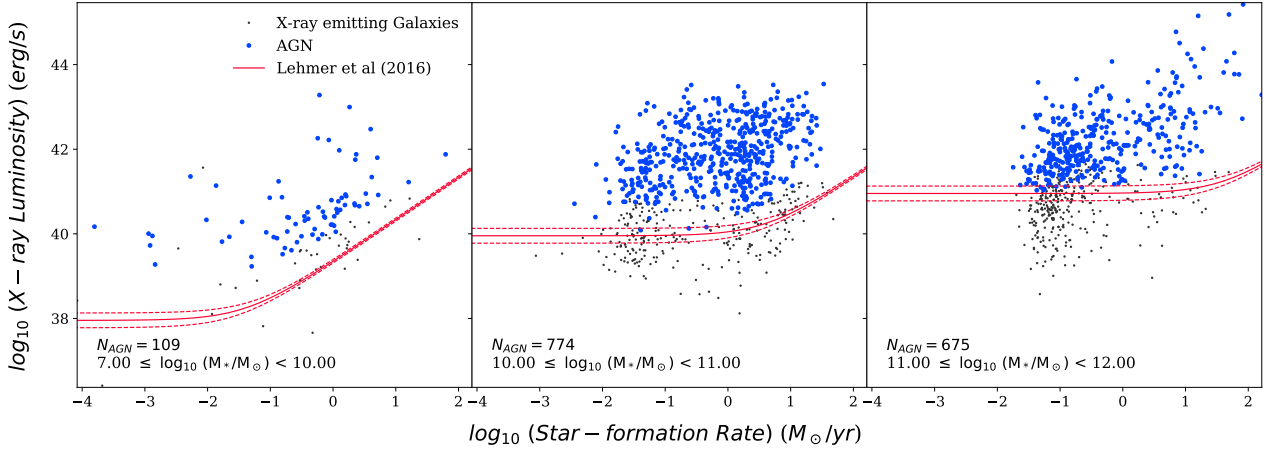
Based on this criterion, 949 X-ray emitting galaxies were classified as AGN. Figure 2 shows a breakdown of this classification process, comparing the observed X-ray luminosity with the SFR for every X-ray emitting galaxy and highlighting whether it has been classified as an AGN. The sample was split into several mass bins for clarity but it also highlights the changing nature of the [Lehmer et al. \(2016\)](#) equation. Each illustrated line was calculated using the mean stellar mass, indicated on the panel, and mean redshift of the bin. So each line should be considered an illustrative threshold. As the stellar mass increases, so too does the normalisation of equation (2) at low SFR. So for the highest mass galaxies, this relationship behaves almost like a luminosity cut of  $> 10^{41}$  erg/s, lower than the typical threshold of  $10^{42}$  erg/s ([Brandt & Alexander 2015](#)). At such low redshifts and SFRs, however, we trust that the selection criterion will continue to identify AGN activity. This increase in normalisation is believed to originate from a growing population of long-lived low mass X-ray binaries that occurs in more massive galaxies. Variation due to uncertainty in coefficients is highlighted by the dashed lines.

To measure an accurate AGN fraction for this sample we had to ensure it was a statistically complete sample of galaxies above a given stellar mass limit. Thus, we took galaxies of all masses in narrow bands of redshifts from MPA-JHU and determined the mass range which contained  $\sim 90\%$  of the galaxies. This process allowed us to create a 90% mass completeness function. Figure 3 shows the distribution of AGN on the redshift and stellar mass plane, with the completeness function highlighted as a red line. We have split this into 6 redshift intervals between 0 and 0.35 - 5 of width 0.05 and the final being 0.1 in width. These bins are highlighted by their colour in the figure. Only those AGN which lie above the completeness function were binned, resulting in the removal of 32 objects.

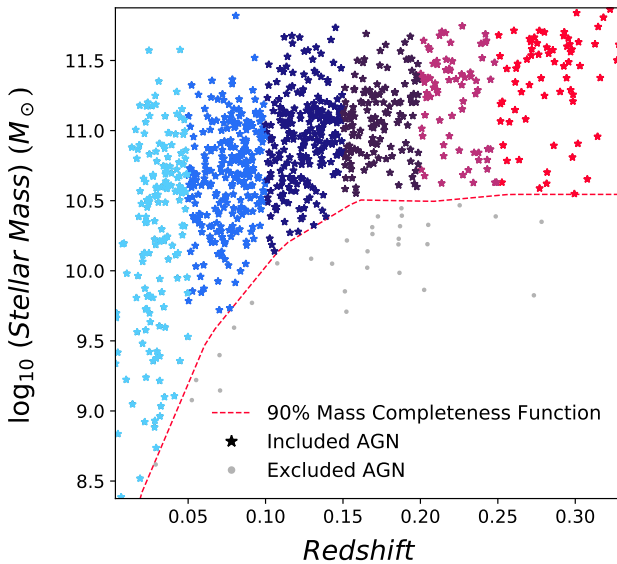
### 3 BPT CLASSIFICATION

AGN activity can impact the host galaxy's emission across the electromagnetic spectrum. The BPT diagnostic ([Baldwin et al. 1981](#)) is a commonly used technique to identify the primary source of ionising radiation in the optical part of the spectrum. By comparing the ratios of various emission lines, we can gain insight into whether star formation, AGN or a composite of both processes dominate in any given galaxy. In [Birchall et al. \(2020\)](#), we found that this diagnostic missed around 85% of our X-ray selected AGN in dwarf galaxies. Now with this larger sample we were able to investigate how the accuracy of this diagnostic changed with stellar mass.

Of the 917 AGN hosts we identified using X-ray selection techniques, 658 had significant detections ( $\frac{\text{Line Flux}}{\text{Line Flux Error}} > 3$ ) in each of the required emission lines so these were used in our



**Figure 2.** Observed X-ray luminosity against SFR for the 1,559 X-ray emitting galaxies which are candidates to host AGN. For clarity, these candidates were split up into three panels of different host galaxy masses (highlight in text on each panel). Objects that met or exceeded the threshold are plotted in blue, those that did not are grey. The red line is the [Lehermer et al. \(2016\)](#) equation, the dominant source of X-ray emission prediction, described in section 2.3. Variation due to uncertainty in coefficients is highlighted by the dashed lines. By comparing it across plots we can clearly see the variation in mass and SFR dependencies.



**Figure 3.** Redshift against stellar mass for the 949 AGN identified with equation (4). This sample has been limited by the 90% mass completeness function (red, dashed line), resulting in the removal of 32 AGN (grey points). The remaining 917 AGN have been split into 6 redshift bins outlined in section 2.3, and separated by colour for clarity.

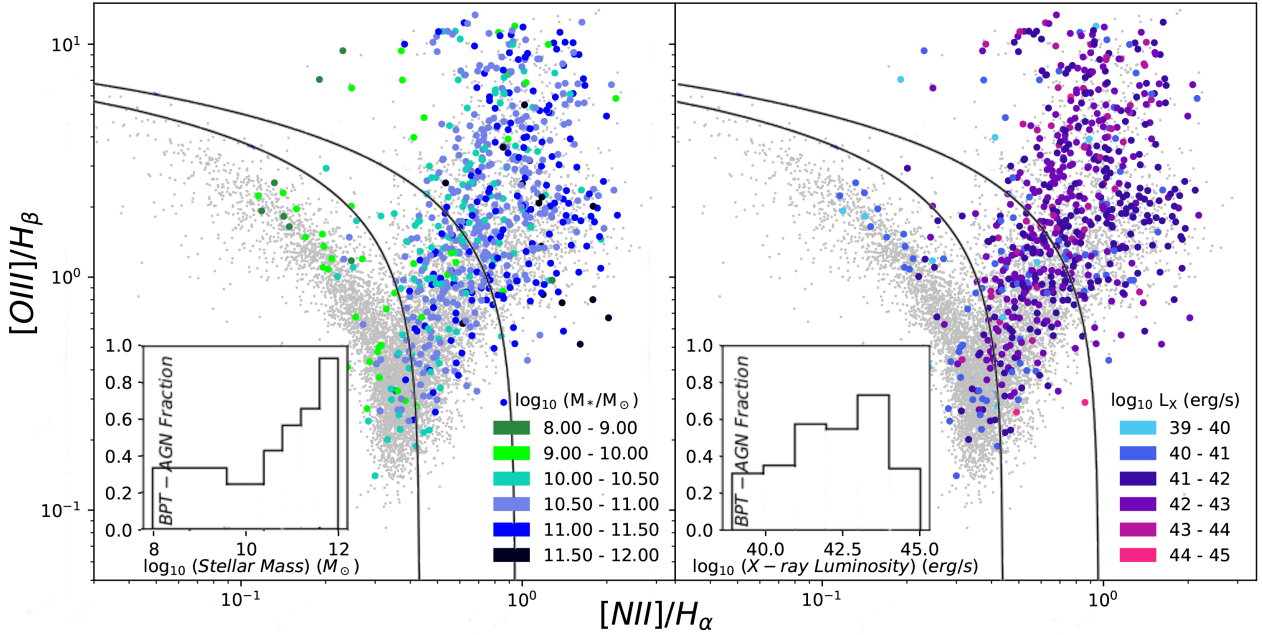
analysis. In the left-hand panel of figure 4, we show the results of the BPT analysis with our AGN sample plotted as large points, coloured to indicate their stellar mass. Black lines separate the AGN hosts into different classifications: objects with ionisation signatures predominately from AGN lie in the top-right, those dominated by star formation in the bottom-left, and those that have composite spectra are in the central region ([Kewley et al. 2001](#); [Kauffmann et al. 2003a](#)). Underneath these points, in grey, are a subset of the MPA-JHU galaxies used to illustrate the underlying BPT distribution. Encouragingly, the vast majority of our AGN appear to lie within the AGN region. However, the colours clearly show that most of the objects found here are higher mass galaxies, with the lower mass points favouring the star-forming

and composite regions. The behaviour exhibited by these low mass galaxies mirrors what we found in [Birchall et al. \(2020\)](#) and is consistent with a growing body of work suggesting that these optical spectroscopic measurements are less effective at identifying AGN in lower mass galaxies ([Agostino & Salim 2019](#); [Cann et al. 2019](#)).

To further interrogate this trend, we calculated the proportion of X-ray-selected AGN with the same BPT classification as a function of stellar mass. Inset into the left-hand panel of figure 4, we can see that the effectiveness of the BPT diagnostic in reproducing our X-ray based classification increases with stellar mass. At the low mass end we can see similar behaviour to that found in our previous work, with the diagnostic missing around 70% of our dwarf galaxy AGN. However, the accuracy rises steadily until it identifies 93% of our AGN in the highest mass host galaxies.

Our previous paper concluded that the AGNs' relatively low luminosity caused their stark mis-classification. In the right-hand panel of figure 4 we investigate how the changing X-ray luminosity affects the BPT classification. Inset in this panel is a distribution showing the changing BPT accuracy with X-ray luminosity. Whilst there is an increase in BPT accuracy, it clearly peaks before reaching the highest luminosity AGN: rising from 31% accurate, in the lowest mass bin, to 73% in the most populous. Clearly then, increasing X-ray luminosity is not the only variable that affects mass dependence of the BPT accuracy.

[Behroozi et al. \(2019\)](#) collated the quenched fraction of galaxies in the nearby universe (based on [Bauer et al. 2013](#); [Moustakas et al. 2013](#); [Muzzin et al. 2013](#)) and showed it increases strongly with stellar mass. This means higher mass galaxies are much less likely to have their emission lines driven by star formation activity. As we move towards higher mass galaxies, the typically higher luminosity AGN combined with this drop in star formation would see AGN-driven signatures dominate and produce the concurrent increase in BPT accuracy.



**Figure 4.** BPT diagnostics for the 658 X-ray selected AGN with significant detections in all the emission lines shown. In the left-hand panel, the colour of each AGN point represents the mass of the host galaxy, outlined in the legend. In the right-hand panel, the colour of each AGN point represents the observed X-ray luminosity of the AGN. Inset into each panel, we see the fraction of X-ray selected AGN with the same classification from the BPT diagnostic, as a function of the relevant observable.

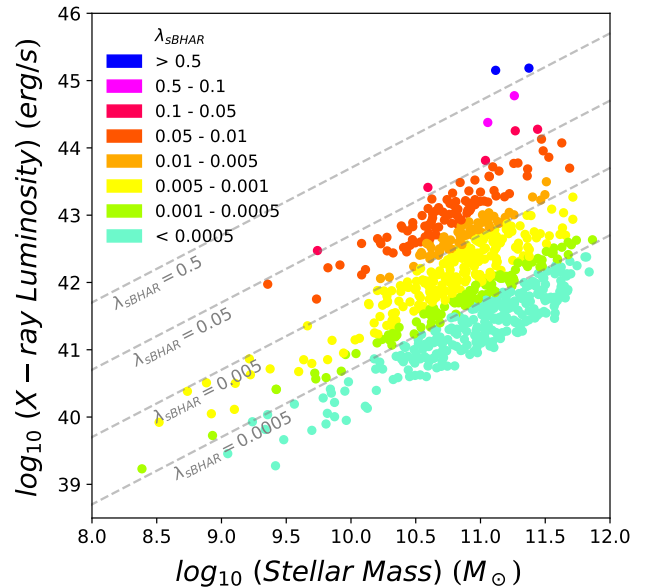
#### 4 SPECIFIC BLACK HOLE ACCRETION RATE

AGN are powered by mass accretion onto a galaxy’s central supermassive black hole however we do not gain any insight into this accretion process by only looking at the observed X-ray luminosity. For example, one black hole growing at a higher accretion rate in a lower mass galaxy could produce a similar X-ray luminosity to another black hole with a lower accretion rate in a high mass galaxy. To break this degeneracy and understand more about how the black hole activity is distributed across our observed AGN sample, we looked at the specific black hole accretion rate (sBHAR),  $\lambda_{\text{sBHAR}}$ . This quantity compares the bolometric AGN luminosity of the galaxy with an estimate of the black hole’s Eddington luminosity to give an indication of how efficiently the black hole is accreting. It is found using,

$$\lambda_{\text{sBHAR}} = \frac{25L_{2-10\text{keV}}}{1.26 \times 10^{38} \times 0.002M_*} \approx \frac{L_{\text{bol}}}{L_{\text{Edd}}} \quad (5)$$

and is taken from [Aird et al. \(2012\)](#). Our use of  $0.002M_*$  implies a correlation between the mass of the stellar bulge and SMBH which is dependent on galaxy morphology, among other properties (e.g. [Blanton & Moustakas 2009](#); [Jahnke et al. 2009](#)). However, our intention behind using this correlation is not to accurately recreate an Eddington ratio, instead it is to present an Eddington-scaled accretion rate quantity for our galaxy sample. By using this scaling to calculate sBHAR - a tracer of the rate of black hole growth relative to the host galaxy’s stellar mass - we can directly compare our results to prior work at higher redshifts.

Figure 5 shows how our observed AGN sample are distributed on the stellar mass and X-ray luminosity plane, with the data point colour giving an indication of the sBHAR of each galaxy. We have also highlighted lines of constant sBHAR to give a sense of what luminosities would be expected from a host galaxy of a given mass



**Figure 5.** Stellar mass against the observed X-ray luminosity for the 917 AGN detected in section 2.3. Each point has been assigned a colour to provide an indication of the accretion rate ( $\lambda_{\text{sBHAR}}$ ) of the galaxy’s central supermassive black hole, calculated in section 4. Several grey lines of constant sBHAR have also been plotted for reference.

at those accretion rates. This figure further highlights that the vast majority of the observed AGN sample lies above stellar masses of  $10^{10} M_{\odot}$ . There is also a clear preference for black holes with relatively low accretion rates. Most of them have accretion rates that are less than 0.5% of their Eddington luminosity and only a handful of the most massive galaxies venturing above 10%.

## 5 COMPLETENESS-CORRECTED PROBABILITY DISTRIBUTIONS

Despite using AGN selection criteria that modify the X-ray luminosity threshold depending on the mass, SFR and redshift of the host galaxy, figure 5 clearly shows that our sample as it is currently constructed preferentially identifies AGN found in the highest mass galaxies. In this section we will summarise and apply the methods developed in [Birchall et al. \(2020\)](#) to reduce the effects of this bias. In doing so we will create a series of probability distributions comparing different configurations of host galaxy properties. From these distributions we aim to understand more about the underlying distribution of AGN in the nearby Universe.

### 5.1 Calculating Completeness Corrections

Using the 3XMM serendipitous sky survey as our X-ray catalogue introduced a significant amount of variation in the detection sensitivity. Such variation brings with it the possibility that lower luminosity AGN within the overlapping region of 3XMM and MPA-JHU could have been missed as the flux limit was insufficient to detect it. To overcome this, we established the fraction of the parent galaxy sample that lie within 3XMM fields with sufficient sensitivity to detect an AGN above a given luminosity threshold.

To perform this analysis, we made use of Flix ([Carrera et al. 2007](#)), 3XMM's upper limits service. It provides upper limits broken down by band and instrument for the whole of 3XMM. From this data we chose the upper limits and observed fluxes from band 8 in the PN camera as it covers the entire energy range of 3XMM. Because of this restriction, we also limited the observed AGN to only those with detections in that same band. Not all objects will meet this detection threshold and some XMM fields will not have been observed by the PN camera so our observed AGN sample was reduced from 917 to 739.

To characterise this sensitivity variation, we needed to find the X-ray detection upper limits of all the galaxies in the region of MPA-JHU that have coverage from 3XMM. There were 28,545 MPA-JHU galaxies found within 3XMM whose co-ordinates were uploaded to Flix. It returned the X-ray flux upper limits at the co-ordinates of 25,949 galaxies – our parent sample. We then extracted the flux upper limits for a given range of stellar masses and redshifts. These fluxes were then converted into luminosities using the redshift of the galaxy associated with each upper limit. Finally, these upper limits were used to construct a cumulative histogram function normalised by the total number of galaxies in the current mass and redshift range. These luminosity sensitivity functions allowed us to determine the fraction of galaxies where an AGN could have been detected above a given X-ray luminosity within a range of masses or redshifts.

### 5.2 Creating the Probability Distributions

We are interested in calculating how the probabilities of finding AGN vary in the nearby Universe as a function of stellar mass and redshift. To illustrate the process of constructing these probability distributions, we will focus on how the probability varies with X-ray luminosity and stellar mass.

For this configuration of host galaxy properties, we split up the parent sample into a series of stellar mass bins. In each of these

stellar mass bins, the data was further broken down as a function of X-ray luminosity, producing an observed AGN count distribution. In addition, we constructed a bespoke luminosity sensitivity function, as outlined in section 5.1. Each bin's count distribution was divided by the correction fractions extracted from this sensitivity function to recover the expected number of galaxies sensitive enough to detect an AGN as a function of luminosity. Finally, we divided these corrected AGN counts by the parent sample size in this bin to produce a series of probability distributions. The results of this process are shown in figure 6. It shows the probability of finding an AGN as a function of luminosity in bins of increasing mass. Some useful reference information is printed on each panel including the sizes of the AGN and parent samples, and the stellar mass range of the galaxies included therein.

By applying corrections extracted from the luminosity sensitivity function to the observed AGN counts we were able to provide robust measurements of the true incidence of AGN within the nearby galaxy population. In figure 6 we see that there is an abundance of AGN across stellar mass despite the clear favouring of higher mass AGN seen in figure 5. Of particular significance are the continuation of the findings from [Birchall et al. \(2020\)](#): AGN populations are well described by power law distributions, with AGN being found predominantly at lower X-ray luminosities (see also [Aird et al. 2012](#), and others).

To calculate the errors in each probability data point we used the confidence limits equations presented in [Gehrels \(1986\)](#) meaning that the size of the error is determined based on the number of detected AGN in a given bin.

One final check was performed on the AGN populations shown in figure 6. Section 2.3 highlights how we have considered contamination from X-ray binaries and hot gas, but there remained the possibility that this emission could have originated from an ultra-luminous X-ray source (ULX), particularly in lower mass galaxies. To check whether a ULX could account for this emission, we used the [Mineo et al. \(2012a\)](#) X-ray luminosity function (XLF). XRB populations, including ULXs, were modelled using a two-part power law normalised by the host galaxy's SFR. In each stellar mass bin, XLFs associated with each individual galaxy were calculated and averaged to show how the predicted number of ULXs compared to our AGN observations. The blue lines in each panel of figure 6 show that the vast majority of the data points do not overlap with the ULX XLFs. The only point at which the two distributions do meet is at the low luminosity end of the lowest stellar mass bin.

Since the overwhelming majority of our sample are all confirmed AGN we could confidently fit power laws to each of these panels. They have the following form,

$$p(X) = A \left( \frac{X}{x'} \right)^k d \log_{10} X \quad (6)$$

where  $p(X)$  is the probability of observing an AGN with corresponding X-axis quantity,  $X$  and centred on a value  $x'$ . For this probability distribution configuration, each power law is centred on the median luminosity for the sample,  $\log_{10} x' = 42.1$  (and  $\log_{10} x' = -2.55$  for sBHAR). The power laws are shown as dashed red lines in each panel and more clearly identify how the



probability of finding an AGN, within the highlighted mass range, changes as a function of X-ray luminosity. The pale red regions surrounding each power law function give an indication of the error in each fit. This was calculated by performing a  $\chi^2$  fit with equation (6) to the black data points in each bin and their associated errors. Fit parameter errors were estimated by taking the square-root of the covariance matrix's diagonal. With this we could outline the extent of the uncertainty in each fit. Encouragingly, in nearly all of the bins the power law fits are appropriate.

### 5.3 Probability Distribution Comparison

Using the method outlined in section 5.2, we created a number of other probability distributions to investigate how different configurations might shed light on the AGN population in the nearby Universe. As can be seen in figure 7, we have added redshift and sBHAR to our properties of interest. By including redshift in our range of host galaxy properties, we can gain some insight into how the changing conditions of the Universe might affect the AGN population. Each configuration has been split up into 6 colour-coded bins of the corresponding property – either redshift (top row) or stellar mass (bottom row) – and had a probability distribution calculated from the data points in that region. For ease of comparison, each set of distributions has been placed in the appropriate column - sBHAR on the left, and X-ray luminosity on the right. The results from figure 6 have been placed in the bottom-right panel of this figure. For clarity, we have only included the fits and their error region in each configuration panel. Appendix A outlines the best-fit coefficients, and associated errors, used in equation (6) to create these probability distributions. In addition, the full set of data points for each configuration can be found in appendix B. Overall, figure 7 shows us that regardless of the host galaxy property configuration, the average number of AGN in the nearby Universe are well described by a power law. However, there are key differences between these configurations which we will outline in this section.

#### 5.3.1 Slope

In each panel of figure 7, there is some evidence of the slope changing with the respective property. Along the top row, we can see a series of slopes steadily steepening with redshift. This steepening effect is most pronounced in the right-hand X-ray luminosity column. However, we see in the top-left panel that AGN at higher redshifts favour more moderate accretion rates. This drives the increase in the incidence of relatively low luminosity AGN emission for the corresponding redshift bin in the top-right panel. A similar steepening can be seen in the bottom-right stellar mass, X-ray luminosity panel, however the trend is much less consistent. It shares the steepening seen in the redshift, X-ray luminosity panel but there is a lot of variation between these two extremes.

#### 5.3.2 Normalisation

As outlined in section 4, studying sBHAR allows us to break observational degeneracies associated with X-ray luminosity so we can see how mass accretion is distributed across the AGN population. What is most striking about these distributions is that those in the left-hand column of figure 7 all appear to be within the same limits. Regardless of redshift or stellar mass, there is little change in the normalisation, implying that the average amount of

material a black hole accretes remains relatively consistent in the nearby Universe.

This trend is in stark contrast to that seen in the luminosity plots, in the right-hand column of figure 7. Both redshift and stellar mass distributions show consistently increasing normalisations. There are a number of processes that could be feeding into this shift: the magnitude limitation of the galaxy sample removing high redshift, low luminosity AGN; and the stellar mass and redshift dependencies of the Lehmer et al. (2016) XLF in equation (2). However, these effects only explain why the lower limits increase and don't account for full shifting of these luminosity-dependent distributions. This can be understood if we consider the lines of constant sBHAR in figure 5: for a fixed accretion rate, stellar mass determines the range of observable luminosities. At higher stellar masses the much more common, lower accretion rate sources have higher observed luminosities. When we consider that these mass and redshift bins have the same accretion limits, it is clear that changing stellar mass is the primary driver of this shift in normalisation, whether caused by explicit binning or the effect of magnitude limitation.

## 6 AGN FRACTIONS

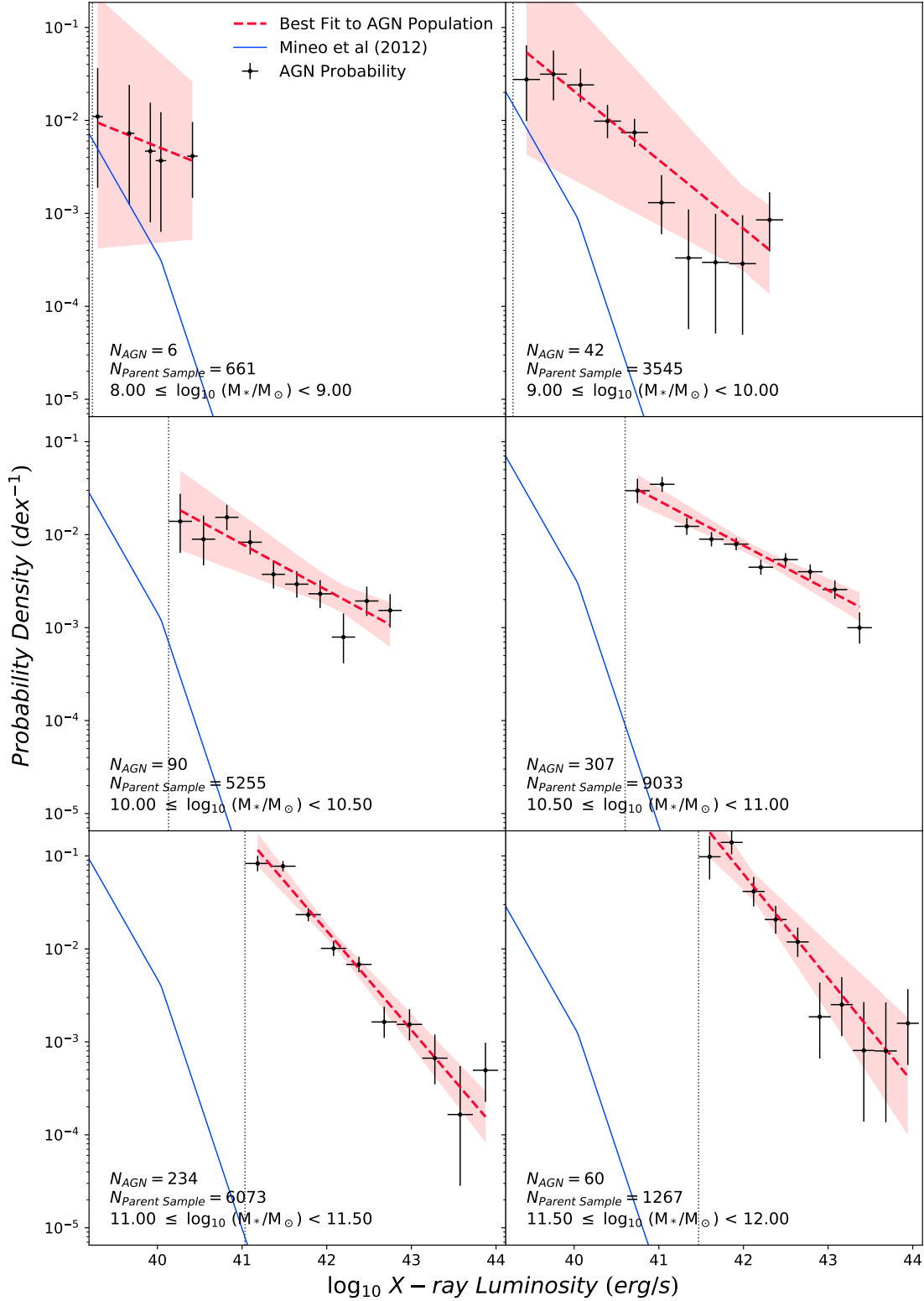
Using the fits produced in section 5.3, we could calculate how the AGN fraction varied with stellar mass and redshift in the nearby Universe. The results of these calculations are shown in figure 8. In this section we will outline this calculation process and the significance of the results.

AGN fractions were calculated by integrating under the probability distributions shown in figure 7. Each distribution produces a single point in the corresponding position, colour and panel of figure 8. Upper and lower limits on each AGN fraction were calculated by integrating both edges of the error region. The limits of these integrations are shown in the bottom-right-hand corner of the stellar mass panel. Consistent limits were chosen to encapsulate the region within which the majority of the distributions are defined and to make sure each point is directly comparable.

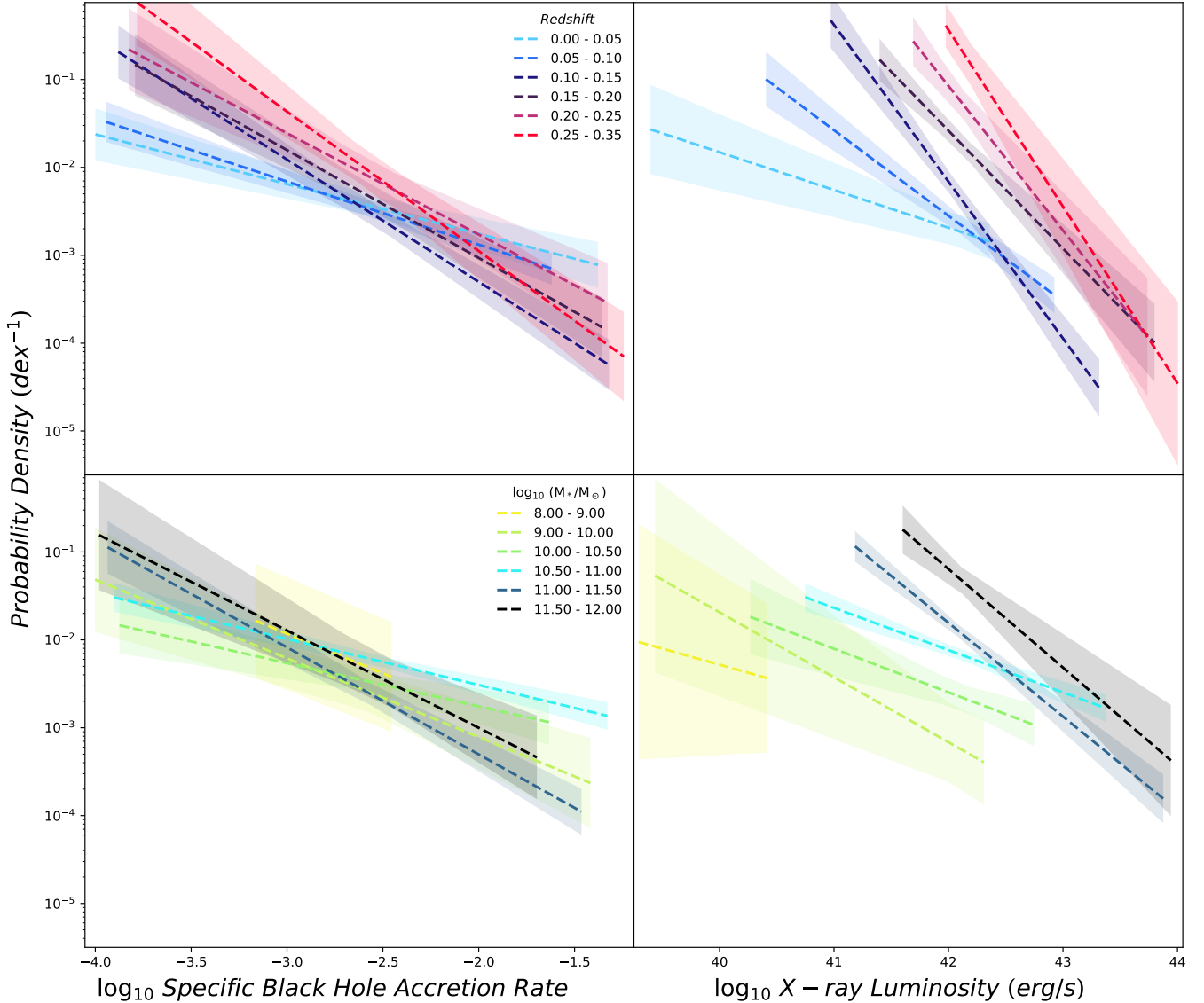
This approach leaves us with two sets of AGN fractions in each galaxy property panel: luminosity-derived fractions (blue stars) and sBHAR-derived fractions (red dots). In effect, we have created two separate AGN definitions to compare. The luminosity-derived points create an observational-style definition whereby an AGN is an object emitting at an X-ray luminosity  $\geq 10^{42}$  erg/s. Whereas the sBHAR-derived points define an AGN as a black hole accreting at rates  $\lambda_{\text{sBHAR}} \geq 10^{-3.5}$ . As is evident from figure 8, these differing definitions have implications for the predicted distribution of AGN.

In the left-hand panel, we can see how the two AGN fraction definitions vary with stellar mass. The sBHAR-derived fractions predict a constant occupation of just over 1% across the stellar mass range. Whereas the luminosity-derived fractions predict a much lower AGN fraction at low stellar mass before rising to agree with sBHAR. Georgakakis et al. (2011) found a similar increase with stellar mass in their sample of X-ray selected AGN with  $\log_{10} L_X > 41$ .

In the right-hand panel, we can see how both AGN fraction definitions vary with redshift. The sBHAR-derived fractions show a steady increase rising from 1% to 10%. The luminosity-derived fractions describe an even steeper increase from about 0.1% to



**Figure 6.** Probability of finding an AGN in the nearby Universe, using the completeness-corrected AGN samples, as a function of observed X-ray luminosity and stellar mass. Power laws (dashed red lines) have been fit to the data in each panel and displayed alongside their  $1\sigma$  uncertainty (pale red region). Information about the number of PN-8-detected AGN and parent galaxies in each mass range are printed on the panels. The blue line shows the average number of high luminosity stellar sources like ULXs. The form of these lines are based on the Mineo et al. (2012a) XLF, scaled by the average SFR of the galaxies in each bin. Dashed grey lines indicate the AGN detection limit in each bin. See section 5.2 for more details on how these plots were constructed.



**Figure 7.** A comparison of the probability distributions and associated error regions calculated using different configurations of host galaxy properties. The left-hand column looks at how the probability varies with sBHAR and the right-hand column shows this with X-ray luminosity. The top row bins the AGN in redshift intervals, the bottom row does this in stellar mass. The fits for the bottom-right hand plot were taken from figure 6, the rest are presented in appendix B.

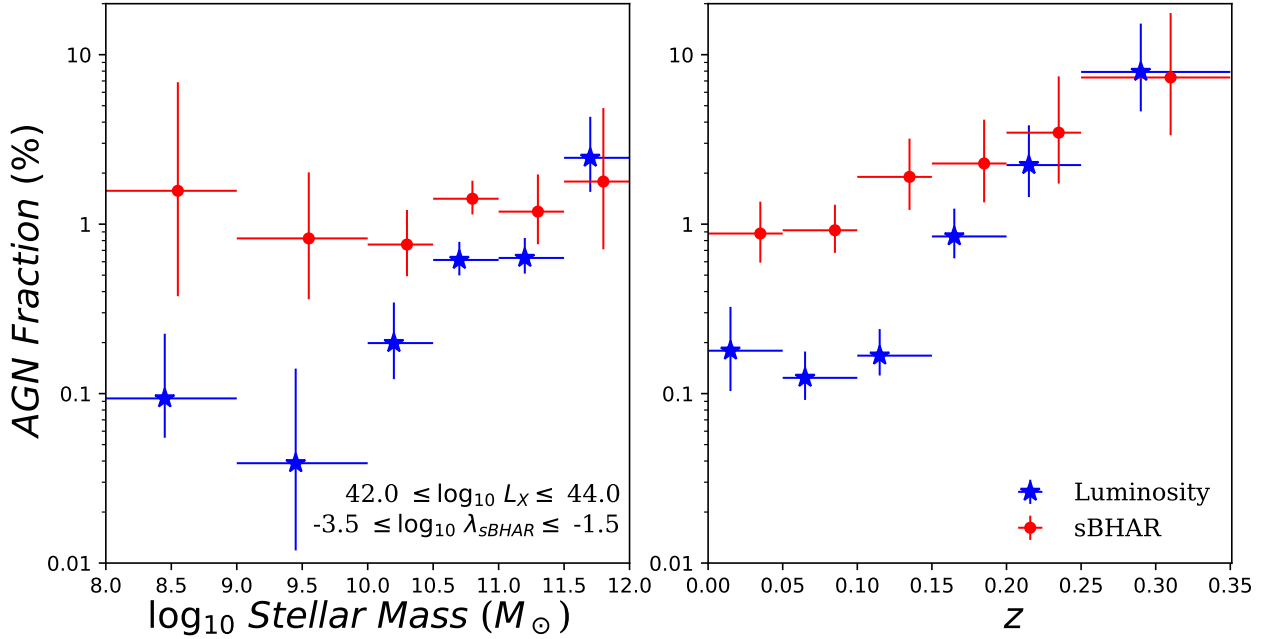
10%, agreeing with the sBHAR-derived data at the highest redshifts.

Further to the discussion at the beginning of section 2, we also calculated these fractions using the Granada values. The trends in both luminosity- and sBHAR-derived AGN fractions seen in figure 8 were also measured when using the Granada data.

Clearly, these two AGN fraction definitions follow different trends. Since the luminosity-derived fractions only consider higher-luminosity AGN, the sample used to calculate these will be restricted and so a drop in AGN fraction is likely. As discussed previously, all our AGN appear to be accreting over a similar range of sBHAR, so stellar mass is the main driver of changes in observed X-ray luminosity. Therefore, we would expect that lower mass AGN, which are much less likely to meet such high luminosities, are the objects most likely being excluded. In addition, figure 3 shows that most of the lowest mass AGN are found at the lowest

redshifts. So this would also concentrate this deficit to the lowest redshift regimes. To pinpoint where this deficit occurs, we first split the AGN by stellar mass - those above and below the median stellar mass of  $10^{10.89} M_{\odot}$  - and recalculated the AGN fractions as a function of redshift. The top row of figure 9 confirms our expectations by showing that this luminosity-derived fraction deficit very clearly occurs in the lowest stellar mass and redshift galaxies. The right-hand, high-mass panel shows that when both limits encapsulate the majority of the AGN sample, they predict very similar trends. Thus, the luminosity-derived fractions are still subject to observational biases despite being calculated from completeness-corrected data. Other studies, with different samples, have observed this effect as well (e.g. Aird et al. 2012; Bongiorno et al. 2016; Weigel et al. 2017).

Whilst the sBHAR-derived points are not free from these biases, selecting AGN based on an Eddington ratio limit has been shown to



**Figure 8.** AGN fractions as a function of stellar mass (left-hand panel) and redshift (right-hand panel) calculated from the probability distributions in figure 7. Each probability distribution was integrated within the relevant limit shown in the bottom-left of the left-hand panel. There is a clear difference in AGN fraction depending on whether the luminosity (blue stars) or sBHAR (red circles) distributions were integrated.

yield a wider range of AGN and host galaxy properties compared to a luminosity limit and so better represents the underlying AGN population (Jones et al. 2017). It is for these reasons that we focus on the sBHAR-derived points when drawing conclusions about the underlying AGN population in the nearby Universe. The left-hand panel of figure 8 showed us that sBHAR-derived AGN fraction was constant across the stellar mass range. Furthermore, the bottom row of figure 9 shows that when we split the AGN by redshift - those above and below the median redshift of 0.11 - and recalculate the AGN fractions as a function of stellar mass, that this flat distribution remains. Thus, the sBHAR-derived AGN fraction does not depend on stellar mass in the nearby universe.

Figure 10 compares our sBHAR-derived AGN fractions with other predictions derived from Eddington ratio-based probability distributions: Aird et al. (2012); Bongiorno et al. (2012); Wang et al. (2017); Aird et al. (2018); Yang et al. (2018); Birchall et al. (2020). We have increased the limits of integration compared to figure 8 to incorporate the generally higher threshold of AGN activity used in these other studies. By shifting the lower limit from  $\log_{10} \lambda_{\text{sBHAR}} = -3.5$  to  $-3$ , we have halved the average AGN fraction predicted from our data to 0.5%. This shows that there is a significant amount of low level accretion in the nearby Universe.

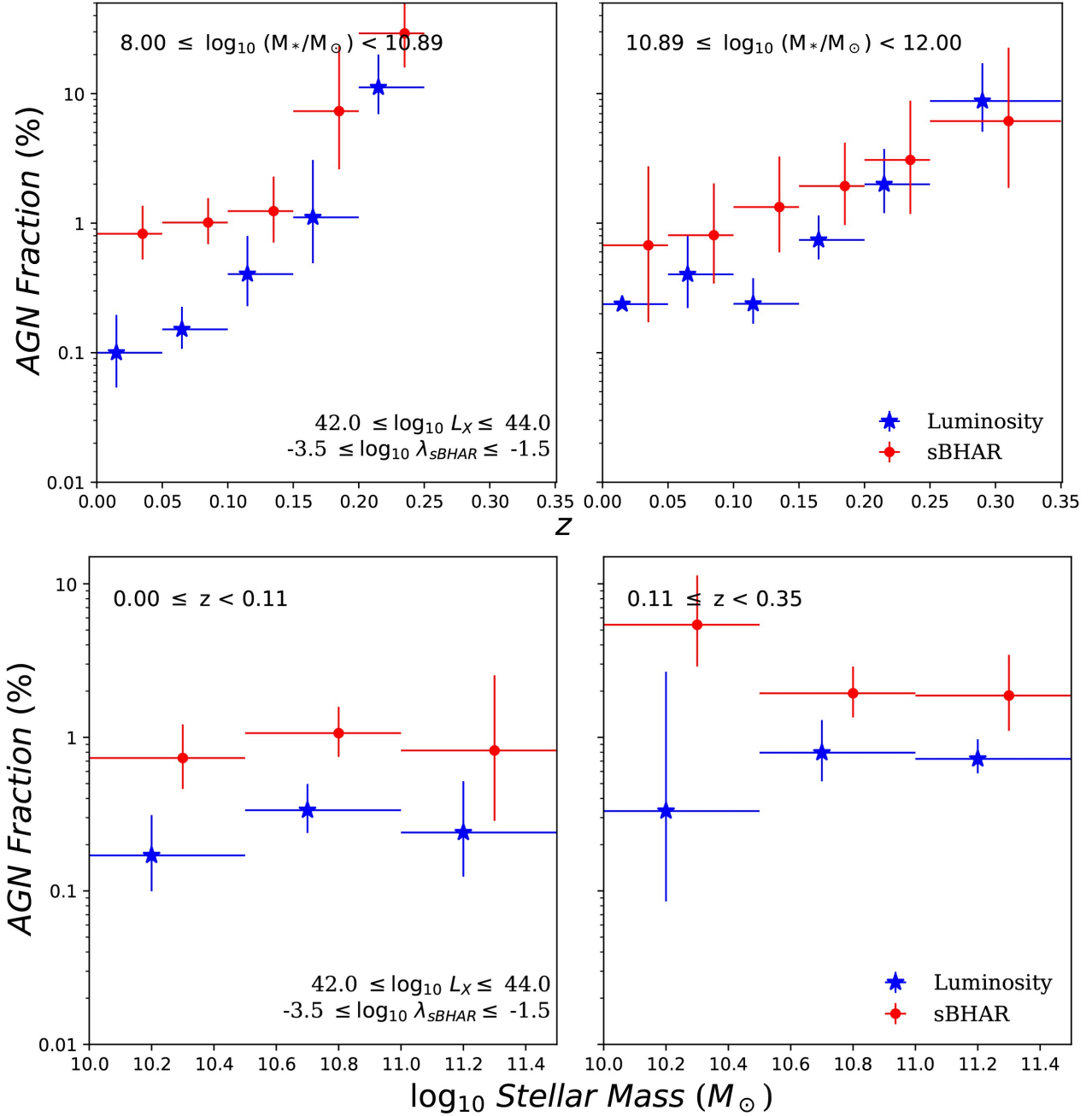
Our results provide one of the first, robust measurements of the incidence of AGN within the nearby galaxy population. To highlight this we presented the median redshift of each paper’s sample in the legend. Overall we see that as redshift increases so too does the AGN fraction. It is thought that the increased availability of cold gas at earlier times (Mullaney et al. 2012; Popping et al. 2012; Vito et al. 2014) would drive this increase in activity. Other AGN samples taken from a range of redshifts and wavelengths appear to show similar patterns of activity (Aird et al. 2018; Delvecchio et al. 2018; Yang et al. 2018). To check whether our

results are consistent with this evolution, we re-integrated equation (12) of Aird et al. (2012) between the same sBHAR limits shown in figure 10. However, this time we modified the redshift-dependent normalisation term by evaluating it at  $z = 0.11$ , the median redshift of our work. This produced a new fraction of 1.1%, consistent with most of our fractions.

However, figure 10 shows that this evolution is suppressed in lower mass galaxies. We can see significant alignment between points accounting for AGN samples measured out to  $z = 1$  in the dwarf galaxy mass range ( $\log_{10} M_*/M_\odot \leq 9.5$ ); consisting of our results, and those from Birchall et al. (2020) and Aird et al. (2018). It is thought that since the potential wells for lower mass galaxies are shallower, gas particles are less likely to fall into the galactic centre and fuel the SMBH (Bellovary et al. 2013). A similar drop in AGN activity at lower stellar masses has also been observed in different AGN samples (Yang et al. 2017; Aird et al. 2018).

## 7 SUMMARY & CONCLUSIONS

This study has rigorously matched a sample of local galaxies from MPA-JHU to their X-ray counterparts in 3XMM using *xmatch*. We re-affirmed that MPA-JHU was the catalogue most suited for this task. Then, before identifying AGN, we confirmed that strong nuclear optical emission from our matched sample did not bias the stellar mass and star-formation rate measurements. With this confirmation we could use these quantities to confidently predict the X-ray luminosity due to emission from X-ray binaries and hot gas. Any galaxies that had observed X-ray emission at least three times larger than the sum of these predictions was considered an AGN host. Finally, any AGN that fell outside of the 90% mass completeness function were removed from the sample. This left us with 917 AGN.

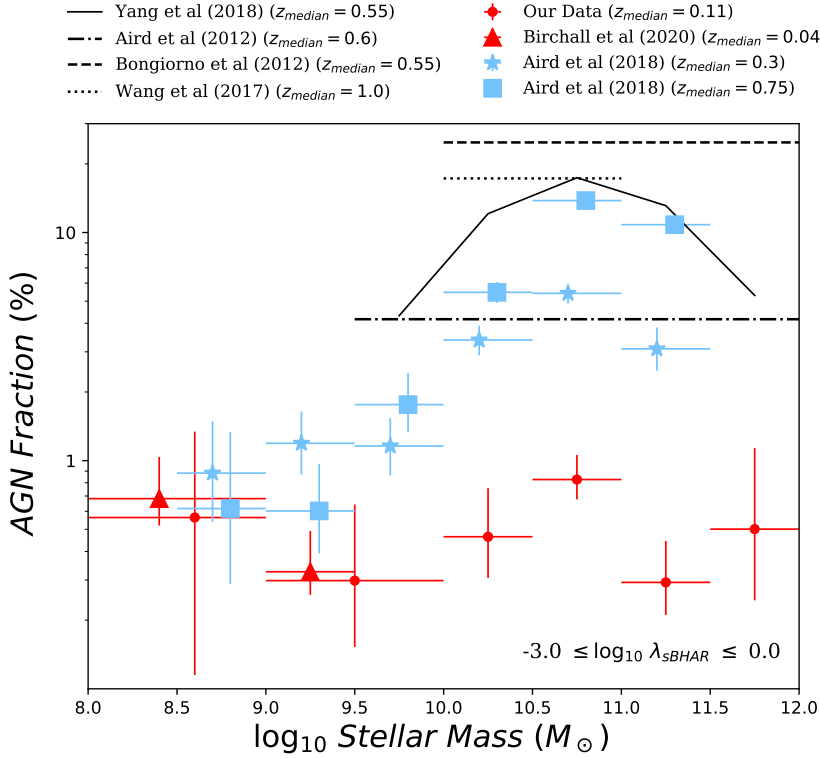


**Figure 9.** The top row depicts AGN fractions as a function of redshift split up into stellar mass bins - those above and below the median mass. Below are the AGN fractions as a function of stellar mass split up into redshift bins - those above and below the median redshift. The limits between which the full AGN sample has been split are shown in the top-left of each panel. As with figure 8, the luminosity-derived fraction are blue stars and the sBHAR-derived fractions are red circles.

We then performed BPT analysis on the 658 AGN with significant detections in all relevant emission lines. We have demonstrated that there is a strong stellar mass dependence in the BPT diagnostic. Higher mass X-ray selected AGN are much more likely to be assigned the same classification by the BPT diagnostic. In contrast, we find little dependence in the BPT-selected fraction of AGN on X-ray luminosity. Since the quenched fraction of galaxies is believed to increase with stellar mass, then the influence of star-formation on these emission lines would drop significantly and

produce the increase in BPT accuracy.

Next we investigated the activity of the black holes powering our AGN by calculating their specific accretion rates. We found that there is a clear preference for black holes with relatively low accretion rates. Most of them have accretion rates that are less than 0.5% of their Eddington luminosity (assuming a nominal scaling between black hole mass and galaxy stellar mass), with only a handful of the most massive galaxies venturing above 10%. Only two AGN have an



**Figure 10.** A comparison of how AGN fraction changes as a function of stellar mass at different redshifts. Our data (red circles) is calculated from the sBHAR-dependent fits. Alongside this are numerous other AGN fractions calculated from Eddington ratio based probability distributions within the limits shown in the bottom-right of the plot. The median sample redshift is highlighted in the legend.

accretion rate that is greater than 50% of their Eddington luminosity.

We then corrected for observational bias in our AGN sample. Drawing on the method used in [Birchall et al. \(2020\)](#), we used `Flix` to determine the flux limits of the 25,949 MPA-JHU galaxies found within 3XMM - our parent sample. These upper limits form the basis from which a sensitivity function can be created and used to correct our observed distributions of AGN. From this we created a series of probability distributions splitting the AGN population by stellar mass and redshift, and looking at trends with observed X-ray luminosity and specific black hole accretion rate. These distributions are well-described by power laws but when compared show distinct trends. As redshift increases, the power laws steepen suggesting more moderate accretion rates and thus luminosities, are favoured at earlier times. A similar steepening trend can be seen with stellar mass, however, the progression is less certain. We also find our galaxies contain AGN accreting over approximately the same, broad range of specific accretion rates ( $-3.5 \lesssim \log \lambda_{\text{sBHAR}} \lesssim -1.5$ ) regardless of stellar mass or redshift.

Finally, integrating under each of these distributions allows us to calculate robust AGN fractions and determine how these vary with host galaxy stellar mass and redshift. Since we have investigated the AGN distribution as a function of both X-ray luminosity and specific black hole accretion rate, we adopt two definitions of an AGN when determining the fraction. The first is pseudo-observational, defining an AGN an object with X-ray luminosity  $\geq 10^{42}$  erg/s; the second defines an AGN as a black hole accreting at rates  $\lambda_{\text{sBHAR}} \geq 10^{-3.5}$ . We find that the pseudo-observational luminosity-limited definition reproduces observational biases inherent in previous AGN studies

so we focus on the accretion rate derived fractions. We find no evidence that our accretion-rate-derived AGN fractions depend on the stellar mass, finding a constant fraction of just over 1% for  $8 \lesssim \log M_*/M_\odot \lesssim 12$ . In addition, we find that AGN fraction increases with redshift, rising from 1% to 10% at  $z = 0.33$ .

We then derived AGN fractions from sBHAR distributions produced by other studies and compared them to our results. Our low redshift results help confirm the previously identified, strong redshift evolution of the AGN fraction in galaxies with  $M_* \gtrsim 10^{10} M_\odot$  brought about by an increase in the availability of cold gas. In contrast, our results confirm a lack of evolution at low stellar masses, indicating a constant AGN fraction out to  $z = 1$  (see also [Birchall et al. 2020](#)) The suppression in this region is thought to be driven by the shallower potential wells in lower mass galaxies that are unable to capture as much gas as their higher mass counterparts.

In conclusion, we have shown that AGN with a broad range of accretion rates are found across the local galaxy population. This work shows that lower mass AGN are more likely to be missed by optical selection but that careful analysis of X-ray observations will help reveal them. Our measurements quantify the AGN fraction in the local Universe and provide crucial limits on their evolution over cosmic time.

**ACKNOWLEDGEMENTS**

KB acknowledges funding from a STFC PhD studentship. JA acknowledges support from an STFC Ernest Rutherford Fellowship (grant code: ST/P004172/1) and a UKRI Future Leaders Fellowship (grant code: MR/T020989/1).

This research has made use of data obtained from the 3XMM XMM-Newton serendipitous source catalogue compiled by the 10 institutes of the XMM-Newton Survey Science Centre selected by ESA. In addition, this research made use of Astropy,<sup>4</sup> a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013, 2018). Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

KB would also like to thank F-X Pineau for his help with constructing the `xmatch` scripts, and Duncan Law-Green for the help accessing the Flix archive.

**DATA AVAILABILITY STATEMENT**

The data underlying this article has been derived from publicly available datasets: the MPA-JHU Catalogue (based on SDSS DR8) & 3XMM DR7. Section 2 outlines where these catalogues are available and how the final sample was derived. The underlying data will also be shared on request to the corresponding author.

**REFERENCES**

- Agostino C. J., Salim S., 2019, *ApJ*, **876**, 12
- Aird J., et al., 2012, *ApJ*, **746**, 90
- Aird J., et al., 2013, *ApJ*, **775**, 41
- 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
- Alexander D. M., Hickox R. C., 2012, *New Astron. Rev.*, **56**, 93
- Astropy Collaboration et al., 2013, *A&A*, **558**, A33
- Astropy Collaboration et al., 2018, *AJ*, **156**, 123
- Baldassare V. F., Reines A. E., Gallo E., Greene J. E., 2017, *ApJ*, **836**, 20
- Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, *ApJ*, **600**, 681
- Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, **93**, 5
- Bauer A. E., et al., 2013, *MNRAS*, **434**, 209
- Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, *MNRAS*, **488**, 3143
- Bellovary J., Brooks A., Volonteri M., Governato F., Quinn T., Wadsley J., 2013, *ApJ*, **779**, 136
- Birchall K. L., Watson M. G., Aird J., 2020, *MNRAS*, **492**, 2268
- Blanton M. R., Moustakas J., 2009, *ARA&A*, **47**, 159
- Bongiorno A., et al., 2012, *MNRAS*, **427**, 3103
- Bongiorno A., et al., 2016, *A&A*, **588**, A78
- Brandt W. N., Alexander D. M., 2015, *Astronomy and Astrophysics Review*, **23**
- Cann J. M., Satyapal S., Abel N. P., Blecha L., Mushotzky R. F., Reynolds C. S., Secrest N. J., 2019, *ApJ*, **870**, L2
- Carrera F. J., et al., 2007, *A&A*, **469**, 27
- Charlot S., Fall S. M., 2000, *ApJ*, **539**, 718
- Coil A. L., et al., 2011, *ApJ*, **741**, 8
- Cool R. J., et al., 2013, *ApJ*, **767**, 118
- Delvecchio I., et al., 2014, *MNRAS*, **439**, 2736
- Delvecchio I., et al., 2018, *MNRAS*, **481**, 4971
- Delvecchio I., et al., 2020, *ApJ*, **892**, 17
- Di Matteo T., Springel V., Hernquist L., 2005, *Nature*, **433**, 604
- Fabian A. C., 2012, *ARA&A*, **50**, 455
- Ferrarese L., Merritt D., 2000, *ApJ*, **539**, L9
- Gebhardt K., et al., 2000, *ApJ*, **539**, L13
- Gehrels N., 1986, *ApJ*, **303**, 336
- Georgakakis A., Nandra K., 2011, *MNRAS*, **414**, 992
- Georgakakis A., et al., 2011, *MNRAS*, **418**, 2590
- Georgakakis A., Aird J., Schulze A., Dwelly T., Salvato M., Nandra K., Merloni A., Schneider D. P., 2017, *MNRAS*, **471**, 1976
- Greene J. E., Setton D., Bezanson R., Suess K. A., Kriek M., Spilker J. S., Goulding A. D., Feldmann R., 2020, *ApJ*, **899**, L9
- Haggard D., Green P. J., Anderson S. F., Constantin A., Aldcroft T. L., Kim D.-W., Barkhouse W. A., 2010, *ApJ*, **723**, 1447
- Hernán-Caballero A., et al., 2014, *MNRAS*, **443**, 3538
- 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
- Jahnke K., et al., 2009, *ApJ*, **706**, L215
- Jones M. L., Hickox R. C., Mutch S. J., Croton D. J., Ptak A. F., DiPompeo M. A., 2017, *ApJ*, **843**, 125
- Kauffmann G., et al., 2003a, *MNRAS*, **346**, 1055
- Kauffmann G., et al., 2003b, *MNRAS*, **346**, 1055
- Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, *ApJ*, **556**, 121
- Kormendy J., Ho L. C., 2013, *Annual Review of Astronomy and Astrophysics*, **51**, 511
- Lehmer B. D., et al., 2016, *ApJ*, **825**, 7
- Lemons S. M., Reines A. E., Plotkin R. M., Gallo E., Greene J. E., 2015, *ApJ*, **805**
- Lusso E., Risaliti G., 2016, *ApJ*, **819**, 154
- Madau P., Dickinson M., 2014, *ARA&A*, **52**, 415
- Magorrian J., et al., 1998, *AJ*, **115**, 2285
- Martin D. C., et al., 2005, *ApJ*, **619**, L1
- Martin D. C., et al., 2007, *ApJS*, **173**, 342
- Mendez A. J., et al., 2013, *ApJ*, **770**, 40
- Mezcua M., Civano F., Marchesi S., Suh H., Fabbiano G., Volonteri M., 2018, *Monthly Notices of the Royal Astronomical Society*, **478**, 2576
- Mineo S., Gilfanov M., Sunyaev R., 2012a, *MNRAS*, **419**, 2095
- Mineo S., Gilfanov M., Sunyaev R., 2012b, *MNRAS*, **426**, 1870
- Moustakas J., et al., 2013, *ApJ*, **767**, 50
- Mullaney J. R., et al., 2012, *ApJ*, **753**, L30
- Muzzin A., et al., 2013, *ApJ*, **777**, 18
- Paggi A., Fabbiano G., Civano F., Pellegrini S., Elvis M., Kim D.-W., 2016, *ApJ*, **823**
- Pardo K., et al., 2016, *ApJ*, **831**
- Pineau F. X., et al., 2017, *A&A*, **597**, A89
- Popping G., Caputi K. I., Somerville R. S., Trager S. C., 2012, *MNRAS*, **425**, 2386
- Rosen S. R., et al., 2016, *A&A*, **590**, A1
- Salim S., et al., 2007, *ApJS*, **173**, 267
- Salim S., et al., 2016, *ApJS*, **227**, 2
- Schawinski K., et al., 2014, *MNRAS*, **440**, 889
- Vanden Berk D. E., et al., 2001, *AJ*, **122**, 549
- Vito F., et al., 2014, *MNRAS*, **441**, 1059
- Wang T., et al., 2017, *A&A*, **601**, A63
- Weigel A. K., Schawinski K., Caplar N., Wong O. I., Treister E., Trakhtenbrot B., 2017, *ApJ*, **845**, 134
- Williams W. L., Röttgering H. J. A., 2015, *MNRAS*, **450**, 1538
- Xue Y. Q., et al., 2010, *ApJ*, **720**, 368
- Yang G., et al., 2017, *ApJ*, **842**, 72
- Yang G., et al., 2018, *MNRAS*, **475**, 1887

<sup>4</sup> <http://www.astropy.org>

## APPENDIX A: PROBABILITY DISTRIBUTION FIT COEFFICIENTS

Table A1 outlines the best-fit coefficients, and associated errors, for equation (6) to create every probability distribution shown in figure 7.

z	sBHAR ( $\log_{10} x' = -2.55$ )		X-ray Luminosity (erg/s) ( $\log_{10} x' = 42.1$ )	
	$\log_{10} A$	$k$	$\log_{10} A$	$k$
0.00 - 0.05	$-2.44 \pm 0.11$	$-0.57 \pm 0.13$	$-2.73 \pm 0.13$	$-0.43 \pm 0.12$
0.05 - 0.10	$-2.48 \pm 0.08$	$-0.72 \pm 0.11$	$-2.65 \pm 0.10$	$-0.97 \pm 0.12$
0.10 - 0.15	$-2.54 \pm 0.12$	$-1.39 \pm 0.14$	$-2.34 \pm 0.10$	$-1.78 \pm 0.19$
0.15 - 0.20	$-2.36 \pm 0.13$	$-1.23 \pm 0.18$	$-1.72 \pm 0.10$	$-1.34 \pm 0.20$
0.20 - 0.25	$-2.13 \pm 0.17$	$-1.15 \pm 0.23$	$-1.24 \pm 0.15$	$-1.64 \pm 0.32$
0.25 - 0.35	$-2.08 \pm 0.19$	$-1.59 \pm 0.24$	$-0.64 \pm 0.20$	$-2.01 \pm 0.39$
$\log_{10}$ Stellar Mass ( $M_{\odot}$ )	$\log_{10} A$	$k$	$\log_{10} A$	$k$
8.00 - 9.00	$-2.24 \pm 0.61$	$-0.90 \pm 0.01$	$-3.05 \pm 0.00$	$-0.37 \pm 0.43$
9.00 - 10.00	$-2.61 \pm 0.22$	$-0.90 \pm 0.26$	$-3.24 \pm 0.43$	$-0.74 \pm 0.25$
10.00 - 10.50	$-2.49 \pm 0.11$	$-0.49 \pm 0.16$	$-2.65 \pm 0.14$	$-0.50 \pm 0.16$
10.50 - 11.00	$-2.23 \pm 0.06$	$-0.52 \pm 0.08$	$-2.16 \pm 0.05$	$-0.48 \pm 0.08$
11.00 - 11.50	$-2.63 \pm 0.12$	$-1.22 \pm 0.13$	$-1.91 \pm 0.07$	$-1.07 \pm 0.12$
11.50 - 12.00	$-2.39 \pm 0.25$	$-1.11 \pm 0.26$	$-1.31 \pm 0.14$	$-1.12 \pm 0.27$

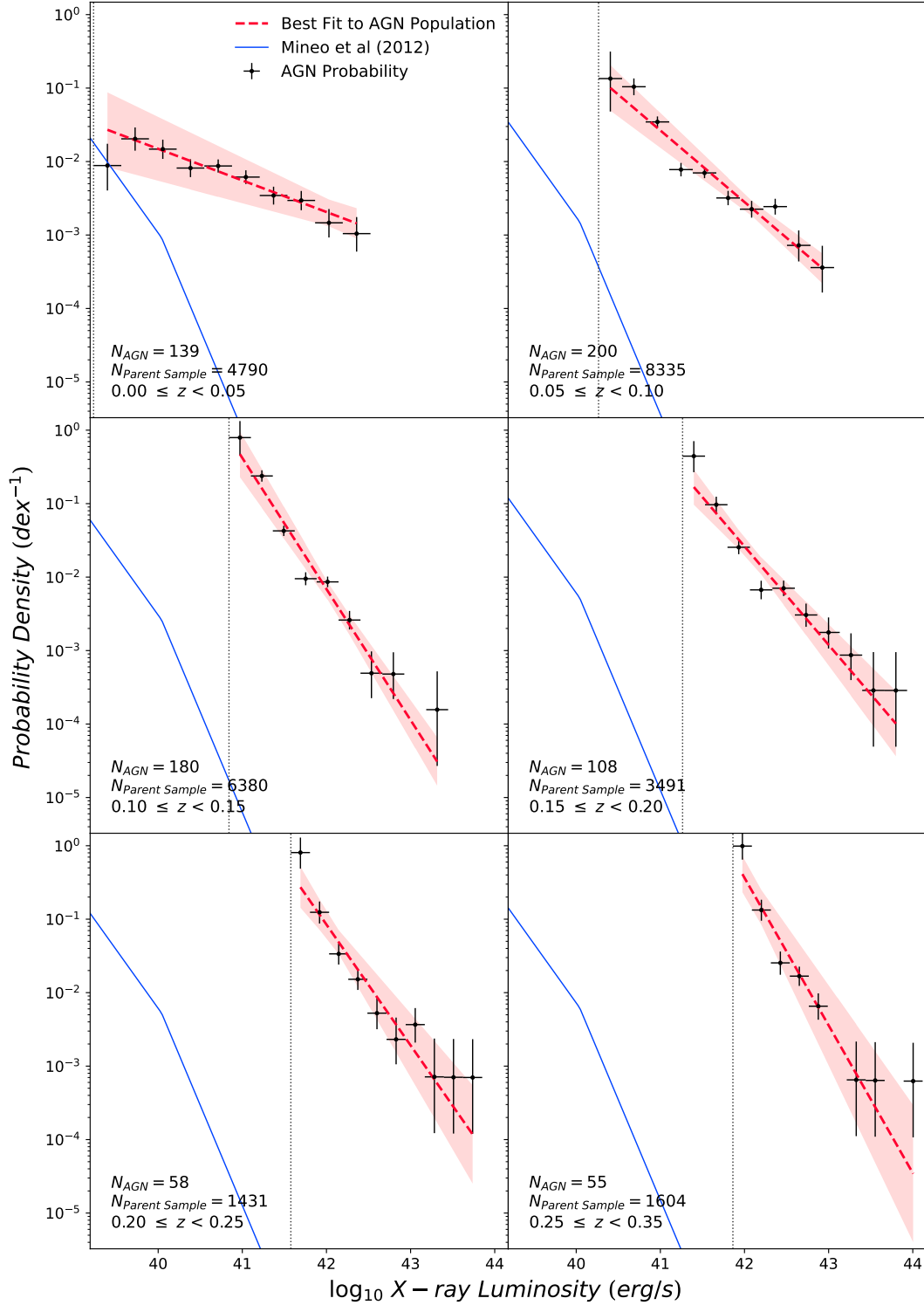
**Table A1:** Best-fit coefficients used to fit equation (6) to all probability distribution configurations.

## APPENDIX B: OTHER PROBABILITY DISTRIBUTIONS

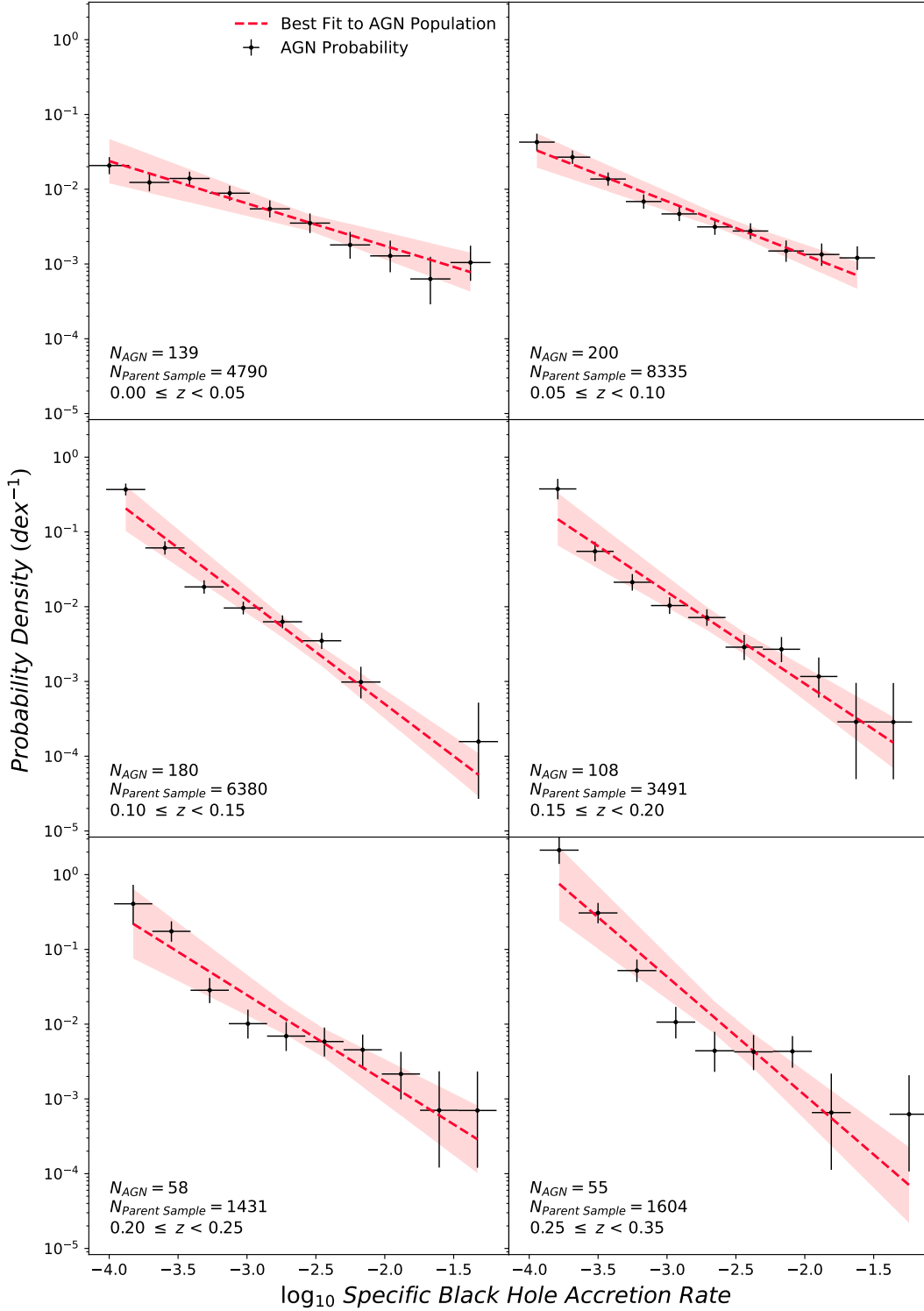
Figures B1 - B1 show the probability distributions split by stellar mass and redshift, and the remaining trends with X-ray luminosity and sBHAR. They have the same form as figure 6 and are included for the sake of transparency, to show the strength of our power law fits to the data.

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.

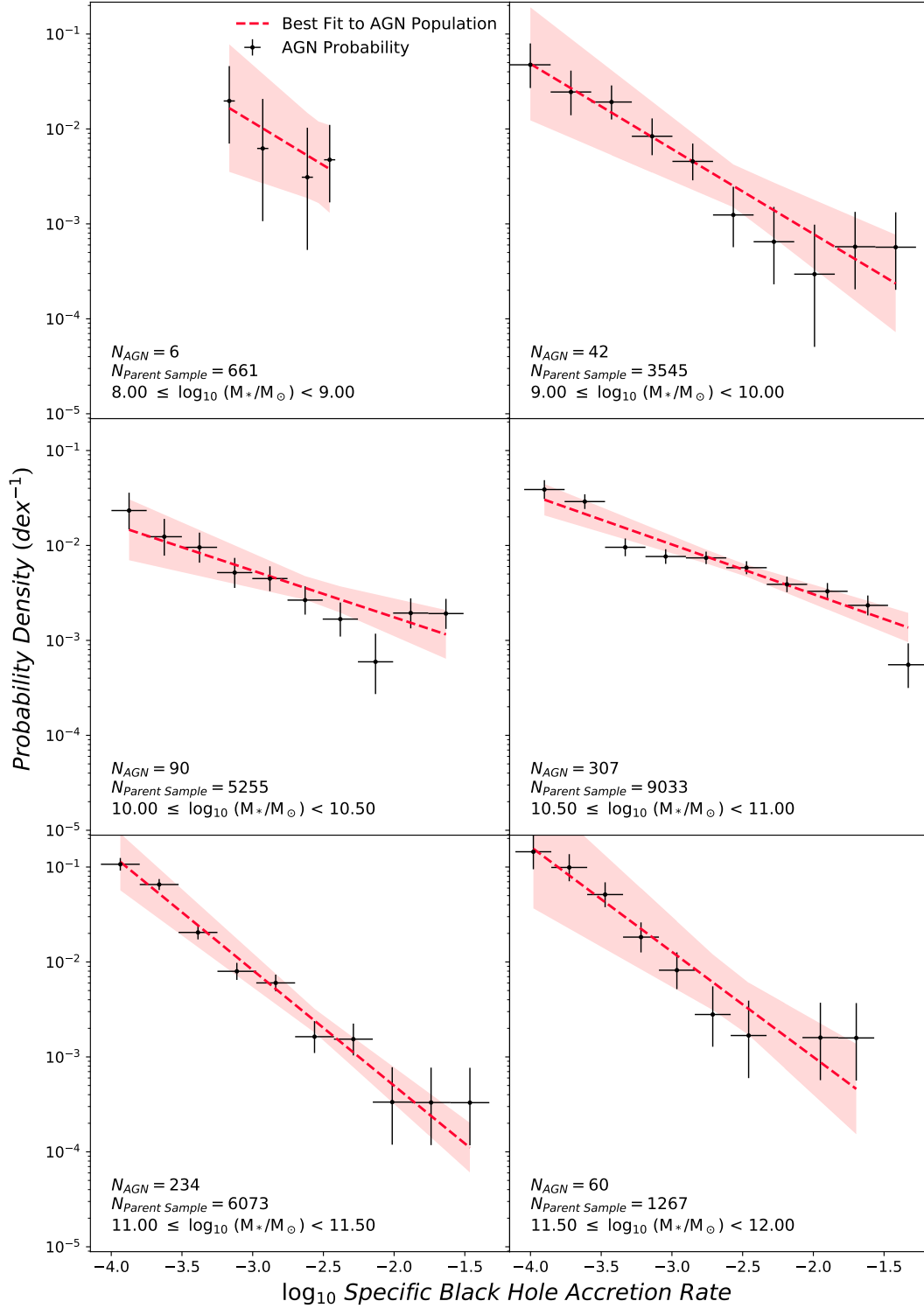




**Figure B1.** Additional probability distribution used to calculate fits and associated error regions for the other host galaxy properties. This figure shows the probability of a galaxy hosting an AGN as a function of X-ray luminosity and split into bins of redshift. As with figure 6, power laws (dashed red lines) have been fit to the data in each panel and displayed alongside their  $1\sigma$  uncertainty (pale red region). These plots were constructed using the method outlined in section 5



**Figure B1.** Additional probability distribution used to calculate fits and associated error regions for the other host galaxy properties. This figure shows the probability of a galaxy hosting an AGN as a function of  $\lambda_{\text{BHAR}}$  and split into bins of redshift. As with figure 6, power laws (dashed red lines) have been fit to the data in each panel and displayed alongside their  $1\sigma$  uncertainty (pale red region). These plots were constructed using the method outlined in section 5



**Figure B1.** Additional probability distribution used to calculate fits and associated error regions for the other host galaxy properties. This figure shows the probability of a galaxy hosting an AGN as a function of  $\lambda_{\text{sBHAR}}$  and split into bins of stellar mass. As with figure 6, power laws (dashed red lines) have been fit to the data in each panel and displayed alongside their  $1\sigma$  uncertainty (pale red region). These plots were constructed using the method outlined in section 5