

# A novel Bayesian approach for decomposing the radio emission of quasars – II. Link between quasar radio emission and black hole mass

B.-H. Yue (岳博涵)<sup>1,2\*</sup>, K. J. Duncan<sup>1</sup>, P. N. Best,<sup>1</sup> M. I. Arnaudova<sup>3</sup>, L. K. Morabito<sup>4,5</sup>, J. W. Petley<sup>2,4</sup>, H. J. A. Röttgering,<sup>2</sup> S. Shenoy<sup>3</sup> and D. J. B. Smith<sup>3</sup>

<sup>1</sup>*Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK*

<sup>2</sup>*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*

<sup>3</sup>*Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK*

<sup>4</sup>*Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

<sup>5</sup>*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

Accepted 2025 January 12. Received 2025 January 10; in original form 2024 November 5

## ABSTRACT

Whether the mass of supermassive black hole ( $M_{\text{BH}}$ ) is directly linked to the quasar radio luminosity remains a long-debated issue, and understanding the role of  $M_{\text{BH}}$  in the evolution of quasars is pivotal to unveiling the mechanism of active galactic nucleus (AGN) feedback. In this work, based on a two-component Bayesian model, we examine how  $M_{\text{BH}}$  affects the radio emission from quasars, separating the contributions from host galaxy star formation (SF) and AGN activity. By modelling the radio flux density distribution of Sloan Digital Sky Survey quasars from the LOFAR Two-metre Sky Survey Data Release 2, we find no correlation between  $M_{\text{BH}}$  and star formation rate (SFR) at any mass for quasars at a given redshift and bolometric luminosity. The same holds for AGN activity across most  $M_{\text{BH}}$  values; however, quasars with the top 20 per cent most massive supermassive black holes (SMBHs) are two to three times more likely to host strong radio jets than those with lower mass SMBHs at similar redshift and luminosity. We suggest defining radio quasar populations by their AGN and SF contributions instead of radio loudness; our new definition unifies previously divergent observational results on the role of  $M_{\text{BH}}$  in quasar radio emissions. We further demonstrate that this radio enhancement in quasars with the 20 per cent most massive SMBHs affects only the  $\sim 5$  per cent most radio bright quasars at a given redshift and bolometric luminosity. We discuss possible physical origins of this radio excess in the most massive and radio-bright quasar population, which remains an interest for future study.

**Key words:** black hole physics – galaxies: active – quasars: general – quasars: supermassive black holes – radio continuum: galaxies.

## 1 INTRODUCTION

Quasars (QSOs) emit powerful emission across the entire electromagnetic spectrum, ranging from X-rays to radio bands. Yet different components of the quasar dominate at different wavelengths. Historically, since the compilation of some of the earliest optically selected quasar catalogues, it was observed that only a small percentage (about 5 to 10 per cent) of the optically selected quasars are also luminous in radio bands when compared to the active galactic nucleus (AGN) emission in optical bands (see e.g. Kellermann et al. 1989), which are then referred to as radio-loud (RL). The remaining optical quasars that are relatively faint in the radio bands are classified as radio-quiet (RQ). Early studies considered the RL and RQ quasars as two distinct populations with different physical origins (e.g. Ivezić et al. 2002; White et al. 2007). However, more recent studies have disputed this point of view, arguing that radio emissions in RL and RQ quasars share a similar origin, and RL quasars are simply the long tail toward

the radio-bright end in a continuous quasar radio power distribution (e.g. Cirasuolo et al. 2003; Baloković et al. 2012; Macfarlane et al. 2021).

Efforts to tackle the questions behind a possible quasar radio bimodality mainly focused on investigating the possible differences in the physical properties of the central engines to quasar radio emission: supermassive black holes (SMBHs). Previous studies have used the AGN bolometric luminosity ( $L_{\text{bol}}$ ) as an indicator of black hole (BH) growth and therefore linked the quasar radio loudness to  $L_{\text{bol}}$  (e.g. Blundell & Rawlings 2001; Ho & Ulvestad 2001; Ulvestad & Ho 2001). However, any observed correlation between quasar radio emission and  $L_{\text{bol}}$  suffers from the degeneracy between the BH's Eddington-scaled accretion rate and BH mass, since both massive and highly accreting BHs can be optically bright. Introducing BH mass into the correlation can help break this degeneracy and uncover the leading factor that contributes to the quasar radio power, be it the BH accretion rate, BH mass, BH spin (Wilson & Colbert 1995), or other related evolutionary effects.

Meanwhile, understanding the role of the BH mass in the production of quasar radio emission is also a critical step in resolving the

\* E-mail: [bohan.yue@ed.ac.uk](mailto:bohan.yue@ed.ac.uk)

evolutionary path of radio quasars. BH masses are found to have a tight correlation with the stellar masses of their host galaxies' bulge (see Kormendy & Ho 2013, and the reference therein), which is now considered to be critical evidence of the co-evolution between the galaxy and the BH through feedback processes (see e.g. Fabian 2012; Heckman & Best 2014). A possible connection (or not) between the BH mass and the quasar radio jet will help us determine the role of AGN jets in the feedback process, which can then serve as a building block in the galaxy evolution model.

Thanks to a series of optical survey campaigns, especially the Sloan Digital Sky Survey (SDSS; York et al. 2000; Schneider et al. 2010; Ross et al. 2012), previous studies have established the robustness of virial BH mass estimators (e.g. Kaspi et al. 2000; Vestergaard & Peterson 2006), and sensible estimations of BH masses can now be derived using a set of broad-line full width half-maxima (FWHM) and the continuum luminosity from single-epoch optical spectra (Shen et al. 2011, 2019). However, despite the expanding sample set and improved observing methods, previous studies have failed to reach an agreement on whether the BH mass is correlated with the quasar radio loudness. Studies including McLure & Jarvis (2004), Seymour et al. (2007), and Whittam et al. (2022) reported an increase in the number fraction of RL AGNs above the BH mass threshold of  $10^8 - 10^9 M_{\odot}$ , suggesting that the RL AGNs tend to host more massive BHs than RQ AGNs, using radio observations from the NRAO VLA Sky Survey (NVSS; Condon et al. 1998), the FIRST survey (Faint Images of Radio Sky at Twenty Centimeters; Becker, White & Helfand 1995), and the MeerKAT International GHz Tiered Extragalactic Exploration (Jarvis et al. 2017), respectively.

On the other hand, using recent low-frequency radio data from LOw-Frequency ARray (LOFAR; van Haarlem et al. 2013), studies including those by Gürkan et al. (2019), Macfarlane et al. (2021), and Arnaudova et al. (2024) failed to find any correlation between BH mass and radio-loudness or differences in average BH masses between the RL and RQ population. Although using different approaches, these studies came to the same conclusion while sharing a similar data set (SDSS-detected quasars observed by LOFAR). Gürkan et al. (2019) compared the radio-loudness- $M_{\text{BH}}$  distribution for LOFAR-detected quasars and the LOFAR non-detections, and found no significant correlation between  $M_{\text{BH}}$  and radio loudness. Macfarlane et al. (2021) separated the quasars into hosting high-mass and low-mass BHs based on whether they were above or below the mean  $M_{\text{BH}}$  of the entire sample set, before fitting their radio flux density distribution to the two-component model that separates the AGN and host galaxy contribution in the quasar radio emission, finding that the AGN component shows little difference between their high-mass and low-mass samples. Arnaudova et al. (2024) adopted a selection criterion of  $R = \log_{10}(L_{1.4\text{GHz}}/L_i) > 1$  for RL quasars (where  $L_{1.4\text{GHz}}$  and  $L_i$  stand for the 1.4 GHz radio luminosity and  $i$ -band optical luminosity, respectively), and found no systematic difference in the distribution of BH mass estimates between RL and RQ quasars. By performing a null hypothesis test on the stacked spectra of quasars in  $z - \mathcal{M}_i - M_{\text{BH}}$  space (where  $\mathcal{M}_i$  is the absolute  $i$ -band magnitude), they also concluded that the BH mass cannot fully explain the difference between the stacked spectra of RL and RQ quasars. Therefore, they argued that BH mass is not the driving force behind quasar radio emission.

This disagreement with previous studies is, at least in part, caused by a lack of consensus on the various components that drive the quasar radio emission; as a result, different works adopted different definitions of radio loudness based on observational results in different frequency bands. Therefore, it is hard to integrate these

different observational data sets into one physical picture. To resolve this issue, based on the two-component model from Macfarlane et al. (2021), we proposed a new Bayesian model in Yue et al. (2024, hereafter Y24) that separates the host galaxy star formation (SF) and the contribution of AGN activity to quasar radio emission using the observed radio flux density distribution of quasars. By fitting the model against the LOFAR-observed radio flux density distribution of SDSS quasars, we were able to give a robust estimation of the mean star formation rate (SFR) of the quasar host galaxy and determine the normalization of a power-law jet luminosity function for any subpopulation of quasars with a given absolute  $i$ -band magnitude ( $\mathcal{M}_i$ ) and redshift ( $z$ ). We were able to explore how the evolution of the host galaxy SFR and AGN activity vary with  $L_{\text{bol}}$  and  $z$ , and also, by binning the quasars by their physical properties, to study the connection between different components of quasar radio emission and quasar colour. This approach can then be expanded to study the individual evolutionary patterns of host galaxy SFR and AGN activity with BH mass and Eddington ratio, under a physically motivated model. Therefore, we aim to answer these questions:

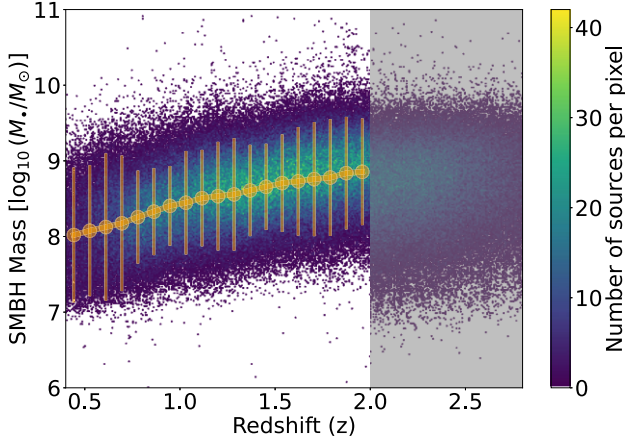
- (i) Is BH mass associated with the radio loudness of quasars?
- (ii) Do the current definitions of quasar radio loudness correctly reflect the underlying physical processes?
- (iii) Does the BH mass/accretion mode govern the AGN jet production mechanism?

This paper is structured as follows: Section 2 describes the data we used to build an optically selected quasar sample observed with LOFAR. Section 3 briefly summarizes the parametric model we proposed to characterize quasar radio emission and presents our result on the evolution of quasar radio emission with different BH masses. Section 4 compares our result with those in the previous literature, which leads to a discussion on the origin of radio excess in more massive quasars. Sections 5 and 6 discuss various scenarios that could explain our results. Finally, a summary of our conclusions can be found in Section 7. Throughout this work, we assume a Lambda cold dark matter cosmology with parameter values published in the WMAP9 result (Hinshaw et al. 2013).

## 2 DATA

### 2.1 Sloan Digital Sky Survey quasar sample

Our parent quasar sample is the SDSS Data Release 16 quasar (DR16Q) catalogue, with the photometric data presented in Lyke et al. (2020) and the associated spectroscopic data presented in Wu & Shen (2022). The DR16Q catalogue is constructed using observations from SDSS-I/II/III and IV epochs, processed with the final eBOSS SDSS reduction pipeline (v5.13.0), and validated spectroscopically following the criteria outlined in table 1 of Pâris et al. (2018). It incorporates previously identified sources from the DR14Q, DR12Q, and DR7Q catalogues (Schneider et al. 2010; Ross et al. 2012; Pâris et al. 2018). For quasars with redshifts ranging from  $0.8 < z < 2.2$ , the identification of quasars is carried out using the decision tree detailed in Dawson et al. (2016); for quasars with redshifts  $z > 2.2$ , the identification is based on the Ly $\alpha$  forest measurements as described in Myers et al. (2015). There are no inconsistencies observed in the characteristics (e.g. redshift and bolometric luminosity) of quasars identified through different methodologies. The total count of detected quasars exceeds 480 000 and 239 000 for the low- and high-redshift categories, respectively. Additional optical characteristics of quasars from the SDSS data set,



**Figure 1.** BH mass-redshift distribution of the sample used in this work. Solid dots mark the average measured BH mass within each redshift bin, while the error bars show the standard deviation of BH masses within the corresponding redshift bin. Average measured BH masses lie between  $M_{\text{BH}} = 10^8 \sim 10^9 M_{\odot}$ , and maintain a good range throughout the redshift bins. The shaded area ( $z > 2$ ) marks the parameter space where only C IV BH masses are available; these values show discrepancy with BH masses measured via Mg II and H  $\beta$ , possibly due to AGN outflows that introduced bias to the emission line FWHM measurement. Therefore sources with  $z > 2$  are excluded from our analysis.

such as absolute  $i$ -band magnitude,<sup>1</sup> are enriched by survey data from, for example, FIRST and the *Wide-field Infrared Survey Explorer* (WISE; Wright et al. 2010) through cross-referencing between data sets; the details of the cross-matching process are described in Wu & Shen (2022).

### 2.1.1 Black hole mass

We adopt the BH mass measurements from the values presented in the spectroscopic SDSS DR16Q catalogue (Wu & Shen 2022). These masses are obtained by fitting the SDSS quasar spectra to a global continuum + emission line model, using the public PYQSOFIT code (Shen et al. 2019). Three BH mass indicators are presented in the catalogue: H  $\beta$ , Mg II, and C IV. The BH masses are estimated from single-epoch observations of these mass indicator emission lines, using the virial relation provided in Schneider et al. (2010):

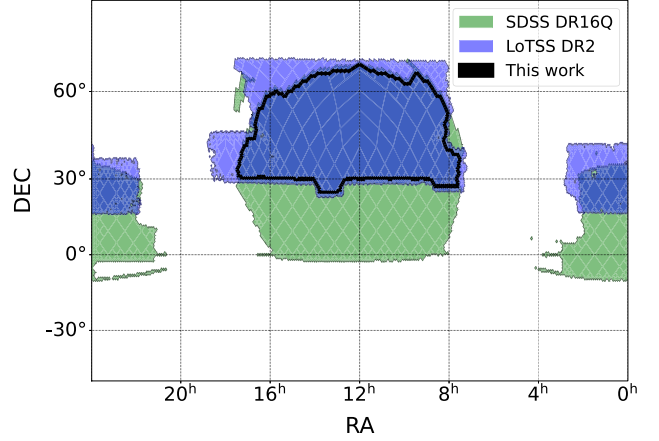
$$\log \left( \frac{M_{\text{BH}}}{M_{\odot}} \right) = a + b \log \left( \frac{\lambda L_{\lambda}}{10^{44} \text{erg s}^{-1}} \right) + 2 \log \left( \frac{\text{FWHM}}{\text{km s}^{-1}} \right), \quad (1)$$

where  $(a, b) = (0.910, 0.50)$  for H  $\beta$ ,  $(0.740, 0.62)$  for Mg II, and  $(0.660, 0.53)$  for C IV.

We present our sample BH mass distribution with redshift up to  $z = 2.8$  in Fig. 1. The H  $\beta$  and Mg II measurements show good consistency in the mass-redshift space, with average values (shown in light yellow dots) lying between  $M_{\text{BH}} = 10^8 \sim 10^9 M_{\odot}$ . The error bars indicate the standard deviation of BH masses within the corresponding redshift bin, from which we can conclude that the BH mass distribution maintains a good dynamical range and shows no signs of selection effects across the entire redshift range.

On the other hand, for sources with redshift beyond  $z = 2$  where only C IV measurements are available, the emission line FWHM measurements appear to be discrepant, leading to a jump in the measured

<sup>1</sup>The  $i$ -band magnitudes in the DR16Q catalogue were  $K$ -corrected to  $z = 2$ , while in our data set, they are  $K$ -corrected to  $z = 0$  using the original data and assuming a spectral index of 0.5 (Richards et al. 2006).



**Figure 2.** A comparison of the sky coverages from LoTSS DR2 (blue) and SDSS DR16Q (green) survey catalogues. The black line encompasses the sky area studied in this paper, which is the overlap between the LoTSS DR2 and SDSS DR16Q survey areas in the north galactic cap.

$M_{\text{BH}}$  values with redshift at  $z = 2$ , as shown by the shaded area in Fig. 1. Previous works, including Coatman et al. (2017), have attempted to investigate this discrepancy and provide ways to correct the bias in these measurements. In particular, Coatman et al. (2017) argued that such discrepancy may be caused by the emission line displacement due to C IV blueshift, which is, in turn, connected to AGN outflow activities. However, none of the correction methods appear to be conclusive, and neither do they resolve the tension between C IV and Mg II  $M_{\text{BH}}$  measurements. In addition, other studies have shown a strong correlation between AGN outflow and enhanced radio emission from AGN (e.g. Morabito et al. 2019; Petley et al. 2022), suggesting that biased mass estimations from C IV emission line are associated with quasar radio excess at  $z > 2$ . As a result, to avoid any biases, we exclude the sources with  $z > 2$  and restrict our samples to lower redshift ( $z \leq 2$ ) where the Mg II and H  $\beta$  estimates are available.

## 2.2 Matching LoTSS data with SDSS quasars

### 2.2.1 LOFAR Two-metre Sky Survey

As in Y24, we obtain the radio measurements for our quasar samples via the LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al. 2017, 2019, 2022), which is a wide-field imaging survey conducted with the LOFAR HBA (high-band antenna) system. It aims to survey the entire northern sky in the 120–168-MHz radio band, with a desired rms noise level of approximately  $100 \mu\text{Jy beam}^{-1}$ , an angular resolution of 6 arcsec, and a positional accuracy better than 0.2 arcsec. The most comprehensive catalogue to date is the LoTSS DR2 catalogue presented by Shimwell et al. (2022), which covers an area of  $5720 \text{ deg}^2$  of the sky and achieves a median rms noise of  $83 \mu\text{Jy beam}^{-1}$ . The current LoTSS DR2 catalogue comprises approximately 4400 000 radio-detected sources. The sky coverage, in comparison with the SDSS DR16 quasar catalogue, is illustrated in Fig. 2. A significant portion of the sky area in LoTSS DR2 overlaps with the SDSS DR16 quasar catalogue, making it an excellent resource for multiwavelength studies.

In the LoTSS DR2 catalogue, radio sources were extracted from LoTSS images using PYBDSF, the Python Blob Detector and Source Finder (Mohan & Rafferty 2015). These sources were pinpointed based on the peak radio flux densities surpassing the  $5\sigma$  threshold in the LoTSS DR2 images. Although PYBDSF is proficient at detecting regions of radio emission, it does not always properly group them



to physical sources. To mitigate misassociations by PYBDSF, a combination of statistical methods and thorough visual examination, facilitated by the LOFAR Galaxy Zoo, was employed to validate that the radio catalogue reflects the genuine distribution of radio sources (Hardcastle et al. 2023). In addition to the revised radio catalogue, Hardcastle et al. (2023) also detail the process of cross-referencing the LoTSS DR2 sources with the optical-infrared counterparts of the WISE (Wright et al. 2010) and DESI Legacy Imaging (Dey et al. 2019) surveys, employing a methodology similar to that outlined in Williams et al. (2019) and Kondapally et al. (2021).

### 2.2.2 Radio flux densities from LoTSS

We obtain the radio flux density measurements for SDSS DR16Q quasars from the LoTSS DR2 catalogue and the original LoTSS mosaic. While the details of the process are described in Y24, we briefly summarize our approach as follows:

(i) For sources with radio flux densities that surpass the detection limit of  $5\sigma$  in LoTSS DR2, we cross-matched the positions of quasars in the SDSS catalogue with those in the source-associated LoTSS catalogue using a matching radius of 1.5 arcsec. Note that whenever feasible, we used the coordinates from the cross-referenced optical catalogues of Hardcastle et al. (2023) for LoTSS positions.

(ii) For sources with radio flux densities below the LoTSS  $5\sigma$  detection limit and thus not recorded in the LoTSS DR2 catalogue, we performed forced photometry on LoTSS mosaics using the method described in Gloudemans et al. (2021). This approach is motivated by the assumption that most quasars are unresolved in LoTSS DR2 images, while the astrometric uncertainty in LoTSS DR2 (0.2 arcsec) is significantly smaller than the LoTSS pixel scale (1.5 arcsec). The extracted radio flux density is considered to be the highest pixel value within a  $3 \times 3 \text{ px}^2$  aperture centred on the source position in the SDSS catalogue. The uncertainty in the flux density is calculated based on the standard deviation of pixel values in a  $100 \times 100 \text{ px}^2$  cut-out region surrounding the central pixel. As discussed in Y24, although these individual measurements are rather noisy, they retain valuable information about the distribution of radio flux densities of the population as a whole.

### 2.3 Building a LoTSS–SDSS quasar sample

We build our quasar sample set based on the method presented in Y24 by extracting LoTSS radio flux density measurements for the parent SDSS DR16Q quasar samples. As described in Y24, we have removed the following sources from the SDSS DR16Q sample set:

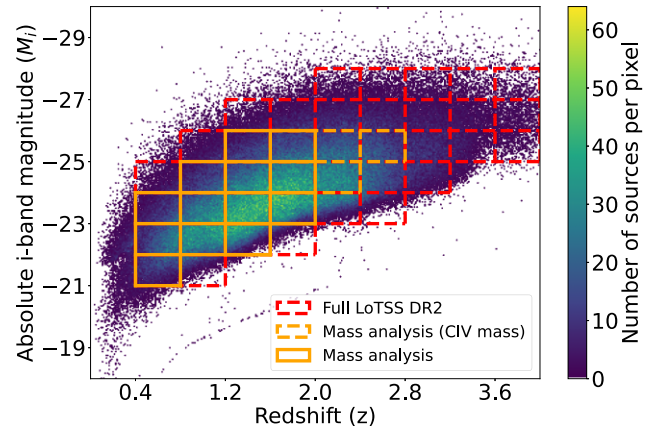
(i) Overly luminous sources (with absolute  $i$ -band magnitude brighter than  $\sim -40$ ) that are likely artefacts of the SDSS data reduction pipeline;

(ii) Sources in the 0h field (as illustrated in Fig. 2) that exhibit consistently higher uncertainties in radio flux density compared to those in the 13 h region, and quasars that are not covered by LoTSS being located either outside the target field or within the spaces between the LoTSS mosaics.

(iii) Sources in the SDSS DR16Q catalogue for which the only selection criteria for spectroscopic observation was the radio properties.

In addition to the restrictions above, we have applied the following additional constraints for the specific analysis in this paper:

(iv) Given that our fitting approach uses sources within each grid in the  $\mathcal{M}_i$ – $z$  plane (see Fig. 3), we focus only on cells within the



**Figure 3.** Coverage of parameter space in  $\mathcal{M}_i$ – $z$  plane. The red-dashed grids show the total LoTSS DR2–SDSS sample set included in Y24, while the solid orange grids show the samples used in this work: Each grid hosts at least 5000 quasars with estimated BH masses from either  $H\beta$  or  $Mg\ II$  measurements. The dashed orange grid cells each host more than 5000 quasar samples but have only C IV BH mass estimations. These grid cells are excluded from this work but may be reintroduced into the sample once high-quality spectroscopic data becomes available in future surveys including WEAVE.

$\mathcal{M}_i$ – $z$  grids that contain more than 5000 sources. This is because during the mass dependence analysis, each  $\mathcal{M}_i$ – $z$  grid needs to be binned into five sub-grids by the BH mass percentiles within each grid, and the minimum sample size required by our fitting approach is  $\sim 1000$  (see Y24).

(v) In Wu & Shen (2022), the final BH masses are adopted from estimations from different mass indicators based on redshift range ( $H\beta$  for  $z < 0.7$ ,  $Mg\ II$  for  $0.7 \leq z \leq 2.0$ ). We discard the sources where only C IV measurements are available due to the reasons listed in Section 2.1.1.

The final sample consists of 222 782 quasars. The samples are characterized by their distribution in the  $\mathcal{M}_i$ – $z$  plane, as shown in Fig. 3. While the sample is naturally biased towards brighter absolute magnitudes at higher redshifts, it still maintains a good sampling range across the entire parameter space.

## 3 THE BLACK HOLE MASS IMPACT ON QUASAR RADIO CONTINUUM EMISSIONS

### 3.1 Modelling the quasar radio flux density distribution

In Y24, we presented a fully Bayesian two-component model that disentangles the host galaxy and the AGN contribution in a given quasar radio flux density distribution. This model can then be adopted to analyse the dependence of the quasar host galaxy and AGN radio emission on BH mass independently. Here, we briefly summarize the structure of our Bayesian model.

In order to translate the observed quasar radio flux density distribution to a parametric statistical function, we followed the methodology developed in Roseboom & Best (2014) that allows us to stack the sources and analyse them as a whole. More specifically, our model calculates the probability  $P(s_j)$  of each input source  $j$  within a given  $\mathcal{M}_i$ – $z$  grid (as shown in Fig. 3) having an observed flux density  $s_j$ . Through Bayes’ theorem, the likelihood of the fitted

model  $\mu$  is defined as

$$P(\mu|\{s_j\}) \propto P(\mu)P(\{s_j\}|\mu) = P(\mu) \prod_j P(s_j|\mu). \quad (2)$$

where  $\{s_j\}$  is the distribution of the observed radio flux density within the  $\mathcal{M}_i-z$  grid.

We then derive the expression of our model probability density function (PDF)  $P(s_j|\mu)$  based on the two-component assumption presented in Macfarlane et al. (2021) and Y24, where every quasar is thought to host radio emission from both host galaxy SF and AGN activity. Assuming that quasars within a  $\mathcal{M}_i-z$  grid share similar physical properties, we model the SF component with a log-Gaussian PDF centred at the 150 MHz radio luminosity  $\log(L_\mu/[W \text{ Hz}^{-1}])$  that traces the typical SFR  $\Psi$  within the grid, and with a scatter of  $\sigma_\mu/\text{dex}$ :

$$P_{\text{SF}}(L)dL = \left[ \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left(-\frac{(\log L - \log L_\mu)^2}{2\sigma_\mu^2}\right) \right] \frac{dL}{L}. \quad (3)$$

The fitted values of  $L_\mu$  and  $\sigma_\mu$  can then be converted to SFR ( $\Psi$ ) and scatter in SFR ( $\sigma_\Psi$ ) using the mass-independent SFR–radio luminosity correlation in Smith et al. (2021). We model the AGN component by extrapolating the luminosity function of RL AGNs to lower luminosities, assuming that both RL and RQ AGNs share the same mechanism for generating radio emissions, only with different power efficiency. The AGN component is therefore characterized by a power-law PDF with a slope of  $\gamma$  and a normalization parameter  $\phi$ :

$$P_{\text{AGN}}(L)dL \propto \phi L^{-\gamma} dL, \quad (4)$$

where  $\phi$  can be translated into the RL fraction  $f$ , defined in Macfarlane et al. (2021) as the fraction of sources with 150 MHz luminosity of the AGN component above  $10^{26} \text{ W/Hz}$ . To maintain consistency with Macfarlane et al. (2021) and Y24, we characterize our best-fitting model with the set of parameters  $\{\Psi, \sigma_\Psi, \gamma, f\}$ .

The left and middle panels of Fig. 4 show the example PDF within one of the  $\mathcal{M}_i-z$  grids that we explored in this work, in log and linear space, respectively. The orange-dotted line shows the fitted underlying radio flux density distribution of host galaxy SF activity, while the green-dashed line shows the underlying distribution of AGN activity. Combining these two components, we have the total underlying radio flux density distribution for the quasar population (pink dash-dotted line). Finally, the PDF convolved with the radio flux density error is shown in the solid brown line. Our convolved PDF shows good agreement with the observed quasar radio flux densities (blue crosses) across all  $\mathcal{M}_i-z$  grids, indicating that our model provides a good description of the quasar sample.

### 3.2 The dependence of quasar radio emission on BH mass

With our model being able to separate SF and AGN features in a radio flux density distribution containing at least 1000 sources (Y24), we can take an  $\mathcal{M}_i-z$  grid cell and further bin it down to sub-samples by their measured BH masses. By comparing the fitting result across different sub-samples within one  $\mathcal{M}_i-z$  grid cell, we can investigate how the SF and AGN components in quasar radio emission vary separately with BH masses. With the average uncertainties of our BH masses being around 0.1 dex and the typical range of masses within a bin being between  $\log(M/M_\odot) = 7$  and 10, we bin the BH masses within a  $\mathcal{M}_i-z$  grid cell into five sub-sample sets, each containing sources with measured BH mass in quintiles 0–20, 20–40, 40–60, 60–80, and 80–100 per cent based on the mass distribution

in the  $\mathcal{M}_i-z$  grid cell.<sup>2</sup> The right panel in Fig. 4 takes the third and fifth quintile BH mass bins as an example to show how different best-fitting model parameters affect the final convolved PDFs and that the parameterization we use gives an accurate reflection of the observed radio flux density distributions within each sub-sample.

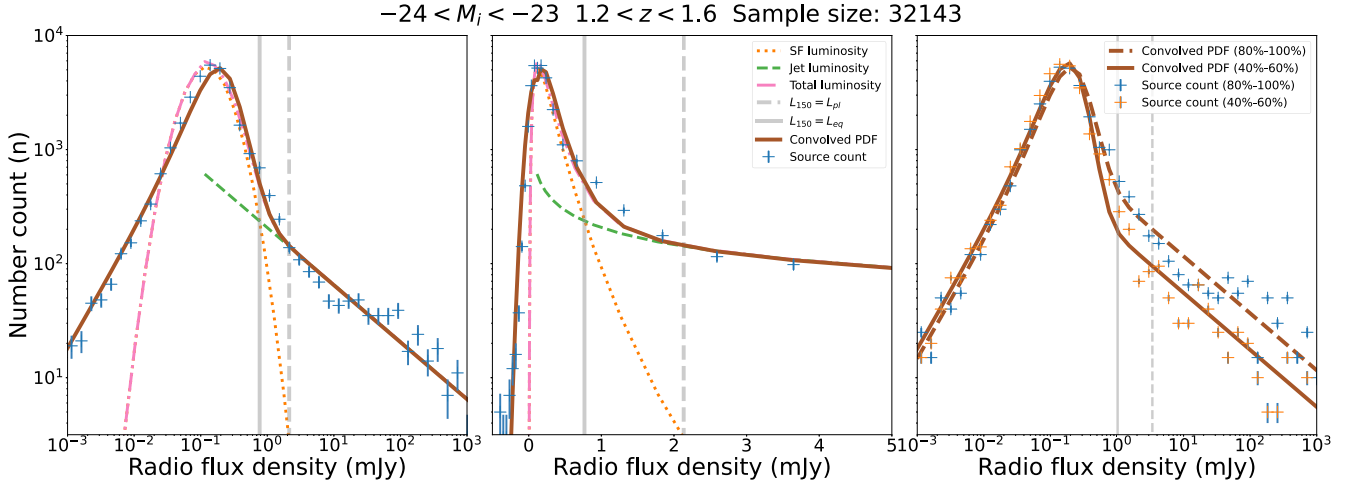
Going into more detail, Fig. 5 shows the continuous evolution of our fitted parameters with BH mass percentiles in all of the  $\mathcal{M}_i-z$  grid cells included in this work. Here, we focus on the parameters  $\Psi$  and  $f$ , as they are key indicators of the levels of host galaxy SF activity and AGN activity, respectively. The parameter  $\gamma$  is fixed at  $\gamma = 1.5$  based on the result from extrapolating the bright end of the distribution function as described in Y24, while  $\sigma_\mu$  remains a free parameter in our fit. The inferred  $\sigma_\mu$  values are typically between 0.2 and 0.3, with less than 5 per cent variation with BH mass; thus, they are excluded from the discussion here. The sub-samples are characterized by the deviation between the average BH mass within the sub-sample set and the average BH mass within the  $\mathcal{M}_i-z$  grid cell, denoted by  $\Delta \log M_{\text{BH}}$  in Fig. 5. Meanwhile, the evolution of quasar radio emission from host galaxy and AGN, i.e. our y-axis, is characterized by the deviation between the fitted value of the parameters in the sub-sample set and the fitted value in the full sample within that  $\mathcal{M}_i-z$  grid cell, denoted as  $\log_{10}(\Psi/\Psi_0)$  and  $\log_{10}(f/f_0)$  for the SF component and the AGN component, respectively.

As seen in Fig. 5, there is little sign of variation with the BH mass for the host galaxy SF activity across the entire sample (left panel) and for AGN activity with most of the BH mass bins (middle panel). However, quasars with the most massive BHs (top 20 per cent of the BH mass distribution) are two to three times more likely to be radio-bright when compared to the quasars of the same redshift and bolometric luminosity while hosting smaller BHs, due to an enhanced AGN activity (an  $\sim 0.4$  dex increase in the fitted value of  $\log(f)$ ). We have also split the variation of AGN activity into low-redshift ( $z < 1.2$ ) and high-redshift ( $z > 1.2$ ) regimes, and showed that the trend of the variation remains consistent in all redshifts (right panel). Note that while the modelling assumes a fixed value of  $\gamma$ , inspection of individual fits shows that this assumption appears reasonable (e.g. the right panel in Fig. 4). This means that the enhancement occurs across a broad range of AGN jet powers.

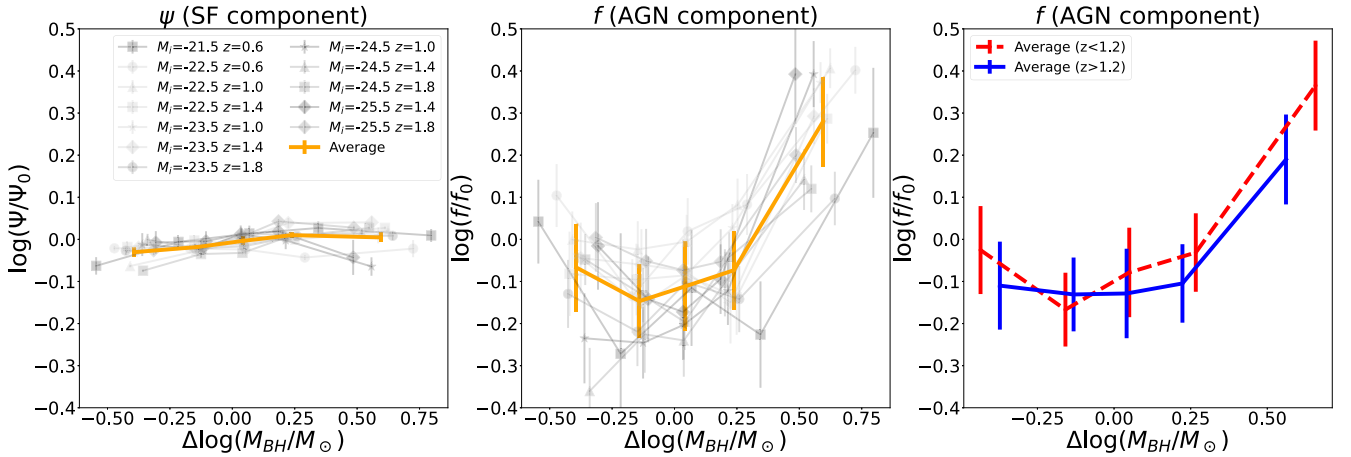
## 4 A COHERENT PICTURE OF QUASAR RADIO LOUDNESS AND BLACK HOLE MASS

In the above section, we have split the sample by BH mass and explored its effect on the quasar radio continuum emission. Although the relation between quasar radio luminosity and BH mass has been visited many times in the previous literature, most of those works split the quasars by their radio properties (e.g. radio loudness) and examined the resulting BH mass distribution. While the model best-fits in Fig. 5 outline the evolution of radio emission within continuous mass percentiles, in this section, we will divide the samples by

<sup>2</sup>We pick this binning method to balance between the ability to see a quantified trend and the possibility of cross-contamination between different bins due to uncertainties in BH mass measurement. The typical width of the BH mass bin is  $\sim 0.25$  dex; while this width is larger than the quoted statistical uncertainties of the BH mass measurements, several studies (e.g. McLure & Jarvis 2002; Shen 2013; Peterson 2014; Chaves-Montero et al. 2022; Shen et al. 2024) indicate that the typical scatter between BH masses obtained from single-epoch observations and reverberation campaigns is  $0.3 \sim 0.5$  dex, which is larger than (or similar to) our typical bin widths. There may therefore be some cross-contamination between the bins, which would weaken any observed trend.



**Figure 4.** Radio flux density distribution and the corresponding model best-fit in one of the representative  $\mathcal{M}_i-z$  bins explored in this study. The left and middle panels show the distribution and fitted relation for the entire quasar population within the grid, in log and linear scales, respectively. The orange-dotted line and green-dashed line represent the SF component from the host galaxy activity (log Gaussian) and the jet component from the AGN activity (single power law), respectively. The pink dashed-dotted line shows the combined PDF of the two-component radio flux density distribution model, and the brown solid line shows the PDF after convolving with the observational error. This convolution spreads the original PDF at the faint end out to negative flux density values, leading to an offset between the convolved and original PDFs at flux densities lower than 100  $\mu\text{Jy}$ . The distribution of measured radio flux densities is shown as blue crosses, indicating a good match between the actual distribution and the best-fitting model result. The vertical solid and dashed grey lines mark the criteria proposed in this work for identifying host galaxy- and AGN-dominated quasars, which will be discussed in Section 4. In the right panel, we present our fitting results for medium BH mass (with values in the 40–60 per cent of the mass distribution) and high BH mass bins (80–100 per cent) in solid and dashed lines, respectively. The measured radio flux density distributions of both bins are mapped against the best-fits as orange and blue crosses.



**Figure 5.** The evolution of SF and AGN contributions to the quasar radio flux density distribution with BH mass, as measured from the model best-fits using quasar samples split by BH mass percentiles. Within each  $\mathcal{M}_i-z$  grid cell, the BH masses are divided into five quintiles (0–20, 20–40, 40–60, 60–80, and 80–100 per cent) and are characterized by the difference between the average BH mass in each percentile bin and the average from the entire  $\mathcal{M}_i-z$  grid cell ( $\Delta \log M_{\text{BH}}$ ). The SF and AGN contributions are characterized by the relative variation of the model parameters ( $\Psi/\Psi_0$ ,  $f/f_0$ ), where  $\Psi$ ,  $f$  are the best-fitting parameter values within each mass quintile and  $\Psi_0$ ,  $f_0$  are the best-fitting value using quasars from the entire  $\mathcal{M}_i-z$  grid cell. The thin grey lines show the correlation in each  $\mathcal{M}_i-z$  bin, and the thick yellow line presents the correlation averaged over the entire  $\mathcal{M}_i-z$  parameter space. The SF contribution shows little variance with increasing BH mass (left panel). Similarly, the AGN contribution within most of the BH mass range shows little variance over most of the BH mass range, but the most massive quasars (top 20 per cent of the mass percentile) show enhanced radio emission from AGN activities (middle panel), which confirms the pattern seen in Fig. 4. The right panel shows the variation of AGN activity split into low- ( $z < 1.2$ ) and high- ( $z > 1.2$ ) redshift ranges, which confirms that this trend is uniformly seen across all redshifts.

their radio brightness instead and examine their respective BH mass distributions. We can therefore bring the detected radio enhancement into the context of the overall radio luminosity and radio brightness of quasars to form a coherent picture and unify the previous literature.

As discussed previously, studies including McLure & Jarvis (2004) and Whittam et al. (2022) are in favour of the hypothesis where BH mass is connected to radio loudness due to RL quasars in their sample

typically hosting more massive BHs ( $M_{\text{BH}} \gtrsim 10^8 M_{\odot}$ ). However, other works including Gürkan et al. (2019) and Arnaudova et al. (2024) have found no correlation between radio loudness and BH mass, nor any difference between the RL and RQ quasar population in their BH mass distribution, therefore arguing that the BH mass is not a primary driving factor behind the radio-loudness of quasars. The main difference between these works is the radio loudness criteria

that they used to separate the RL and RQ quasars. For studies such as McLure & Jarvis (2004) that use high-frequency radio data, the widely adopted definition is  $R_{\text{flux}} = f_{5\text{GHz}}/f_{4400\text{\AA}}$ , where  $R_{\text{flux}}$  is the radio loudness, and  $f_{5\text{GHz}}$  and  $f_{4400\text{\AA}}$  are the observed fluxes from AGN in radio and optical bands, respectively. The threshold for RL/RQ quasar separation is  $R_{\text{flux}} = 10$ , with  $R_{\text{flux}} > 10$  quasars classified as RL and the rest as RQ. More recent studies using LOFAR data (e.g. Gürkan et al. 2019; Arnaudova et al. 2024) used the definition based on radio luminosities:  $R_{\text{lum}} = \log_{10}(L_{1.4\text{GHz}}/L_i)$ , where  $L_{1.4\text{GHz}}$  is the 1.4 GHz radio luminosity converted from LOFAR 150 MHz radio measurements assuming a radio slope of  $\alpha = 0.7$ ,<sup>3</sup> and  $L_i$  is the luminosity observed by the SDSS  $i$ -band filter; the RL/RQ threshold is defined as  $R_{\text{lum}} = 1$  (e.g. Baloković et al. 2012; Arnaudova et al. 2024). Gürkan et al. (2019) performed a correlation analysis between  $R_{\text{lum}}$  and  $M_{\text{BH}}$  where no distinction was made between the RL and RQ population; however, they still characterized their quasar population with the peak values in the  $R_{\text{lum}}$  distribution ( $\log_{10}(L_{150\text{MHz}}/L_i) \sim 1.6$ ). Note that while both are commonly referred to as ‘ $R$ ’ in the literature and we sometimes use  $R$  when discussing radio loudness, in general, in this section, we distinguish the two definitions ( $R_{\text{flux}}$  and  $R_{\text{lum}}$ ) with different subscripts for clarity. Alternative definitions based on radio excess are used in works including Whittam et al. (2022), where RL quasars are defined as having radio emission above the  $2\sigma$  threshold of the infrared–radio correlation (IRRC). This threshold implies that radio emission cannot be produced from SF alone for sources classified as RL.

However, definitions using radio flux or luminosity ratios are generally based on thresholds or boundaries in observational quantities rather than physical models. They group quasars over a wide range of redshifts and luminosities together to provide a single threshold, whereas the quasar radio emission is known to scale (non-linearly) with both of these parameters (e.g. Y24). Therefore, in this work, based on the two-component model presented in Section 3, we propose a new physically motivated quasar classification derived from the fitted contribution of the SF and AGN components in the quasar radio emission. We define the expression for the threshold radio luminosity  $L_{\text{eq}}$  (shown as the vertical solid line in Fig. 4), where  $P_{\text{SF}}(L_{\text{eq}})dL = P_{\text{AGN}}(L_{\text{eq}})dL$ , using the expressions for  $P_{\text{SF}}(L)dL$  and  $P_{\text{AGN}}(L)dL$  presented in equations (3) and (4) (note that the expressions of  $P_{\text{SF}}(L)dL$  and  $P_{\text{AGN}}(L)dL$  depend on the model best-fits and therefore values of  $L_{\text{eq}}$  vary between different grid cells). By definition, the AGN contribution to radio emission in sources with radio luminosities below  $L_{\text{eq}}$  is very likely to be smaller than the SF contribution ( $P_{\text{AGN}} < P_{\text{SF}}$ ), and therefore these sources can be classified as SF-dominated (traditionally RQ) quasars.

On the other hand, as shown by the vertical-dashed line in Fig. 4, we use  $L_{\text{pl}}$  as a radio luminosity lower limit for AGN-dominated sources (traditionally RL), since their fitted PDFs of radio flux density distribution are dominated by the single power-law AGN component.  $L_{\text{pl}}$  is defined as the luminosity threshold where the power-law component strongly dominates the PDF:  $P_{\text{AGN}}(L > L_{\text{pl}})dL > 0.95 P(L > L_{\text{pl}})dL$  so that 95 per cent of. This is a conservative approach to minimize the contamination of this AGN-dominated sample by radio intermediate (RI) sources lying at the bright end of the Gaussian SF distribution, as their radio emissions could be dominated by either SF or AGN. As a result, RI sources have radio luminosities between  $L_{\text{eq}}$  and  $L_{\text{pl}}$ . Since the origin of radio emission in RI quasars is indistinguishable by the model, we exclude them from the following analysis. To avoid any confusion,

<sup>3</sup>Here,  $\alpha$  is defined with  $L_\nu \propto \nu^{-\alpha}$ .

since our quasar classification is based on actual contributions of SF and AGN activities rather than radio-loudness thresholds, we will use the terms ‘SF-dominated’ and ‘AGN-dominated’ instead of RQ/RL in the following analysis.

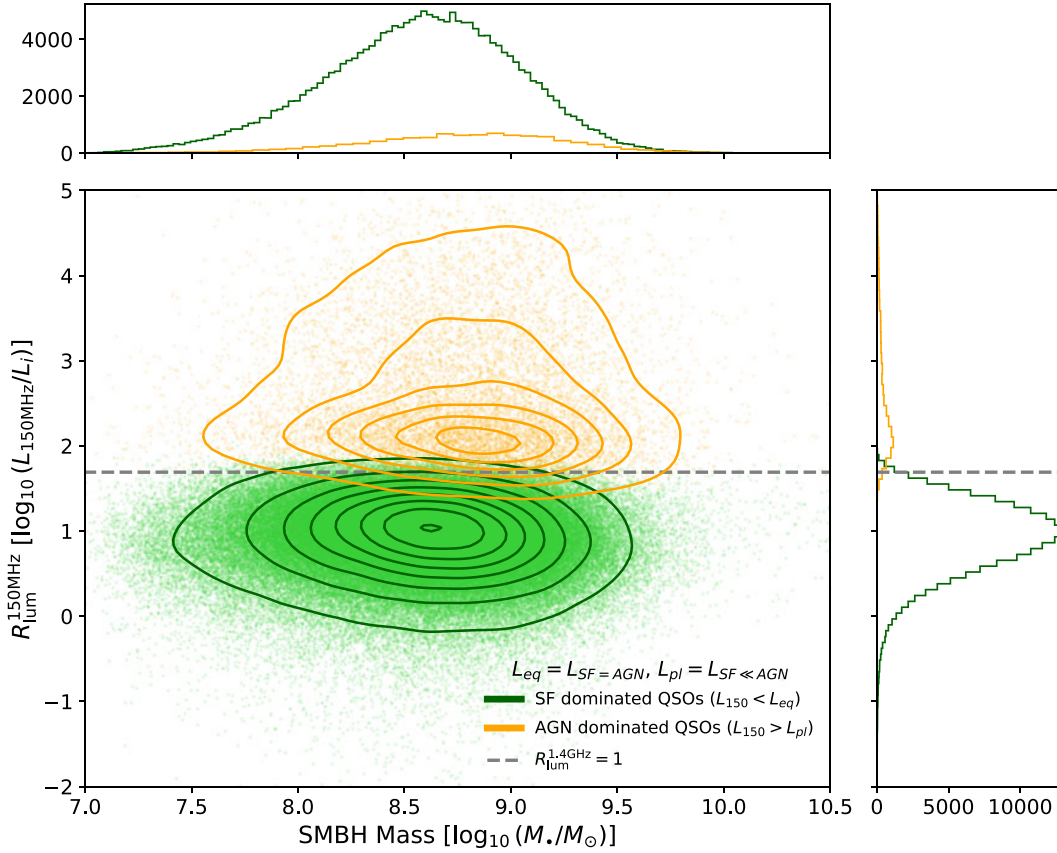
Fig. 6 shows the distribution of the SF-dominated (dark green) and AGN-dominated (light orange) quasars under our new definition, plotted in the radio-loudness-BH-mass space, where the radio loudness is defined as  $R = \log_{10}(L_{150\text{MHz}}/L_i)$  in order to compare with previous studies. The contours highlight the number density in the distribution by marking the same percentiles for both populations. As shown by the contours, both the RL and the RQ quasar population span across the entire BH mass range, suggesting that there is no sharp bimodality in the BH mass distribution of the RL and RQ quasars. This being said, RL quasars under our classification tend to have a higher BH mass on average, as suggested by the histogram on the top panel.

The marginalized distributions of both populations in radio loudness and in BH mass space are shown as histograms on the side and top panels. As seen from the dashed line on the right histogram (which illustrates  $R_{\text{lum}} = \log_{10}(L_{1.4\text{GHz}}/L_i) = 1$ ), our definition is comparable to the criteria used in the previous literature, since the dashed line falls on the tails of both the SF- and AGN-dominated population. However, unlike previous definitions, the SF- and AGN-dominated quasar populations overlap in their radio loudness values (see the histogram on the right) due to their different levels of SF and AGN activity in the  $\mathcal{M}_i - z$  space. This addresses the difference between a more physically based classification and the traditional RL/RQ quasar classification based solely on the radio loudness.

In Fig. 7, we plot the cumulative distribution function (CDF) of the BH mass for SF- and AGN-dominated quasars under the new definition we proposed, across the explored  $\mathcal{M}_i - z$  grid cells. The corresponding characteristic luminosities  $L_{\text{eq}}$  and  $L_{\text{pl}}$  for each grid cell are shown in Table A1 of Appendix A as a reference. Within each grid cell, the BH mass CDF of SF-dominated quasars is shown as a green solid line, whereas for AGN-dominated quasars, the CDF is shown as an orange-dashed line. Both distributions are characterized by their median BH masses, shown by the green and orange vertical lines, respectively. Across many  $\mathcal{M}_i - z$  grid cells, the cumulative distributions of BH mass show little difference between SF- and AGN-dominated quasars in the low-mass regime (well below the median RQ quasar BH masses); on the other hand, the AGN-dominated quasar population tends to host more massive BHs than SF-dominated quasars in the high-mass regime, as shown by the difference in average BH masses for the SF- and AGN-dominated quasars. To quantify the difference between SF- and AGN-dominated quasar populations, we perform K-S tests between BH mass CDFs for SF- and AGN-dominated quasars in each grid cell, and the  $p$ -values are shown in Fig. 7 for each grid cell. We are able to reject the null hypothesis that the two CDFs are identical ( $p$ -value  $< 0.05$ ) in most grid cells.

Our result is in agreement with the results of McLure & Jarvis (2004) and Whittam et al. (2022). It also agrees with the result in Fig. 5 where the quasars that host the most massive BHs are more likely to be radio-bright when compared to the rest of the population. Note that although  $L > L_{\text{pl}}$  proves to be a good proxy for AGN-dominated sources in all grid cells, the exact shape of the CDF can still be affected by minor features in the AGN tail of the radio flux density distribution that deviate from the single power-law assumption used in our model – for example, the SF- and AGN-dominated distributions arguably come closer together for quasar samples in more luminous  $\mathcal{M}_i$  grid cells, where  $p$ -values rise slightly above 0.05.





**Figure 6.** The distribution of AGN-dominated (light orange) and SF-dominated (dark green) quasars under our new model-driven classification in the  $R_{\text{lum}} - M_{\text{BH}}$  space, where  $R_{\text{lum}}$  is the radio loudness defined as  $R_{\text{lum}} = \log_{10}(L_{150\text{MHz}}/L_i)$ . With characteristic luminosity defined as  $L_{\text{eq}}$  where  $P_{\text{SF}}(L_{\text{eq}})dL = P_{\text{AGN}}(L_{\text{eq}})dL$ , SF-dominated (radio-faint) quasars are defined as quasars with radio luminosities  $L < L_{\text{eq}}$ , and AGN-dominated (radio-bright) quasars are defined as having radio luminosities  $L > L_{\text{pl}}$ , where the power-law component  $P_{\text{AGN}}(L)dL$  dominates over 95 per cent of the total PDF when  $L > L_{\text{pl}}$ . The contours track the same fractions of the AGN-dominated and SF-dominated quasar populations in each class, while the histograms show marginalized distributions in the  $R$  and  $M_{\text{BH}}$  spaces. The histogram at the top shows the marginalized distribution of  $M_{\text{BH}}$  for AGN-dominated (light orange) and SF-dominated (dark green) quasars, and the histogram on the right shows the marginalized distribution of  $R$  following the same colour code for the SF- and AGN-dominated population. Note that the bin widths of the histograms for SF- and AGN-dominated populations are defined with Scott’s rule and are thus different from each other. The criterion for RQ quasars used in Baloković et al. (2012) and Arnaudova et al. (2024) ( $R_{\text{lum}} < 1$ ) is shown as the horizontal grey-dashed line in the right-hand histogram and the scatter plot.

Given the broadly consistent offset between SF- and AGN-dominated populations in the  $\mathcal{M}_i - z$  space (as shown in Fig. 7), we here combine the  $\mathcal{M}_i - z$  grid cells to get a more compact and detailed look at the results, while also investigating the impact of different RL/RQ classifications from the literature.

The BH mass CDFs under our model-driven quasar classification are stacked across the  $\mathcal{M}_i - z$  grid cells in Fig. 7 and plotted in the top left panel of Fig. 8. This is consistent with what we show in Fig. 7, with  $p$ -value = 0.0054 rejecting the null hypothesis. In the bottom panels of Fig. 8, we bin the same quasar sample set into RQ and RL quasars based on the different traditional radio loudness criteria mentioned above and plot the corresponding BH mass CDFs for the RQ and RL population. We include the equivalent of Fig. 7 for these classification criteria in Fig. B1 of Appendix B to give an overview of the evolution of their BH mass CDFs in the  $\mathcal{M}_i - z$  space.

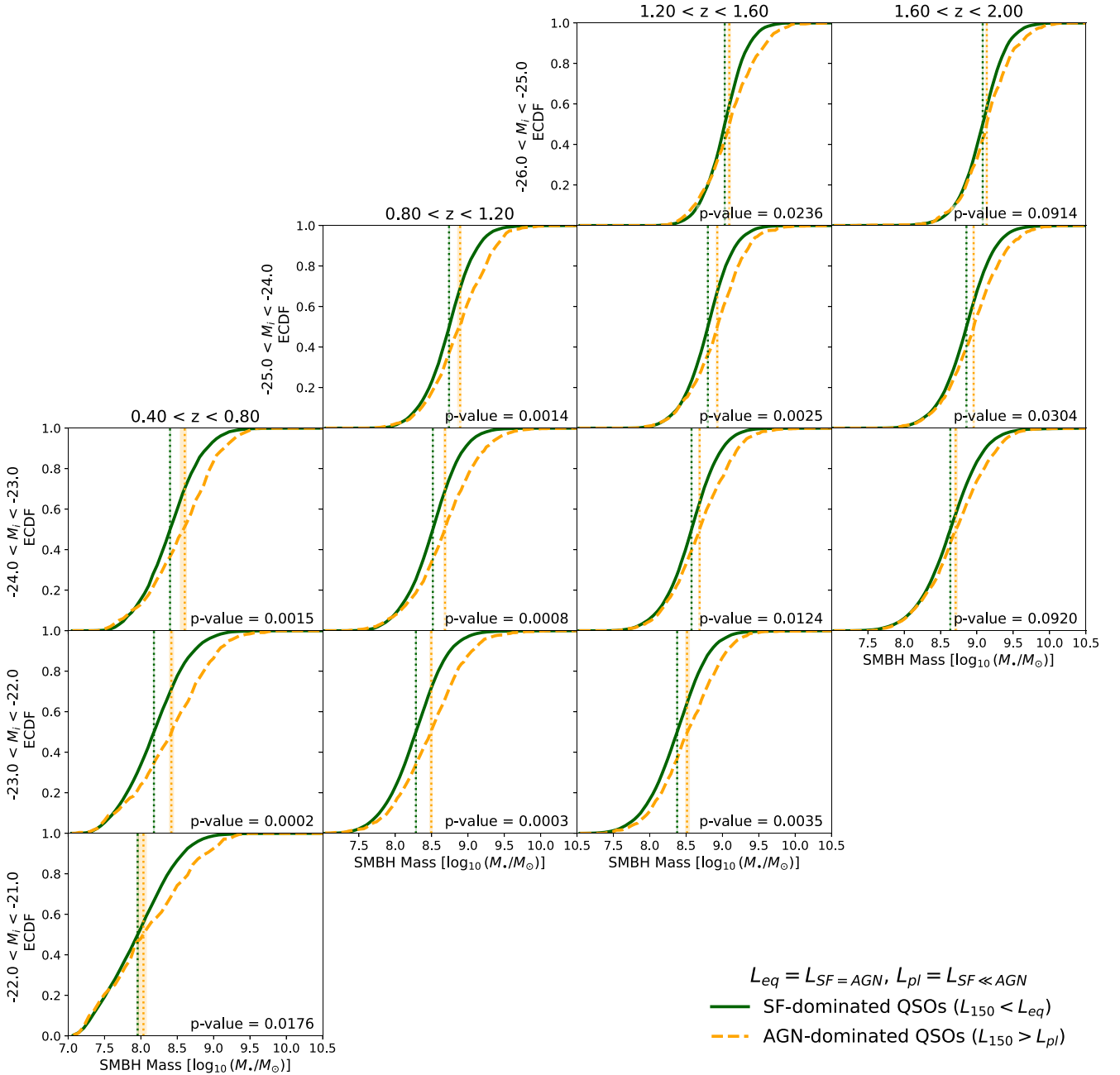
Specifically, in the bottom left panel of Fig. 8, we examine the radio loudness indicator used in McLure & Jarvis (2004):  $R_{\text{flux}} = f_{5\text{GHz}}/f_{4400\text{\AA}}$ , where the solid deep green line denotes the RQ quasars with  $R_{\text{flux}} < 10$ , and the dashed orange line shows the RL quasars with  $R_{\text{flux}} > 10$ . The mean BH mass for each population is shown as a vertical dotted line for RL quasars and a dash-dotted line for

RQ quasars. Here, the 5 GHz radio flux density is converted from the LOFAR 150 MHz radio flux density by assuming a radio spectral slope of  $\alpha = 0.7$ ,<sup>4</sup> and the 4400 Å flux density is calculated from the fitted rest-frame 3000 Å continuum luminosity in Wu & Shen (2022) by assuming an optical spectral index of  $\alpha_{\text{opt}} = 0.5$ . Despite the different radio-loudness indicators used, the trend remains consistent with Fig. 7 where AGN-dominated quasars host more massive BHs in the high-mass regime, which is reflected by the difference in the mean BH mass of the RL and RQ samples. For comparison, the bottom right panel shows similar CDFs of RQ and RL quasar BH masses under the classification criteria in Baloković et al. (2012) and Arnaudova et al. (2024) instead. Compared to the bottom left panel, the difference between the RL and RQ population is less significant ( $p$ -value = 0.20), which is consistent with their conclusions.

Going into more detail, Table 1 compares the number statistics of quasars classified into different populations according to different criteria. Each row shows one of the three quasar populations

<sup>4</sup>We pick this value to keep in consistency with previous works including Gürkan et al. (2019), Smith et al. (2021), and Arnaudova et al. (2024).

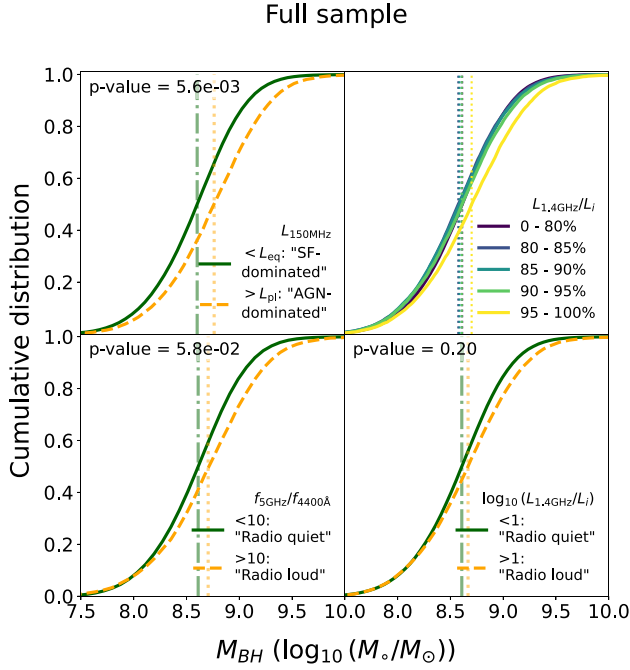




**Figure 7.** Cumulative mass distributions of AGN- (light orange) and SF-dominated (dark green) quasars defined in this work within each  $M_i-z$  grid. The grid cells are arranged by their average  $i$ -band luminosity and redshift – grid cells with fainter luminosity are placed towards the bottom, and those with lower redshift are placed towards the left. The vertical solid and dashed lines show the median BH mass for SF-dominated and AGN-dominated quasars, respectively, and the uncertainty is determined from the 16 and 84 per cent confidence intervals when bootstrap sampling the distributions. We quantify the similarity between two populations with  $p$ -values calculated from K-S tests:  $p$ -values less than 0.05 reject the null hypothesis that two distributions are identical. Our new definition leads to a universal difference in the BH mass distribution between radio-faint and radio-bright quasars, suggesting that AGN-dominated quasars tend to be more massive than SF-dominated quasars.

under our model-driven definition: SF-dominated (SF), RI, and AGN-dominated (AGN); the columns outline the classification under the luminosity ( $R_{\text{lum}} = \log_{10}(L_{1.4\text{GHz}}/L_i)$ ) and flux ( $R_{\text{flux}} = f_{5\text{GHz}}/f_{4400\text{\AA}}$ ) definition, where  $R_{\text{lum}} < 1$  ( $R_{\text{flux}} < 10$ ) quasars are defined as RQ and  $R_{\text{lum}} > 1$  ( $R_{\text{flux}} > 10$ ) quasars are defined as RL. The luminosity definition provides a clean selection of SF-dominated quasars in their RQ population, with 94.2 per cent of the RQ quasars dominated by SF; only 0.6 per cent of which are AGN dominated.

However, the single-value threshold introduces a fair amount of SF-dominated and RI quasars into the RL population, with only 56.3 per cent of the RL quasars are being securely AGN dominated. This contaminated selection significantly reduces the BH mass difference between the SF- and AGN-dominated quasars when they compared their RL samples with RQ samples. The flux definition, on the other hand, provides a cleaner selection in their RL population, with 83.1 per cent of them being AGN dominated. However, their



**Figure 8.** Difference in the BH mass CDF between RL/RQ (SF/AGN dominated) quasars under different definitions. We carefully examined the origin of quasar radio emission at different  $M_{\text{BH}}$  using our two-component model, and reproduce the difference in the cumulative BH mass distributions of SF- and AGN-dominated quasars (upper left panel). By binning quasars into different radio loudness percentiles (upper right panel), we show that only the top 5 per cent in quasar radio-loudness is responsible for the difference in the BH mass CDF. In the lower panels, quasars are classified into RL/RQ based on their flux density ratio (top left panel; e.g. McLure & Jarvis 2004) and luminosity ratio (top right panel; e.g. Baloković et al. 2012; Arnaudova et al. 2024) between radio and optical bands, with the thresholds listed in the corresponding panels. We quantify the difference of cumulative BH mass distributions between RL (AGN-dominated) and RQ (SF-dominated) quasars with  $p$ -values calculated from K-S tests performed on the two populations; smaller  $p$ -values indicate that the two distributions are less likely to be identical.

RL quasar selection also misses out 2672 AGN-dominated quasars, which accounts for 23.1 per cent of the entire AGN-dominated quasar population – twice the number of the 1328 dropouts under the luminosity definition. Despite the incomplete selection of AGN-dominated quasars, RL and RQ populations under the flux definition still reflect the difference in their BH mass distribution, since the number of contamination sources introduced into the RL sample is 26.8 per cent less than the luminosity definition.

To better understand these differences, in the upper right panel of Fig. 8, we divide the quasars by their percentiles in the radio loudness distribution (here defined as  $R_{\text{lum}} = L_{1.4\text{GHz}}/L_i$ ), and the quasars within the top 20 per cent of the radio loudness distribution are further binned into 5 per cent wide percentile bins. This is motivated by the definition of RL/RQ AGNs in Whittam et al. (2022), where they argued that AGNs at  $1\sigma$  beyond the IRRC – which corresponds to the 70th percentile in the  $L_{\text{radio}}/L_{\text{bol}}$  distribution – are probably not powerful enough to be classified as RL, and they adopted the  $2\sigma$  radio excess – which corresponds to the 95th percentile in the distribution – as the threshold for RL AGNs.

From Fig. 8, we can see that the difference in BH mass CDF occurs only in 5 per cent of the quasar population in this study; this fraction is consistent with the fraction of AGN-dominated quasars in the entire sample (5.6 per cent; see Table 1). This confirms that the differences

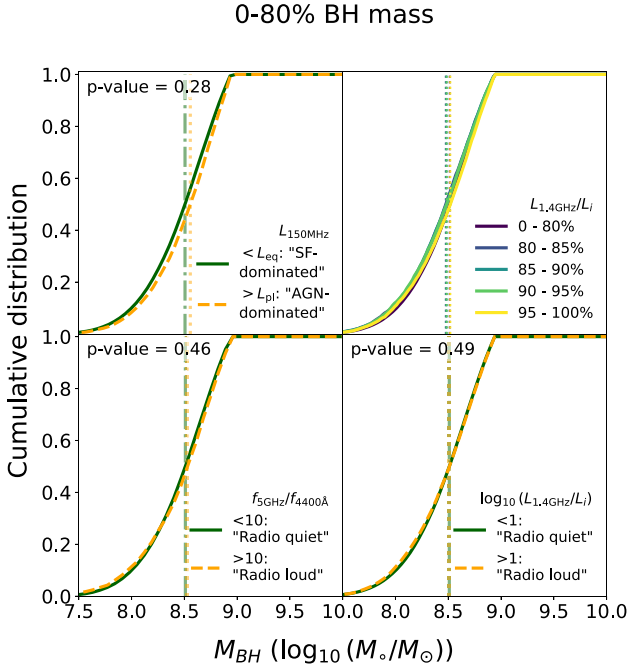
**Table 1.** Number statistics of quasars being classified into RL/RQ under the luminosity definition ( $R_{\text{lum}} = \log_{10}(L_{1.4\text{GHz}}/L_i)$ ) or flux definition ( $R_{\text{flux}} = f_{5\text{GHz}}/f_{4400\text{Å}}$ ), and quasars being classified into SF-dominated (SF), radio intermediate (RI), and AGN-dominated (AGN) under the two-component model definition presented in this work. The bottom row shows the total number of quasars being classified into RL/RQ under either  $R_{\text{lum}}$  or  $R_{\text{flux}}$  definition, which are then used as the benchmark for calculating the percentage of RL/RQ quasars being classified into SF, RI, or AGN in this work (shown in brackets below the numbers). The rightmost column shows the total numbers and percentages of the SF, RI, and AGN quasar population.

	$R_{\text{lum}} = \log_{10}\left(\frac{L_{1.4\text{GHz}}}{L_i}\right)$		$R_{\text{flux}} = \frac{f_{5\text{GHz}}}{f_{4400\text{Å}}}$		Total
	RQ	RL	RQ	RL	
SF	192 717 (94.2 per cent)	1890 (10.4 per cent)	194 566 (91.7 per cent)	41 (0.4 per cent)	194 607 (87.4 per cent)
RI	10556 (5.2 per cent)	6048 (33.3 per cent)	14834 (7.1 per cent)	1770 (16.5 per cent)	16604 (7.4 per cent)
AGN	1328 (0.6 per cent)	10243 (56.3 per cent)	2672 (1.2 per cent)	8899 (83.1 per cent)	11571 (5.2 per cent)
Total	204 601 (100 per cent)	18 181 (100 per cent)	212 072 (100 per cent)	10 710 (100 per cent)	222782 (100 per cent)

in BH mass only occur in the AGN tail of the radio flux density distribution, not the SF population. Furthermore, as noted earlier, this difference in BH mass CDF only becomes more pronounced for BH with masses larger than  $\log(M/M_{\odot}) = 8.5$ , which is in line with the trend presented in Fig. 5 where the quasars with the most massive BHs tend to be more RL. These results show that AGN-dominated quasars have a  $M_{\text{BH}}$  distribution that differs from their SF-dominated counterparts, suggesting a connection between  $M_{\text{BH}}$  and jet production.

Finally, we check the consistency between this analysis and the earlier analysis from our Bayesian model fitting. In Fig. 5, we had shown that among the AGN-dominated subset, it was only the quasars within the top 20 per cent of the BH mass distribution in each grid cell that contributed to the excess radio emission. Therefore, in Fig. 9, we remove all quasars in the top 20 per cent of the BH mass distribution in each grid cell, and again plot the BH mass CDFs from Fig. 8. By applying the mass cut, the BH mass distribution becomes consistent between RL and RQ quasars under all classification methods, with  $p$ -values rising above 0.2 for all criteria. This is in full consistency with our results from Fig. 5. We therefore further secured the origin of the excessive radio loudness associated with BH mass down to the most massive (top 20 per cent in the BH mass distribution) and most RL (top 5 per cent in the radio loudness distribution, dominated by the AGN component) quasars alone.

Combining the analysis from Sections 3 and 4, we can conclude that under most scenarios there is no fundamental connection between quasar radio emission and BH mass since the radio emission from the host galaxy does not depend on  $M_{\text{BH}}$ , and radio emission from the AGN remains the same across most BH mass percentiles; however, for the ‘genuine’ RL quasars where powerful AGN jets dominate the quasar radio emission, there is a boost in the level of AGN jet activity in quasars above a certain BH mass threshold. Since all of the panels in Fig. 8 trace the BH mass distribution of the same quasar population, the disagreement between previous results originates solely from definitions of quasar radio loudness that fail to reflect the physical processes behind the scene. More



**Figure 9.** The BH mass CDF after removing the top 20 per cent in the BH mass distribution, separated into RL (AGN-dominated) and RQ (SF-dominated) quasars based on various definitions of radio loudness. In contrast to Fig. 8, there is no difference between RL and RQ population under all radio loudness definitions ( $p$ -value  $> 0.05$ ), which confirms that the excess in quasar radio loudness only comes from the most massive (top 20 per cent) quasars.

specifically, the reason that Gürkan et al. (2019) and Arnaudova et al. (2024) did not identify a strong dependence on  $M_{\text{BH}}$  between the RQ and RL populations is that the assumed RL threshold classified  $\sim 30$  per cent RI sources into RL when their radio emission is not necessarily dominated by the AGN component, and that the  $R_{\text{lum}}-M_{\text{BH}}$  distribution is dominated by the SF-dominated quasars; as a result, the BH mass distribution of the RL samples in Arnaudova et al. (2024) shows a higher level of similarity to their RQ samples, and Gürkan et al. (2019) failed to find any correlation between  $R_{\text{lum}}$  and  $M_{\text{BH}}$ . Note that Arnaudova et al. (2024) report a slightly higher  $M_{\text{BH}}$  for RL quasars when using H  $\beta$  and [O III] line measurements, which is also in line with our result in Fig. 8. On the other hand, the RL quasar population defined in McLure & Jarvis (2004), although incomplete, was not significantly contaminated by non-AGN-dominated quasars, resulting in a difference in the BH mass distribution between RL and RQ quasars. Our new understanding of traditional radio loudness definitions helps unify these different observational results into a coherent framework.

## 5 TESTING FOR AGN OUTFLOW SIGNALS

Outflow structures in quasar emission lines can lead to systematically higher measured FWHMs (and hence BH masses) within the published SDSS catalogues, as shown in Coatman et al. (2017) for C IV BH mass measurements. Although C IV BH masses are already excluded from our sample, quasars with the most extreme outflow activity could still be misclassified into more massive BH mass bins (i.e. with BH masses being overestimated); in turn, this would bring a radio excess into the quasar population hosting most massive BHs, since outflow activities are associated with higher radio detection

fraction of quasars (e.g. Richards et al. 2021; Petley et al. 2024, Escott et al. 2025).

To examine the potential observational bias of the BH mass in our quasar sample and to quantify the impact of AGN outflow activities on our conclusion, we perform an analysis of the stacked properties of quasars in our sample set. We start by separating our quasars into four quadrants based on their BH masses in the SDSS catalogue and the dominant source of their radio emission:

- (i) Quadrant 1: quasars hosting moderate mass BHs (0–80 per cent in the BH mass distribution of quasars within the corresponding  $M_i - z$  grid cell) with radio emission dominated by SF activity;
- (ii) Quadrant 2: quasars hosting massive BHs (top 20 per cent in the BH mass distribution of quasars within the corresponding  $M_i - z$  grid cell) with radio emission dominated by SF activity;
- (iii) Quadrant 3: quasars hosting moderate mass BHs with radio emission dominated by AGN activity; and
- (iv) Quadrant 4: quasars hosting massive BHs with radio emission dominated by AGN activity.

In total, 157 229 quasars are classified into quadrant 1, 37 378 quasars are classified into quadrant 2, 8156 quasars are classified into quadrant 3, and 3415 quasars are classified into quadrant 4. The rest of quasars either are defined as RIs or have spurious BH mass measurements ( $M_{\text{BH}} \leq 0$ ) in the SDSS DR16Q catalogue and are thus not included in the classification.

To robustly compare the difference in physical properties between different quadrants while ensuring that the differences are not driven by underlying correlations with  $M_i$  or  $z$ , we match the quasars of different quadrants in the  $M_i - z$  space. We start by defining the distance in the  $M_i - z$  space as  $d = \sqrt{\Delta M_i^2 + (\epsilon \Delta z)^2}$ , where the coefficient  $\epsilon$  is defined as the length ratio of each  $M_i - z$  grid cell,  $\epsilon = 1/0.4$ , to compensate for the relative difference in the scale of  $M_i$  and  $z$  in the Euclidean space. Next, we define the ‘closest neighbour’ pair as two quasars having the smallest distance in the  $M_i - z$  space. We then match every quasar in quadrant 4 with their closest neighbour in quadrants 1, 2, and 3. The matched quasars in quadrants 1, 2, and 3 therefore share the same redshift and bolometric luminosity properties as quasars of quadrant 4.<sup>5</sup> We then use the matched quasars in quadrants 1, 2, 3, and 4 for our analysis below.

First, we focus on the potential bias in the BH mass measurements by investigating the difference in the stacked spectra of the Mg II emission line, which is used to obtain single-epoch BH mass estimations in our sample.

As briefly discussed above, strong AGN outflows will be visible from the shape of the emission lines as an extended wing on the blue end. Such line asymmetry will lead to an overestimation in the single-epoch emission line FWHM and hence to an overestimation in the BH mass. Although the individual quasar spectra from SDSS DR16Q are often noisy, we can take advantage of the large size of our quasar sample to test for such a systematic effect by stacking the spectra of quasars with similar redshifts and luminosities. We select a narrow ( $\Delta \log M_{\text{BH}} = 0.2$  dex) range of BH mass at the high  $M_{\text{BH}}$  end, and stack the quadrant 2 and 4 quasars in that  $M_{\text{BH}}$  range for different  $M_i - z$  grid cells. We then compare the stacked emission line profiles to check for any asymmetry caused by SF- or AGN-dominance in quasars.

We find that the Mg II emission line profiles are similar across all  $M_i - z$  grids, with no evidence of systematically different outflow

<sup>5</sup>In the remaining part of this section, we will use quadrant 1 to 4 quasars to refer to the *matched* quasar population within each quadrant.



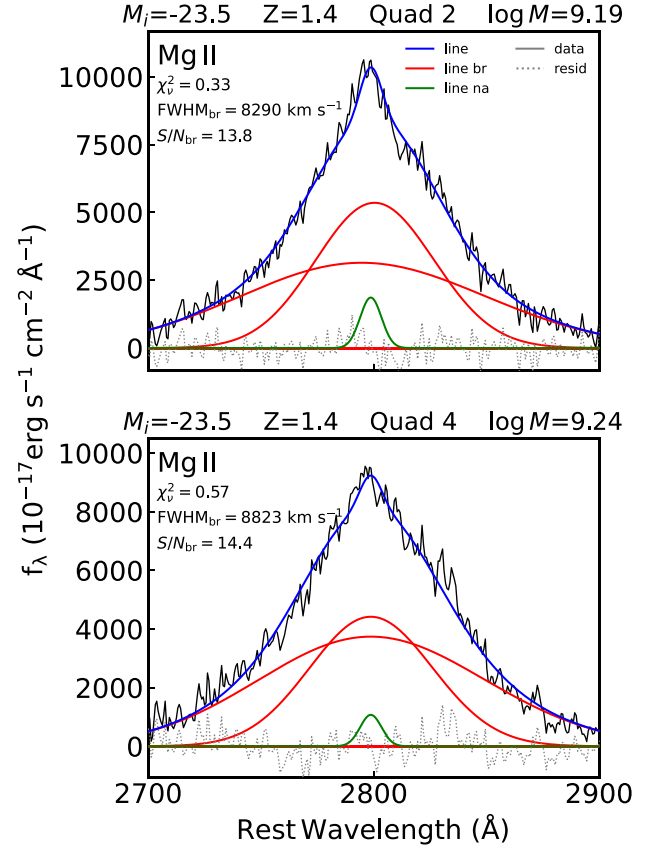
components in AGN-dominated (quadrant 4) quasars compared to SF-dominated (quadrant 2) quasars, indicating that outflow features in radio-bright quasars do not have a systematic effect on the BH mass measurements in our sample. As an example, within a representative  $\mathcal{M}_i-z$  grid cell, we select quadrants 2 and 4 quasars with a catalogue BH mass of  $9.1 < \log(M_{\text{BH}}/M_{\odot}) < 9.3$ , and we show the Mg II components of the stacked SDSS spectra of these quasars in Fig. 10. Here, we adopt the Mg II BH masses from Wu & Shen (2022) when selecting quasars in the stack. The average catalogue values of the BH mass are  $\log(M_{\text{BH}}/M_{\odot}) = 9.19$  for both stacks. To test the recovery of  $M_{\text{BH}}$ , we fit the entire spectrum with the PYQSOFIT code (Shen et al. 2019) while adopting the Mg II BH mass recipe from Vestergaard & Osmer (2009):

$$\log(M_{\text{BH}}/M_{\odot}) = 6.86 + 2 \log\left(\frac{\text{FWHM}(\text{Mg II})}{1000 \text{ km} \times \text{s}^{-1}}\right) + 0.5 \log\left(\frac{L_{3000}}{10^{44} \text{ erg} \times \text{s}^{-1}}\right), \quad (5)$$

where  $L_{3000}$  is the monochromatic luminosity  $\lambda L_{\lambda}$  at  $3000 \text{ \AA}$ . When computing the fitted BH masses, we use the best-fitting values from PYQSOFIT for  $\text{FWHM}(\text{Mg II})$ , and we average the  $L_{3000}$  values obtained from Wu & Shen (2022) for quasars in the stack as input for  $L_{3000}$ . We find that recovered BH masses are very close to the median input mass. For the specific example shown in Fig. 10, the stacked Mg II line for quadrant 4 quasars has a higher FWHM than that measured with quadrant 2 quasars, leading to a 0.05 dex difference between the Mg II BH mass calculated from the stack ( $\log M_{\text{BH}}/M_{\odot} = 9.24$ ) and the average Mg II BH mass from catalogue. This suggests that the BH masses for quasars in quadrant 4 might actually be underestimated in this case (the opposite of the potential bias we were concerned about).

Next, we examine the interplay between the level of outflow activity, BH mass, and the dominant source in radio emission using the properties of C IV emission line as a tracer for AGN outflow. Recent works (e.g. Richards et al. 2011; Rankine et al. 2020; Richards et al. 2021; Temple et al. 2023) have explored the diagnostic power of 2D C IV equivalent width (EW) and blueshift space through the definition of ‘C IV distance’.<sup>6</sup> In Richards et al. (2011), the 2D parameter space of C IV EW and blueshift is proposed to have a connection with quasar accretion modes, where disc-dominated accretions have higher C IV EW and low blueshifts, whereas wind-dominated accretions tend to reside in the parameter space with low C IV EW and high blueshifts. Richards et al. (2021) further quantified this difference in C IV space by converting the 2D distribution into a single value of C IV distance. When calibrated against the SDSS-RM quasar samples, the C IV distance itself can act as a proxy for ‘disc’ versus ‘wind’ activity (Richards et al. 2011; Rankine et al. 2020), with lower C IV distance being associated with ‘disc’ model and higher C IV distance to ‘wind’ model. Through extensive modelling of quasar spectral energy distribution (SEDs), Temple et al. (2023) quantified the differences in C IV and He II emission line properties and the underlying theoretical accretion model within the Eddington ratio ( $L/L_{\text{Edd}}$ ) and  $M_{\text{BH}}$  plane. Quasars with the highest C IV EW are found in the parameter space with a high Eddington ratio ( $L/L_{\text{Edd}} \gtrsim 0.2$ ) and a relatively low BH mass ( $\log(M_{\text{BH}}/M_{\odot}) < 9.0$ ), or with a low Eddington ratio and a high BH mass. Quasars with high C IV blueshifts are observed in a different part of the parameter space with a high Eddington ratio and a high

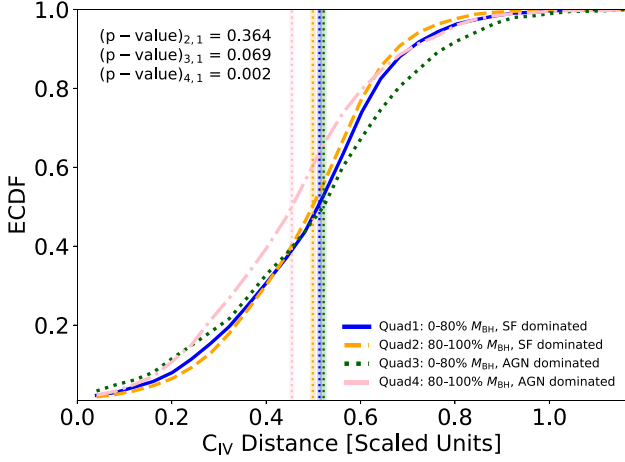
<sup>6</sup>We refer the readers to Richards et al. (2021) for the exact definition and calibration of C IV distance used in this work.



**Figure 10.** We show one representative  $\mathcal{M}_i-z$  bin ( $-24 < \mathcal{M}_i < -23$ ,  $1.2 < z < 1.6$ ) in which we stack the SDSS spectra of quasars that are matched in  $\mathcal{M}_i-z$  space and have catalogue BH masses of  $9.1 < \log(M_{\text{BH}}/M_{\odot}) < 9.3$  (within the highest 20 per cent BH mass bin), while being dominated by SF (quadrant 2, upper panel) and AGN (quadrant 4, lower panel) components, respectively. The stacked spectra (grey solid line) appear very similar. Both are then fitted with PYQSOFIT, with the best fits of broad emission line components and narrow line component plotted in red and green solid lines, respectively; the residual of the fit is plotted as grey dotted line. The blue solid line shows the best fit to the Mg II line, combining broad and narrow components. The BH mass values shown on top of each panel is measured from the stacked spectra, calculated with the broad line FWHM best fit (shown in the legend) using the recipe in Vestergaard & Osmer (2009); the BH masses from stack broadly recovers the median input value..

BH mass. Temple et al. (2023) also suggested a likely discrepancy between the existing simple SED models and the observation when the Eddington ratio drops below  $L/L_{\text{Edd}} \approx 0.2$ , which may be the result of a transition phase in the accretion geometry (Giustini & Proga 2019).

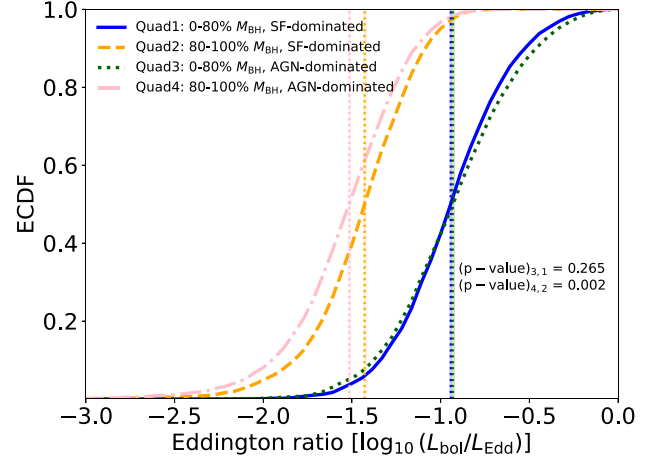
To characterize the distribution of quasars from the four quadrants in the C IV space, we use the recipe from Richards et al. (2021) to compute C IV distances for every quasar in our matched sample where C IV emission lines are visible from SDSS spectra. Fig. 11 shows the CDF of C IV distances for quasars within each quadrant, where quadrants 1 to 4 are colour-coded with blue solid line, orange-dashed line, green-dotted line and pink dash-dotted line, respectively. The vertical lines show the average values within each quadrant, and the uncertainty is determined by bootstrap resampling of the sources. The difference between the distributions is quantified by the  $p$ -values of the K-S tests performed between quadrant 1 and the other quadrants.



**Figure 11.** The CDF of quasar C IV distances in quadrants 1 to 4, calculated from the best-fitting line in Richards et al. (2021). Note that this figure does not include quasars with  $\log C\text{ IV EW} > 2.5$  as they fall out of the domain of definition in Richards et al. (2021). Quadrants 1 to 4 are marked with blue solid line, orange-dashed line, green dotted line and pink dashed-dotted line, respectively. The mean values of each distribution are marked as solid-dotted lines, and the uncertainty is determined from the 16 and 84 per cent confidence intervals when bootstrap sampling the distributions. We have performed K-S tests between the C IV distance distribution from quadrant 1 and the rest of quadrants, and the corresponding  $p$ -values are shown in the upper left corner of the plot. While quadrants 2 and 3 show a higher similarity with quadrant 1, the  $p$ -values between quadrant 1 and 4 fall below 0.05, hence we reject the null hypothesis that the C IV distance distributions from these two quadrants are identical.

For the SF-dominated quasars in quadrants 1 and 2, their distribution in the C IV space shows minimal difference ( $p\text{-value} > 0.05$ ) since their C IV distance distributions are almost identical, suggesting that a common accretion mode is shared between all SF-dominated quasars. AGN-dominated quasars with moderate BH mass are more likely to reside at the high C IV distance end (C IV distance  $> 0.5$ ) than SF-dominated quasars. The most significant difference is between the AGN-dominated quasars with high BH mass and the rest of the populations: quadrant 4 quasars tend to reside in the lower C IV distance parameter space, which is reflected by both a lower average C IV distance than the rest and a low  $p$ -value (0.002) between the C IV distance distribution of quadrant 4 and quadrant 1 quasars, which in principle rejects the null hypothesis. However, note that our analysis does not take into account the significant uncertainties in individual measurements. The difference in quadrant 4 quasars suggests that these systems are less outflow dominated than the other quasar populations and that there is no evidence for an enhancement in AGN outflow in the AGN-dominated high-BH mass sample. This agrees with an increased radio detection fraction at the low C IV distance end reported in Richards et al. (2021). On the other hand, Richards et al. (2021) also discovered an increase of radio detection fraction at the high C IV distance end; since the AGN-dominated sources under our classification are by definition jet-dominated, such an increase can instead be attributed to the dominance of AGN wind and/or outflow in the radio emission of RI quasars.

We also investigated the difference in the BH accretion efficiency traced by the Eddington ratio ( $L/L_{\text{Edd}}$ ). Note that our samples are matched in quasar luminosity and that the Eddington ratio scales as  $L/M_{\text{BH}}$ , therefore this effectively traces the difference in  $M_{\text{BH}}$  between different quadrants. Hence, quasars with moderate BH mass (quadrants 1 and 3) inherently have a higher  $L/L_{\text{Edd}}$  than



**Figure 12.** The cumulative distribution of Eddington ratio in quadrants 1 to 4 quasar populations, plotted in blue solid line, orange dashed line, green dotted line and pink-dashed-dotted line, respectively. The vertical lines mark the mean value of Eddington ratio in each population. The  $p$ -values at the lower right corner are calculated from the K-S tests performed between quadrants 1 and 3, and quadrants 2 and 4, respectively. The Eddington ratio distribution is identical between quadrants 1 and 3; however, we reject the null hypothesis for quadrants 2 and 4.

quasars with high BH mass (quadrants 2 and 4), due to selection biases.

Fig. 12 plots the distribution of the Eddington ratio for quasars matched in the four quadrants, with the same colour code as in Fig. 11. For quasars with moderate BH masses, the SF-dominated and the AGN-dominated populations show no difference in their Eddington ratio distributions; this is expected since these quasars are matched in luminosities and redshifts while occupying the same BH mass percentiles. However, for quasars with massive BHs, the Eddington ratio distribution of AGN-dominated quasars is different from SF-dominated quasars ( $p\text{-value} = 0.002$ ), and AGN-dominated quasars with high BH masses tend to have lower Eddington ratios in general. This agrees with the findings in Section 4 where AGN-dominated quasars tend to be more massive than SF-dominated quasars (hence the lower Eddington ratio in matched samples), and such difference only occurs in quasars hosting the 20 per cent most massive BHs.

Combining the results in Figs 11 and 12, we can map the quadrant 4 quasars to the quasar population highlighted in Temple et al. (2023) with low C IV distances, low Eddington ratio and high BH masses. This suggests that alternative accretion physics in the broad-line region (e.g. Mitchell et al. 2023; Temple et al. 2023) may play a role in the uniqueness of this AGN-dominated massive quasar population.

## 6 DISCUSSION

We have investigated the relationship between BH mass and quasar radio loudness from two different perspectives: In Section 3, we binned the quasars by their BH mass and found that quasars with the most massive BHs have an excess in radio emission, while in Section 4 we show that quasars with extreme radio excess are also more likely to host more massive BHs. Combining the analysis from these two sections, we have singled out the quasar population that is the top 5 per cent in radio loudness and top 20 per cent in BH mass; only quasars of these highest BH masses show a different distribution of radio flux densities, and this distribution only differs within the

top 5 per cent of the radio luminosities, corresponding to the radio luminosities dominated by the AGN/jet component. This suggests a difference in the nature of these quasars compared to the rest of the population. In Section 5, we have gone on to show that this is not driven by a bias in  $M_{\text{BH}}$  measurements; yet these quasars do show different C IV line EWs, blueshifts and Eddington-scaled accretion rates than the rest of the quasar sample.

Although the data analysed in this paper do not allow us to draw a robust conclusion on the origin of this difference in radio flux densities, in this section, we consider a variety of different mechanisms or effects that might be driving this difference:

(i) BH accretion mode. Low accretion rate BHs are thought to host hot accretion flows, e.g. the advection-dominated accretion flow (ADAF; Narayan & Yi 1994, 1995), while high accretion objects host cold accretion flows generated from thin disc accretion. AGNs are then classified into jet-mode and quasar-mode (radiative-mode) based on their accretion mode (e.g. Heckman & Best 2014). Previous studies have constructed a fundamental plane of jet-mode AGNs using BH mass, core jet radio luminosity and disc X-ray luminosity (e.g. Merloni, Heinz & di Matteo 2003), which then links the radio jet production to the accretion rate and BH mass. However, for high accretion rate objects including radio quasars, such a fundamental plane is not well established (see e.g. Merloni & Heinz 2008; Plotkin et al. 2012; Inoue et al. 2017), implying multiple BH accretion modes may be at play. In Ichikawa & Inayoshi (2017), the authors suggested that for radio quasars, the difference between the RL fraction under different  $M_{\text{BH}}$  may be linked to the alternative jet production mechanism of AGNs, where a large enough BH mass is required for a powerful ( $L_{\text{mech}} > L_{\text{bol}}$ ) radio jet to form and maintain itself. Under this scenario, the gas flow at the core region of the SMBH converts into an ADAF from the original accretion disc. This suggests a transitional phase where a jet with enough mechanical power can inject itself into the ambient gas background and form shock regions where the radio emission is generated by non-thermal electrons, while the exterior of the accretion disc remains stable, or the AGN accretion disc becomes fragmented due to gravitational instability. As the formation of ADAF requires an Eddington fraction below  $\sim 0.1$ , for sources with a given bolometric luminosity this leads to a high BH mass. Studies have found a strong BH mass dependence in the jet-mode AGN activity (see Kondapally et al. 2022, and references therein). Meanwhile, Sikora & Begelman (2013) also suggested a mixture of accretion modes in the triggering of RL quasars, but through an alternative path where an episode of hot accretion is followed by a massive, cold accretion event. As a result, the large magnetic flux is trapped within the inner radius of the newly formed cold accretion disc and forms a magnetically choked accretion flow (Narayan, Igumenshchev & Abramowicz 2003; Reynolds, Garofalo & Begelman 2006; McKinney, Tchekhovskoy & Blandford 2012). This path is most likely associated with major merger activities involving a giant elliptical galaxy and a disc galaxy.

(ii) BH spin. It is widely believed that quasar radio jets are generated from a combination of magnetic field and rotational energy (e.g. Blandford & Znajek 1977, the ‘BZ’ mechanism); under this scenario, the jet power is correlated with BH spin. Results from general relativistic magnetohydrodynamic simulations (e.g. Hawley & Krolik 2006; Martínez-Sansigre & Rawlings 2011b) helped constrain the BH spin influence on the radio jet power from the theoretical side. Meanwhile, on the observational side, Martínez-Sansigre & Rawlings (2011a) parameterized the spin distribution of BHs with  $\log_{10}(M_{\text{BH}}/M_{\odot}) > 8$  from radio observations, and found

a small population of high excitation galaxies with both maximal spin (inferred from parameters associated with BH spin; indicating high radio loudness) and high accretion efficiency. Schulze et al. (2017) used the [O III] luminosity as an indirect tracer of BH spin, and suggested that RL quasars have systematically higher radiative efficiencies than RQ quasars, therefore higher BH spin. However, in these works, the sample selection is biased by the detection limit of FIRST survey, and their estimation of BH spin is limited by model uncertainties. Previous studies have associated high BH spin objects with major merger events (e.g. Wilson & Colbert 1995; Chiaberge & Marconi 2011; Chiaberge et al. 2015), implying a higher BH mass when compared to low spin BHs. Although whether BH spin governs the quasar radio loudness is still under debate (see e.g. Sikora & Begelman 2013; Inoue et al. 2017; Macfarlane et al. 2021), an alternative scenario could be where the power from magnetically driven outflow is distributed differently between the jet and disc component due to a combined effect of both BH spin and magnetic field configuration, as shown by Mehdipour & Costantini (2019) through performing high-resolution X-ray spectroscopy on RL AGNs, albeit with a small sample size. They have found an inverse correlation between the AGN wind and jet components (radio loudness), and also predicted disc truncation as a likely consequence of this scenario.

(iii) Environmental factors. Studies including Mandelbaum et al. (2009), Ross et al. (2009), and Shen et al. (2009) show that RL quasars tend to reside in more massive haloes and denser environments, while the link between quasar environment, BH mass (stellar mass) and jet power remains unclear. Further evidence for this was provided by Retana-Montenegro & Röttgering (2017), who used sources from SDSS and FIRST surveys and found that RL quasars are more strongly clustered than RQ quasars, while quasars powered by the most massive BHs also cluster more strongly than quasars with less massive BHs. It has been argued that denser environments can lead to a boosting of the radio luminosity (for fixed jet power) due to the confinement of radio lobes preventing energy losses of the relativistic electrons in radio lobes through adiabatic expansion (e.g. Barthel & Arnaud 1996); in addition Hatch et al. (2014) argue that the dense environment fosters the formation of radio jets. However, the conclusions of these studies could be affected by the different selection bias and definitions of radio loudness in AGN samples used by different studies (as highlighted in our analysis above) and the contamination from star-forming galaxies (see e.g. discussion in Magliocchetti 2022).

The nature of this extreme quasar population remains an interest for future studies, since the data required to distinguish between potential scenarios that explain such differences are not yet available. Robust measurements of the quasar emission line and continuum through next-generation optical spectroscopic surveys such as WHT Enhanced Area Velocity Explorer Surveys (WEAVE; Jin et al. 2024) – specifically the WEAVE–LOFAR survey (Smith et al. 2016) – and the Black Hole Mapper programme in SDSS-V will provide the necessary information to identify observational biases in the current quasar sample and provide extended BH mass calibrations. Together with accurate photometric measurements of host galaxy morphologies, stellar mass, and environmental factors from *Euclid* (Euclid Collaboration 2024), we can gather critical evidence to distinguish between these scenarios and determine the physical mechanism behind the radio excess in the most massive and RL quasars.



## 7 SUMMARY

In this work, we extend the fully Bayesian two-component model presented in Y24 to study the evolution of host galaxy SF and AGN jet activities with BH mass. We model the quasar radio emission by assuming that every quasar hosts both SF activity through their host galaxies (characterized by mean SFR  $\Psi$ ) and jet activity through central BHs (characterized by AGN jet normalization  $f$ ). We bin the quasar samples into grid cells based on their location in the parameter space of  $\mathcal{M}_i$ ,  $z$  and BH mass, before fitting the radio flux density distribution within each grid cell against our two-component model to get the best fits for host galaxy SFR and AGN jet normalization (relative AGN activity level compared to host galaxy SF) in our quasar sample. Both host galaxy and jet activities show little difference in most of the BH mass bins; however, we detect an excess in the jet activity level for quasars hosting the most massive (top 20 per cent in  $M_{\text{BH}}$ ) BHs, where the normalization of the power-law jet luminosity function ( $f$ ) is increased by a factor of  $\sim 2.5$ , suggesting these quasars are 2.5 times more likely to host powerful jets when compared to the rest.

In order to build a coherent picture on the relation between quasar radio emission and BH mass that explains the differences in the previous studies on the BH mass dependence of radio loudness, we present a series of new definitions on the radio quasar population – AGN-dominated ( $L_{150\text{MHz}} > L_{\text{pl}}$ ) and SF-dominated ( $L_{150\text{MHz}} < L_{\text{eq}}$ ) – based on the dominant component in quasar radio emission instead of radio loudness. Our results suggest that previous disagreements in the inferred dependence of radio loudness on BH mass originate from different definitions of RL/RQ quasars applied in their analysis, leading to contamination of the RL sample by SF-dominated sources, which significantly weaken any observed dependence. Our new definition, motivated by the actual physical model, can reconcile this disagreement. Through our analysis, the origin of the radio excess correlated with the BH mass can be attributed to just the most massive quasars (with top 20 per cent BH mass) and affect only the most RL (with top 5 per cent radio loudness) quasars dominated by AGN jets, at a given redshift and bolometric luminosity. We further prove that the differences in this particular quasar population are not driven by bias in BH mass measurements, and these quasars also show differences in their Eddington ratios and their distribution in the C IV space.

We present potential scenarios that could explain these observed differences, relating to BH accretion mode changes, BH spin, quasar environment, or observational biases in the  $M_{\text{BH}}$  measurements driven by outflows. Distinguishing between these scenarios is not yet possible, but remains of huge interest for future studies.

## ACKNOWLEDGEMENTS

BY would like to thank for the support from the University of Edinburgh and Leiden Observatory through the Edinburgh-Leiden joint studentship. KJD acknowledges funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 892117 (HIZRAD) and support from the STFC through an Ernest Rutherford Fellowship (grant number ST/W003120/1). LKM is grateful for support from the Medical Research Council (MR/T042842/1). Several authors are grateful for support from the UK Science and Technology Facilities Council (STFC) via grants ST/V000594/1 (PNB), ST/Y000951/1 (PNB), ST/V000624/1 (MIA), ST/T506047/1 (JP), ST/V506643/1 (JP), ST/X508408/1 (SS), ST/V000624/1 (DJBS),

and ST/Y001028/1 (DJBS). BY gratefully acknowledges Yuming Fu for the useful discussions on PYQSOFIT.

LOFAR data products were provided by the LOFAR Surveys Key Science project (LSKSP; <https://lofar-surveys.org/>) and were derived from observations with the International LOFAR Telescope (ILT). LOFAR (Haarlem et al. 2013) is the Low Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources), and which are collectively operated by the ILT foundation under a joint scientific policy. The efforts of the LSKSP have benefited from funding from the European Research Council, NOVA, NWO, CNRS-INSU, the SURF Co-operative, the UK Science and Technology Funding Council and the Jülich Supercomputing Centre. This research made use of the University of Hertfordshire high-performance computing facility and the LOFAR-UK computing facility located at the University of Hertfordshire and supported by STFC (ST/P000096/1).

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions.

SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is [www.sdss4.org](http://www.sdss4.org).

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics|Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

## DATA AVAILABILITY

The data sets used in this paper were derived from sources in the public domain: the LOFAR Two-Metre Sky Surveys ([www.lofar-surveys.org](http://www.lofar-surveys.org)) and the Sloan Digital Sky Survey ([www.sdss.org](http://www.sdss.org)). Solutions to  $L_{\text{eq}}$  and  $\sigma_{\mu}$  in  $\mathcal{M}_i-z$  grid cells not included in this work can be provided upon reasonable request to the lead author.

## REFERENCES

- Arnaudova M. I. et al., 2024, *MNRAS*, 528, 4547
- Baloković M., Smolčić V., Ivezić Z., Zamorani G., Schinnerer E., Kelly B. C., 2012, *ApJ*, 759, 30
- Barthel P. D., Arnaud K. A., 1996, *MNRAS*, 283, L45
- Becker R. H., White R. L., Helfand D. J., 1995, *ApJ*, 450, 559
- Blandford R. D., Znajek R. L., 1977, *MNRAS*, 179, 433

- Blundell K. M., Rawlings S., 2001, *ApJ*, 562, L5
- Chaves-Montero J. et al., 2022, *A&A*, 660, A95
- Chiaberge M., Marconi A., 2011, *MNRAS*, 416, 917
- Chiaberge M., Gilli R., Lotz J. M., Norman C., 2015, *ApJ*, 806, 147
- Cirasuolo M., Magliocchetti M., Celotti A., Danese L., 2003, *MNRAS*, 341, 993
- Coatman L., Hewett P. C., Banerji M., Richards G. T., Hennawi J. F., Prochaska J. X., 2017, *MNRAS*, 465, 2120
- Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, *AJ*, 115, 1693
- Dawson K. S. et al., 2016, *AJ*, 151, 44
- Dey A. et al., 2019, *AJ*, 157, 168
- Escott E. et al., 2025, *MNRAS*, 536, 1166
- Euclid Collaboration, 2024, preprint (arXiv:2405.13491)
- Fabian A. C., 2012, *ARA&A*, 50, 455
- Giustini M., Proga D., 2019, *A&A*, 630, A94
- Gloude-mans A. J. et al., 2021, *A&A*, 656, A137
- Gürkan G. et al., 2019, *A&A*, 622, A11
- Hardcastle M. J. et al., 2023, *A&A*, 678, A151
- Hatch N. A. et al., 2014, *MNRAS*, 445, 280
- Hawley J. F., Krolik J. H., 2006, *ApJ*, 641, 103
- Heckman T. M., Best P. N., 2014, *ARA&A*, 52, 589
- Hinshaw G. et al., 2013, *ApJS*, 208, 19
- Ho L. C., Ulvestad J. S., 2001, *ApJS*, 133, 77
- Ichikawa K., Inayoshi K., 2017, *ApJ*, 840, L9
- Inoue Y., Doi A., Tanaka Y. T., Sikora M., Madejski G. M., 2017, *ApJ*, 840, 46
- Ivezić Z. et al., 2002, *AJ*, 124, 2364
- Jarvis M. J. et al., 2017, preprint (arXiv:1709.01901)
- Jin S. et al., 2024, *MNRAS*, 530, 2688
- Kaspi S., Smith P. S., Netzer H., Maoz D., Jannuzi B. T., Giveon U., 2000, *ApJ*, 533, 631
- Kellermann K. I., Sramek R., Schmidt M., Shaffer D. B., Green R., 1989, *AJ*, 98, 1195
- Kondapally R. et al., 2021, *A&A*, 648, A3
- Kondapally R. et al., 2022, *MNRAS*, 513, 3742
- Kormendy J., Ho L. C., 2013, *ARA&A*, 51, 511
- Lyke B. W. et al., 2020, *ApJS*, 250, 8
- Macfarlane C. et al., 2021, *MNRAS*, 506, 5888
- Magliocchetti M., 2022, *A&AR*, 30, 6
- Mandelbaum R., Li C., Kauffmann G., White S. D. M., 2009, *MNRAS*, 393, 377
- Martínez-Sansigre A., Rawlings S., 2011a, *MNRAS*, 414, 1937
- Martínez-Sansigre A., Rawlings S., 2011b, *MNRAS*, 418, L84
- McKinney J. C., Tchekhovskoy A., Blandford R. D., 2012, *MNRAS*, 423, 3083
- McLure R. J., Jarvis M. J., 2002, *MNRAS*, 337, 109
- McLure R. J., Jarvis M. J., 2004, *MNRAS*, 353, L45
- Mehdipour M., Costantini E., 2019, *A&A*, 625, A25
- Merloni A., Heinz S., 2008, *MNRAS*, 388, 1011
- Merloni A., Heinz S., di Matteo T., 2003, *MNRAS*, 345, 1057
- Mitchell J. A. J., Done C., Ward M. J., Kynoch D., Hagen S., Lusso E., Landt H., 2023, *MNRAS*, 524, 1796
- Mohan N., Rafferty D., 2015, Astrophysics Source Code Library, record ascl:1502.007
- Morabito L. K. et al., 2019, *A&A*, 622, A15
- Myers A. D. et al., 2015, *ApJS*, 221, 27
- Narayan R., Yi I., 1994, *ApJ*, 428, L13
- Narayan R., Yi I., 1995, *ApJ*, 444, 231
- Narayan R., Igumenshchev I. V., Abramowicz M. A., 2003, *PASJ*, 55, L69
- Päris I. et al., 2018, *A&A*, 613, A51
- Peterson B. M., 2014, *Space Sci. Rev.*, 183, 253
- Petley J. W. et al., 2022, *MNRAS*, 515, 5159
- Petley J. W. et al., 2024, *MNRAS*, 529, 1995
- Plotkin R. M., Markoff S., Kelly B. C., Körding E., Anderson S. F., 2012, *MNRAS*, 419, 267
- Rankine A. L., Hewett P. C., Banerji M., Richards G. T., 2020, *MNRAS*, 492, 4553
- Retana-Montenegro E., Röttgering H. J. A., 2017, *A&A*, 600, A97
- Reynolds C. S., Garofalo D., Begelman M. C., 2006, *ApJ*, 651, 1023
- Richards G. T. et al., 2006, *AJ*, 131, 2766
- Richards G. T. et al., 2011, *AJ*, 141, 167
- Richards G. T., McCaffrey T. V., Kimball A., Rankine A. L., Matthews J. H., Hewett P. C., Rivera A. B., 2021, *AJ*, 162, 270
- Roseboom I. G., Best P. N., 2014, *MNRAS*, 439, 1286
- Ross N. P. et al., 2009, *ApJ*, 697, 1634
- Ross N. P. et al., 2012, *ApJS*, 199, 3
- Schneider D. P. et al., 2010, *AJ*, 139, 2360
- Schulze A., Done C., Lu Y., Zhang F., Inoue Y., 2017, *ApJ*, 849, 4
- Seymour N. et al., 2007, *ApJS*, 171, 353
- Shen Y., 2013, *Bull. Astron. Soc. India*, 41, 61
- Shen Y. et al., 2009, *ApJ*, 697, 1656
- Shen Y. et al., 2011, *ApJS*, 194, 45
- Shen Y. et al., 2019, *ApJS*, 241, 34
- Shen Y. et al., 2024, *ApJS*, 272, 26
- Shimwell T. W. et al., 2017, *A&A*, 598, A104
- Shimwell T. W. et al., 2019, *A&A*, 622, A1
- Shimwell T. W. et al., 2022, *A&A*, 659, A1
- Sikora M., Begelman M. C., 2013, *ApJ*, 764, L24
- Smith D. J. B. et al., 2016, 271, preprint (arXiv:1611.02706)
- Smith D. J. B. et al., 2021, *A&A*, 648, A6
- Temple M. J. et al., 2023, *MNRAS*, 523, 646
- Ulvestad J. S., Ho L. C., 2001, *ApJ*, 558, 561
- van Haarlem M. P. et al., 2013, *A&A*, 556, A2
- Vestergaard M., Osmer P. S., 2009, *ApJ*, 699, 800
- Vestergaard M., Peterson B. M., 2006, *ApJ*, 641, 689
- White R. L., Helfand D. J., Becker R. H., Glikman E., de Vries W., 2007, *ApJ*, 654, 99
- Whittam I. H. et al., 2022, *MNRAS*, 516, 245
- Williams W. L. et al., 2019, *A&A*, 622, A2
- Wilson A. S., Colbert E. J. M., 1995, *ApJ*, 438, 62
- Wright E. L. et al., 2010, *AJ*, 140, 1868
- Wu Q., Shen Y., 2022, *ApJS*, 263, 42
- York D. G. et al., 2000, *AJ*, 120, 1579
- Yue B. H. et al., 2024, *MNRAS*, 529, 3939 (Y24)

## APPENDIX A: CHARACTERISTIC LUMINOSITIES OF SF- AND AGN-DOMINATED QUASARS

In Table A1, we present the values of the characteristic 150 MHz radio luminosities  $L_{\text{eq}}$  and  $L_{\text{pl}}$  (and their corresponding flux densities  $f_{\text{eq}}$  and  $f_{\text{pl}}$ , as evaluated from the midpoint of each redshift bin) defined in Section 4, for the  $\mathcal{M}_i - z$  grid cells included in this work. These values can be applied directly to classify quasars from other sources as long as their rest-frame  $i$ -band magnitudes and redshifts lie within the grid cells and 150 MHz (or equivalent) radio flux densities measurements are available. Following the instructions in Section 4, readers can also calculate  $L_{\text{eq}}$  and  $L_{\text{pl}}$  for the rest of the grid cells in Y24 (while not included in this work) from the model best fits presented in table A1 of Y24.

Since the definitions of  $L_{\text{eq}}$  and  $L_{\text{pl}}$  are independent of the quasar selection or the parameter space specific to this work, our classification method can be extended to other parameter spaces (defined with  $\mathcal{M}_i - z$  or other physical properties) as long as best-fitting of the two-component model in Y24 ( $\{\Psi, \sigma_{\psi}, \gamma, f\}$ ) can be obtained from a certain distribution of quasar radio flux densities.

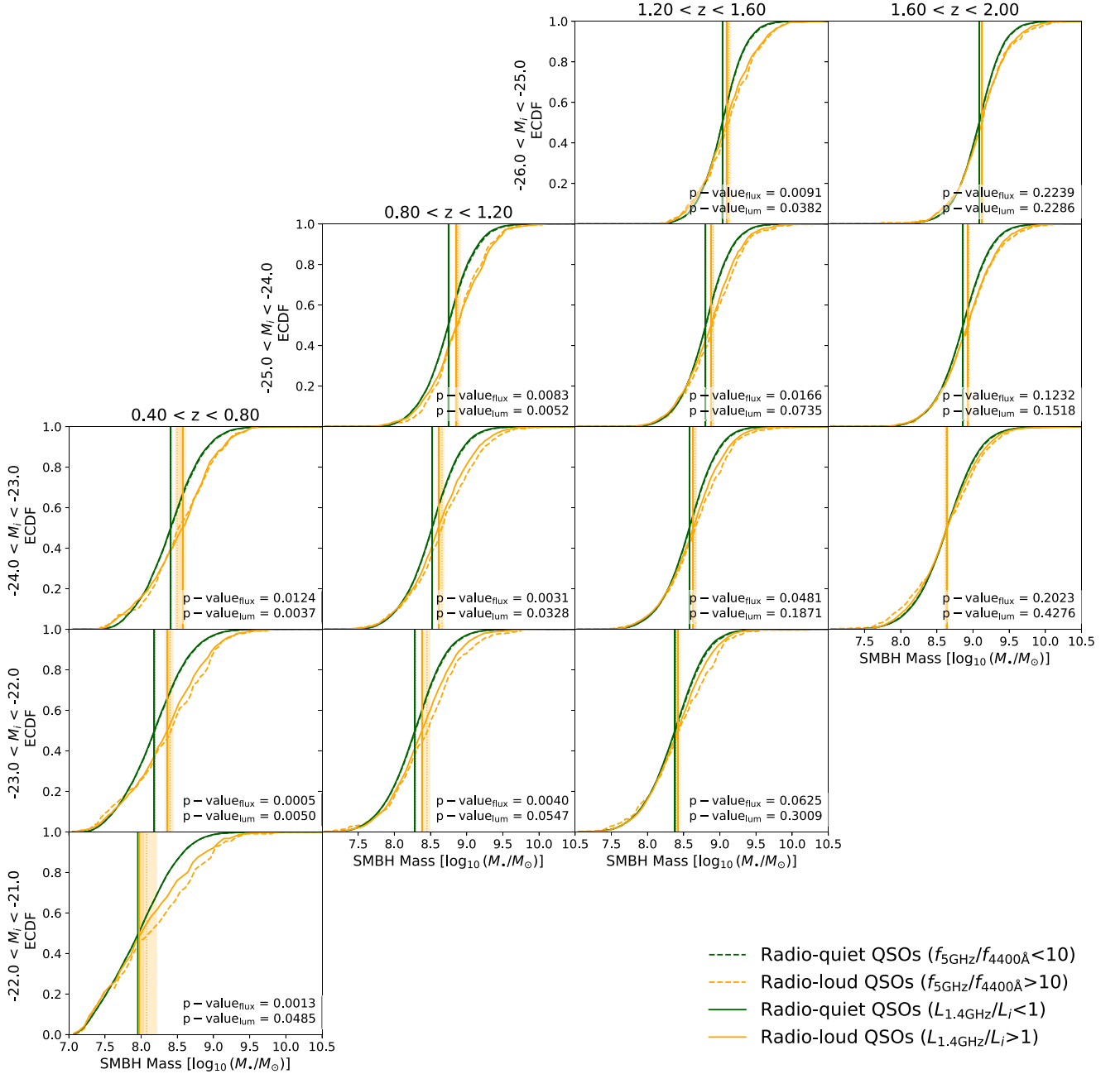
**Table A1.** The characteristic 150 MHz radio luminosity  $L_{\text{eq}}$  (upper luminosity threshold for SF-dominated quasars) and  $L_{\text{pl}}$  (lower luminosity threshold for AGN-dominated quasars) in the  $\mathcal{M}_i-z$  grid cells covered by this work. The  $\mathcal{M}_i$  and  $z$  values in this table represent the central point for each grid cell, while each grid cell spans 1 mag in rest frame  $i$ -band magnitude and 0.4 in redshift. The radio flux densities  $f_{\text{eq}}$  and  $f_{\text{pl}}$  are calculated from the mid-point redshift of each  $\mathcal{M}_i-z$  grid cell.

$\mathcal{M}_i$ ( $z = 0$ )	$z$	$\log L_{\text{eq}}$ ( $\text{W Hz}^{-1}$ )	$f_{\text{eq}}$ (mJy)	$\log L_{\text{pl}}$ ( $\text{W Hz}^{-1}$ )	$f_{\text{pl}}$ (mJy)
-21.5	0.6	23.86	0.541	24.19	1.152
-22.5	0.6	24.23	1.261	24.66	3.406
-23.5	0.6	24.41	1.890	24.98	7.104
-22.5	1.0	24.50	0.704	24.81	1.455
-23.5	1.0	24.65	0.998	25.07	2.610
-24.5	1.0	24.64	0.973	25.18	3.397
-22.5	1.4	24.66	0.463	24.95	0.898
-23.5	1.4	24.86	0.734	25.22	1.700
-24.5	1.4	24.94	0.891	25.38	2.400
-25.5	1.4	25.02	1.051	25.55	3.639
-23.5	1.8	24.97	0.525	25.23	1.104
-24.5	1.8	25.16	0.815	25.55	2.017
-25.5	1.8	25.23	0.941	25.70	2.825

## APPENDIX B: ADDITIONAL PLOT FOR BLACK HOLE MASS DISTRIBUTIONS

Fig. B1 shows the cumulative BH mass distribution across the  $\mathcal{M}_i-z$  parameter space for RL and RQ quasar populations defined in the previous literature, using  $R_{\text{flux}} = f_{5\text{GHz}}/f_{4400\text{\AA}}$  (e.g. McLure & Jarvis 2004) and  $R_{\text{lum}} = \log_{10}(L_{1.4\text{GHz}}/L_i)$  (e.g. Baloković et al. 2012; Arnaudova et al. 2024). Based on both criteria, the BH mass distribution differs between the RL and RQ samples, where the RL quasars tend to be more massive ( $p$ -value  $< 0.05$ ) in some grids; however, there is no unified trend in such a difference, suggesting that radio loudness alone is not enough to characterize the relationship between quasar radio emission and  $M_{\text{BH}}$ . According to the criterion used by Arnaudova et al. (2024) where quasars with  $R_{\text{lum}} < 1$  are classified as RQ quasars and quasars with  $R_{\text{lum}} > 1$  are classified as RL (dotted lines), there is little difference ( $p$ -value  $> 0.05$ ) between the BH mass distributions of RL and RQ quasars in many of the more-populated lower luminosity grid cells, which leads to the disagreement between their conclusion and the result in McLure & Jarvis (2004). For the less populated grid cells in the higher luminosity regime, both criteria agree with each other due to the different host galaxy SFR and AGN jet normalization in those parameter space.





**Figure B1.** Cumulative mass distribution within each  $\mathcal{M}_i - z$  grid cell, separated by the radio-loudness of quasars. The grid cells are arranged by their average  $i$ -band luminosity and redshift, following the pattern in Fig. 7. For the distribution drawn in solid lines, the RL/RQ classification is based on the values of  $R_{\text{lum}} = \log_{10}(L_{1.4\text{GHz}}/L_i)$ : Quasars with  $R_{\text{lum}} < 1$  are classified as RQ quasars, while quasars with  $R_{\text{lum}} > 1$  are classified as RL. The distribution in dashed lines reflects the RL and RQ quasar population defined with  $R_{\text{flux}} = f_{5\text{GHz}}/f_{4400\text{\AA}}$ , where the RQ quasars have  $R_{\text{flux}} < 10$  and the RL quasars have  $R_{\text{flux}} > 10$ . Within each grid cell, we include the  $p$ -values calculated from K-S tests on the BH mass CDFs of RL and RQ quasars under different definitions.

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