



Universiteit
Leiden
The Netherlands

Faint quasars at very low frequencies

Retana Montenegro, E.F.

Citation

Retana Montenegro, E. F. (2019, October 16). *Faint quasars at very low frequencies*. Retrieved from <https://hdl.handle.net/1887/79263>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/79263>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The following handle holds various files of this Leiden University dissertation:
<http://hdl.handle.net/1887/79263>

Author: Retana Montenegro, E.F.

Title: Faint quasars at very low frequencies

Issue Date: 2019-10-16

Chapter 5

The luminosity function of LOFAR radio-selected quasars at $1.4 \leq z \leq 5.0$ in the NDWFS-Boötes field

Abstract

We present an estimate of the luminosity function (LF) of LOFAR radio-selected quasars at $1.4 < z < 5.0$ in the 9.3 deg^2 NOAO Deep Wide-field survey (NDWFS) of the Boötes field. Selection is based on optical and mid-infrared photometry used to train three different machine learning (ML) algorithms (Random forest, SVM, Bootstrap aggregation). Objects identified as quasars by the ML algorithms are required to be detected at 5σ significance in deep radio maps. Optical imaging comes from the Sloan Digital Sky Survey and the Pan-STARRS1 3π survey, mid-infrared photometry is taken from the Spitzer Deep, Wide-Field Survey, and radio data is obtained from deep LOFAR

Retana-Montenegro and Röttgering, submitted to A&A, (2019)

imaging of the NDWFS-Boötes field. The sample comprises 134 objects, including both photometrically-selected candidate quasars (51) and spectroscopically confirmed quasars (83). The depth of our LOFAR observations allow us to detect the radio-emission of quasars that otherwise would be classified as radio-quiet. Around 65% of the quasars in the sample are fainter than $M_{1450} < -24.0$, a regime where the luminosity function of quasars selected through their radio emission (radio-selected quasars, RSQs) has not been investigated in detail. It is demonstrated that in cases where mid-infrared wedge-based AGN selection is not possible due to a lack of appropriate data, the selection of quasars using ML algorithms trained with optical/infrared photometry in combination with LOFAR data provides an excellent approach for obtaining samples of quasars. The LF of RSQs can be described by pure luminosity evolution at $z < 2.4$, and a combined luminosity and density evolution at $z > 2.4$. The faint-end slope, α , becomes steeper with increasing redshift. This trend is consistent with previous studies of faint quasars ($M_{1450} < -22.0$). We demonstrate that RSQs show an evolution that is very similar to the exhibited by faint quasars. We find evidence that supports a scenario where RSQs produce only a few per cent of the photons required to ionize the intergalactic medium at $z > 3$. By comparing the spatial density of RSQs with that of the *total* (radio-detected plus radio-undetected) faint quasar population at similar redshifts, we find that RSQs may compose up to 30% of the whole faint quasar population. This fraction within uncertainties is constant with redshift. Finally, we discuss how the compactness of the RSQs radio-morphologies and their steep spectral indices could provide valuable insights how quasar and radio activity are triggered in these systems.

5.1 Introduction

A good determination of the quasar luminosity function (QLF) is important to understand several aspects of the cosmological evolution of supermassive black holes (SMBHs). These aspects include: i) the build-up of black hole (BH) demography; ii) the integrated UV contribution from quasars to the ionization of the intergalactic medium; iii) the accretion history of BHs across cosmic time; iv) triggering and fueling

quasar mechanisms and their co-evolution with host galaxies.

The cosmological evolution of quasars has been studied in detail over a wide range of optical luminosities at $z < 3$ (e.g. Richards et al. 2006; Croom et al. 2009b; Ross et al. 2013). These studies show that the comoving space density of quasars evolves strongly, with lower-luminosity quasars peaking in their space density at lower redshift than higher-luminosity quasars. This result is interpreted as a downsizing evolutionary scenario for SMBHs in which very massive BHs were already in place at very early times, whereas less massive BHs evolve predominantly at lower redshifts. These results provide valuable benchmarks to BH formation models (e.g. Volonteri & Rees 2005; Somerville et al. 2008).

At $z > 3$, the shallow flux limits of current quasar surveys restricts our understanding of BH growth to the brightest optical objects. As a result, there is little information on how faint quasars evolve. The BH downsizing behavior in the early universe is not well understood, and the role of faint quasars in the cosmic reionization of hydrogen remains poorly constrained. While studies considering only the brightest quasars found that their contribution to cosmic reionization is not significant (e.g. Haardt & Madau 2012), other authors which take into account faint quasars claim that potentially they can produce the high emissivity rate required to ionize the intergalactic medium (e.g. Glikman et al. 2011; Giallongo et al. 2015). For a good picture of the quasar phenomena at high- z , it is important to study significant numbers of these low-luminosity objects at high redshifts.

The selection of quasars at $z > 2.2$ is challenging, particularly for low-luminosity quasars with optical magnitudes close to the flux limit. In this regime, the photometric errors broaden the stellar locus, and color distributions of quasars and stars are hard to distinguish using color selection. One approach to circumvent this difficulty, is to build quasar samples using optical/infrared color selection combined with a radio detection. The main advantages of radio selection over typical color selection is that stellar contamination is reduced significantly as radio stars are very rare (Kimball et al. 2009). The application of this technique has therefore led to the discovery of quasars outside the typical color boxes used to select them (see Hook et al. 2002; McGreer et al. 2009;

Zeimann et al. 2011; Bañados et al. 2015).

One caveat of using radio selection is that the quasars selected may not be representative of the entire quasar demographics. In fact, the majority of studies of the luminosity function of radio-selected quasars (RSQs) (Shaver et al. 1996; Vigotti et al. 2003; McGreer et al. 2009; Carballo et al. 2006; Tuccillo et al. 2015) include only radio-loud quasars (RLQs) with luminosities $L_{1.4GHz} \gtrsim 1 \times 10^{26}$ W/Hz, that are selected using shallow all-sky radio and optical surveys such as FIRST (Becker et al. 1995) and SDSS (York et al. 2000), respectively. Interestingly, the works presented by McGreer et al. (2009) and Tuccillo et al. (2015) found that the luminosity function of RSQs shows a flattening of the bright-end that is similar to the whole quasar population at $3.5 < z < 4.4$ (Richards et al. 2006). Moreover, the analysis by Cirasuolo et al. (2005) suggests a decrement of the space density of faint RSQs from $z \simeq 1.8$ to $z = 2.2$ by a factor of 2. An issue is the origin of the radio emission in radio-quiet quasars (RQQs). It is still a matter of debate whether it is linked to star-forming activity occurring in the host galaxy (Kimball et al. 2011; Padovani et al. 2011; Condon et al. 2013; Bonzini et al. 2013) or non-thermal processes near the SMBH (Prandoni et al. 2010; Herrera Ruiz et al. 2016).

The number of known quasars has increased dramatically in the course of the last two decades (e.g. Croom et al. (2009a); Pâris et al. (2018)). This number will continue increasing with forthcoming facilities such as the Large Synoptic Survey Telescope (LSST, Ivezić et al. 2019), WFIRST (Spergel et al. 2015), and Euclid (Laureijs et al. 2011), expected to deliver millions of quasars (e.g. Ivezić et al. 2014). One of the major challenges is the identification of quasars without spectroscopic observations, which are costly in terms of telescope time for such large samples. Several machine learning (ML) techniques have been proposed to photometrically create large samples of quasars, including artificial neural networks (Yèche et al. 2010; Tuccillo et al. 2015), random forest (Schindler et al. 2017), support vector machine (Gao et al. 2008; Peng et al. 2012; Jin et al. 2019), extreme deconvolution (Bovy et al. 2011; Myers et al. 2015), bayesian selection (Kirkpatrick et al. 2011; Peters et al. 2015; Yang et al. 2017), and bootstrap aggregation (Timlin et al. 2018).

Many of these quasars will be detected by the next generation of radio-surveys (Norris 2017). Particularly, low-frequency radio telescopes such as the Low Frequency Array (LOFAR, van Haarlem et al. 2013) open a new observational spectral window to study the evolution of quasar activity. The LOFAR Surveys Key Science Project (LSKSP, Röttgering et al. 2011) aims to map the entire Northern Sky down to $\lesssim 100 \mu\text{Jy}$, while for extragalactic fields, greater than a few square degrees in size and with extensive multi-wavelength data, the target rms noise is of a few tens of μJy . In this paper, we investigate the evolution of the luminosity function of RSQs. For this purpose, we take advantage of the deep optical, infrared and LOFAR data available for the NOAO Deep Wide-field survey (NDWFS) Boötes field (Jannuzi & Dey 1999).

This paper is organized as follows. In Section 5.2, we present the surveys used in this work. In Section 5.3, we briefly discuss the classification algorithms employed to compile our quasar sample. We explain the method utilized to compute the photometric redshifts for our photometrically-selected candidate quasars in Section 5.4. In Section 5.4.5, we compare the LOFAR and wedge-based mid-infrared selection for objects classified as quasars by the ML classification algorithms. In Section 5.5, we present the luminosity function of our RSQs, and compare our results with previous works from the literature in Sections 5.6.1 and 5.6.2. The comoving spatial density of RSQs is studied in Section 5.6.3. The contribution of RSQs to photoionization of the intergalactic medium (IGM) is discussed in Section 5.6.4. Section 5.7 discusses how the compactness of the RSQs radio-morphologies and their steep spectral indices could provide insights into how quasar and radio activity are triggered. Section 5.7.3 discusses the possible location of RSQs in different spectroscopic parameter spaces. Finally, we summarize our conclusions in Section 5.8. For simplicity, we henceforth refer to photometrically-selected candidate quasars as photometric quasars and spectroscopically confirmed quasars as spectroscopic quasars. Also, we refer to published samples of quasars with $M_{1450} > -22.0$ (Siana et al. 2008; Glikman et al. 2011; Masters et al. 2015; Yang et al. 2018; McGreer et al. 2018; Akiyama et al. 2018) as faint quasars. Through this paper, we use a Λ cosmology with the matter density $\Omega_m = 0.30$, the cosmological constant $\Omega_\Lambda = 0.70$, the Hubble constant $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. We

assume a definition of the form $S_\nu \propto \nu^{-\alpha}$, where S_ν is the source flux, ν the observing frequency, and α the spectral index. To estimate radio and optical luminosities, we adopt optical and radio indices $\epsilon_{\text{opt}} = 0.5$ and $\alpha = 0.7$, respectively. All the magnitudes are expressed in the AB magnitude system (Oke & Gunn 1983).

5.2 Data

In this section, we introduce the datasets that will be utilized for the selection of quasars, and for the estimation of photometric redshifts for objects without spectroscopy.

5.2.1 NOAO Deep Wide-field survey

The NOAO Deep Wide-field survey (NDWFS) is a deep imaging survey that covers approximately two 9.3 deg^2 fields (Jannuzi & Dey 1999). One of these regions, the Boötes field has a large wealth of data available at a range of observing windows including: X-ray (*Chandra*; Kenter et al. 2005), UV/optical ($NUV, U_{\text{spec}}, B_W, R, I, Z_{\text{Subaru}}$ bands; Jannuzi & Dey 1999; Martin et al. 2005; Cool 2007; Bian et al. 2013), infrared (Y, J, H, K, K_s bands, *Spitzer*; Ashby et al. 2009; Jannuzi et al. 2010), and radio (150-1400MHz; de Vries et al. 2002; Williams et al. 2013, 2016; Retana-Montenegro et al. 2018a). We use the *Spitzer*/IRAC- $3.6 \mu\text{m}$ matched photometry catalog presented by Ashby et al. 2009. This catalog contains 677522 sources detected at 5σ limiting magnitudes measured in a $4''$ diameter (aperture-corrected) of 22.56, 22.08, 20.24, and 20.19 at 3.6, 4.5, 5.8, and $8.0 \mu\text{m}$, respectively. The $3.6 \mu\text{m}$ and $4.5 \mu\text{m}$ magnitudes are converted to AB units using the relations: $[3.6 \mu\text{m}]_{AB} = [3.6 \mu\text{m}]_{Vega} + 2.788$ and $[4.5 \mu\text{m}]_{AB} = [4.5 \mu\text{m}]_{Vega} + 3.255^2$. To select our RSQs, we use the deep 150MHz LOFAR observations of the Boötes field presented by Retana-Montenegro et al. (2018a). The image obtained covers more than 20 deg^2 and was based on 55 hours of observations. The central rms noise of the mosaic is $55 \mu\text{Jy}$ with an angular resolution of $3.98'' \times 6.45''$. The final radio catalog contains 10091 sources detected above a 5σ peak flux density threshold. There are 170 extended sources in the catalog, whose com-

²http://irsa.ipac.caltech.edu/data/COSMOS/gator_docs/scosmos_irac_colDescriptions.html

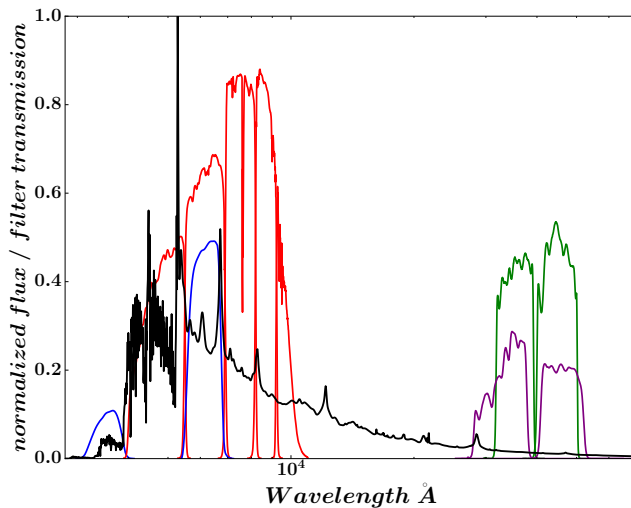


Figure 5.1: The transmission curves of the filters used in this work. *Blue lines*: SDSS-*u* and SDSS-*r* filters; *red lines*: the Pan-STARRS1 filter set: $g_{\text{PS}}, r_{\text{PS}}, i_{\text{PS}}, z_{\text{PS}}, y_{\text{PS}}$; *green lines*: *Spitzer-IRAC* $[3.6 \mu\text{m}]$ and $[4.5 \mu\text{m}]$ bands; *purple lines*: WISE W1 and W2 bands; **and** the solid black line shows a simulated quasar spectrum from our library at $z = 3.4$ (See Section 5.5.2).

ponents were merged according to a visual inspection. This reduces the possibility of missing sources without detected cores in the LOFAR mosaic. A total of 5646 LOFAR sources are found in the *Spitzer/IRAC*- $3.6 \mu\text{m}$ matched catalog using a matching radius of $2''$.

5.2.2 SDSS, Pan-STARRS1, WISE, and Spitzer surveys

The Sloan Digital Sky Survey (SDSS, York et al. 2000) is a multi-filter imaging and spectroscopic survey conducted with the 2.5m wide-field Sloan telescope (Gunn et al. 2006) located at the Apache Point observatory in New Mexico, USA. The SDSS-DR14 (Abolfathi et al. 2018) provides photometry for 14955 deg^2 in five broad-band optical filters (u, g, r, i, z ; Fukugita et al. 1996). The magnitude limits (95% completeness for point sources) in the five filters are $u = 22.0$, $g = 22.2$, $r = 22.2$, $i = 21.3$, and $z = 20.5$ mag, respectively.

We also use optical/near-infrared imaging from the 1.8m Pan-STARRS1 telescope

(Hodapp et al. 2004) located on the summit of Haleakala on the Hawaiian island of Maui, which provides five band photometry ($g_{\text{PS}}, r_{\text{PS}}, i_{\text{PS}}, z_{\text{PS}}, y_{\text{PS}}$). The Pan-STARRS1 first and second data releases (Chambers et al. 2016) are dedicated to the 3π survey, which observed, for almost four years, the sky north of -30° declination reaching 5σ limiting magnitudes in the $g_{\text{PS}}, r_{\text{PS}}, i_{\text{PS}}, z_{\text{PS}}, y_{\text{PS}}$ bands of 23.3, 23.2, 23.1, 22.3, 21.3, respectively. Pan-STARRS1 provides deeper imaging in overlapping optical bands (except the SDSS $-u$ band) and has the near-IR filter y_{PS} . SDSS has the u band covering wavelengths between 3000-4000 Å, which contains the Lyman alpha emission at redshifts $1.3 \lesssim z \lesssim 2.2$. This makes the SDSS $-u$ band important for the selection of $z \lesssim 2.2$ quasars. For these reasons, we combine the SDSS $-u$ band with the Pan-STARRS1 filter set ($g_{\text{PS}}, r_{\text{PS}}, i_{\text{PS}}, z_{\text{PS}}, y_{\text{PS}}$) to have wavelength coverage from 3000 Å to 10800 Å (see Figure 5.1).

As a first step to obtaining mid-infrared photometry for the spectroscopic quasars, we combine the observations from all deep Spitzer-IRAC surveys including: XFLS (Lacy et al. 2005), SERVS (Mauduit et al. 2012), SWIRE (Lonsdale et al. 2003), S-COSMOS (Sanders et al. 2007), SDWFS (Ashby et al. 2009), SHELA (Papovich et al. 2016), and SpIES (Timlin et al. 2016). We follow the same procedure described by Richards et al. (2015) to combine all the Spitzer-IRAC observations. The final catalog contains over 6.2 million Spitzer-IRAC sources. In cases where an IRAC source has been observed multiple times, we use only the deepest IRAC observation.

The mid-infrared photometry for spectroscopic quasars outside the footprint of Spitzer-IRAC surveys comes from observations by NASA's Wide Infrared Survey Explorer (WISE, Wright et al. 2010). WISE mapped the sky at 3.4, 4.6, 12 and 22 μm (known as W1, W2, W3, and W4). After the cryogenic fuel of the satellite was exhausted in 2010, WISE continued its observations as part of the post-cryogenic NEOWISE mission phase using only its two shortest bands (W1 and W2). We use the SDSS/unWISE forced photometry catalog by Lang (2014). This catalog provides forced photometry of custom WISE coadds, at the positions of over 400 million SDSS primary sources. This approach provides WISE flux measurements for sources blended in WISE coadds, but resolved in SDSS images and non-detected objects below the "official" WISE magnitude

limits (i.e. ALLWISE; Cutri 2013). We use only the W1 and W2 bands, as the other two bands are shallower, and thus have lower detection rates.

We retrieve the SDSS, Pan-STARRS1, and WISE photometry from the SDSS database via CasJobs³ and the Mikulski Archive for Space Telescopes (MAST) with PS1 CasJobs⁴. We make sure that the objects in our samples have clean photometry, by excluding sources with the SDSS bad photometry flags described in Richards et al. (2015). However, we opt to keep objects with the flag `BLENDED`, as high- z quasars could be flagged as `BLENDED` in some instances despite being isolated objects (e.g. McGreer et al. 2009). Only `PRIMARY` sources are selected from the SDSS data. The flags that describe the quality of the Pan-STARRS1 sources are taken from Table 2 in Magnier et al. (2016). For SDSS and Pan-STARRS1, we use PSF magnitudes, and adopt the `w1mag` and `w2mag` columns from the unWISE catalog as the WISE measurements for the W1 and W2 bands, respectively. These WISE magnitudes are converted to AB units using the relations: $W1_{AB} = W1_{Vega} + 2.699$ and $W2_{AB} = W2_{Vega} + 3.339^5$. We consider only WISE sources meeting the following criteria: $w1_prochi2 \leq 2.0$ & $w2_prochi2 \leq 2.0$ (to avoid sources with low-quality profile fittings), and $w1_profracflux \leq 0.1$ & $w1_profracflux \leq 0.1$ (to exclude sources with fluxes severely affected by bright neighbors). The SDSS `cMODEL`⁶ magnitudes are also retrieved to investigate the separation of point and extended sources (Section 5.4.4). The SDSS magnitudes in the u filter, originally in inverse hyperbolic sine magnitudes (Lupton et al. 1999), are converted to the AB system using $u_{AB} = u_{SDSS} - 0.04$ (Fukugita et al. 1996). The WISE-W1 and WISE-W2 photometry is converted to the IRAC 3.6 μm and 4.5 μm bands, respectively, using the transformations derived by Richards et al. (2015). We crossmatch the WISE and Spitzer-IRAC catalogs using a radius of $2''$. If the crossmatch is positive, we keep only the IRAC measurement. The SDSS, Pan-STARRS1 and IRAC magnitudes are corrected for Galactic extinction using the prescription by Schlafly & Finkbeiner (2011). Figure 5.1 shows the transmission curves of the filters utilized in this work.

³<http://skyserver.sdss.org/CasJobs/>

⁴<https://mastweb.stsci.edu/ps1casjobs/>

⁵<http://wise2.ipac.caltech.edu/docs/release/allsky/expsup/>

⁶<https://www.sdss.org/dr14/algorithms/magnitudes/>

5.2.3 Spectroscopic quasars with optical and mid-infrared photometry

To efficiently discover new quasars using ML techniques, requires the compilation of a large and representative sample of spectroscopic quasars (e.g. (Richards et al. 2015; Pasquet-Itam & Pasquet 2018; Jin et al. 2019)). For this purpose, we use the Million Quasars (Milliquas) catalog v6.2 2019⁷ by Flesch (2015). This catalog contains more than 600000 type-I quasars and active galactic nuclei (AGN) from the literature, and is updated on a regular basis. The majority of quasars included in the Milliquas catalog were discovered as part of SDSS/BOSS (Schneider et al. 2010; Pâris et al. 2018), LAMOST (Yao et al. 2019), ELQS (Schindler et al. 2017), 2QZ (Croom et al. 2004), 2SLAQ (Croom et al. 2009a), and many other surveys (e.g. Papovich et al. 2006; Trump et al. 2009; Kochanek et al. 2012; Maddox et al. 2012; McGreer et al. 2013). We consider only quasars with magnitudes measured for each band, to maximize the use of the multi-dimensional color information available.

5.3 Classification

In this section, we explain how the training and target (objects to be classified) samples are compiled, and the different algorithms used for the classification of quasars in the NDWFS-Boötes field. We also assess the performance of the classification algorithms by calculating their efficiency and completeness.

5.3.1 Training sample

A critical success factor for any ML technique to classify astronomical sources is the use of an appropriate training sample to identify new objects in the target sample. The training sample must have measurements in the relevant filters to identify the characteristic spectral features (e.g. colors) of the sources of interest (e.g. quasars) in order to map their parameter space. At the same time, the training sample has to be representative of the target data. This means not only including a significant number of

⁷<http://quasars.org/>

the sources of interest, but also the other types of astronomical objects expected to be part of the target sample (i.e. stars and galaxies). In particular for quasars, the training samples require several thousands of these objects to robustly extract their color trends as a function of redshift (Yèche et al. 2010; Richards et al. 2015; Nakoneczny et al. 2019; Pasquet et al. 2019). Unfortunately, there are only 2042 quasars with $0.126 < z < 6.12$ in the NDWFS-Boötes field, with only 1259 of these quasars having redshifts larger than 1.4. The Boötes quasars in this catalog are drawn mainly from the AGN and Galaxy Evolution Survey (AGES, Kochanek et al. 2012), but other quasar surveys (Cool et al. 2006; McGreer et al. 2006; Glikman et al. 2011; Pâris et al. 2018) have been used as well. These objects are included in the Milliquas catalog (Flesch 2015). However, this sample is too small and sparse, to properly map the parameter space of quasars in the NDWFS-Boötes field. Instead, of relying only on the NDWFS-Boötes quasar sample to identify new quasars, we create a training sample using as starting point the Milliquas catalog presented in Section 5.2.3.

We restrict the redshift range of our analysis to $1.4 < z < 5.0$ for the following reasons. First, the host galaxies of some $z < 1$ quasars can be detected in the NDWFS images. This implies that the contribution of the host galaxy component to the overall quasar spectra has to be considered for these sources. This makes it difficult to separate low- z galaxies and quasars using morphological classification based on standard photometric criteria. Secondly, low- z contaminants such as star-forming, blue, and emission-line galaxies can mimic the colors of high- z quasars due to their 4000 Å breaks (Smith et al. 1993; Richards et al. 2002). Therefore, we set $z = 1.4$ as the limit for the lower redshift interval in our analysis. This choice is a compromise between reducing contamination by galaxies and excluding good candidate quasars from the sample, but ensures that the degree of contamination due to galaxies across the redshift interval considered is as low as possible. Thirdly, at redshifts $z > 5$ the number of known quasars is significantly low compared with lower redshifts. Thus, we set $z = 5.0$ as the upper redshift limit for our analysis. In the training sample, we include spectroscopic quasars with redshifts slightly larger than the redshift limits of our analysis. This helps to reduce the possibility of quasars located at the edges of the redshift intervals not

being identified by the classification algorithms. Therefore, we include spectroscopic quasars with $1.2 < z < 5.5$ in our training sample.

The first step to create our training sample is to crossmatch the entire Milliquas catalog with the SDSS, Pan-STARRS1, WISE, and IRAC surveys as described in Section 5.2.2. We make sure that the quasars in the training sample have clean photometry, by excluding quasars with SDSS, IRAC, and WISE bad photometry flags (see Section 5.2.2). We limit the spectroscopic quasar sample to magnitudes $16.0 \leq i_{PS} \leq 23.0$ as this range contains 99.9% of all optical/mid-infrared quasars in the training sample. In total, our sample contains 328956 spectroscopic quasars with $1.2 < z < 5.5$. These quasars have clear photometry and are detected in the optical/near-infrared ($u, g_{PS}, r_{PS}, i_{PS}, z_{PS}, y_{PS}$), and mid-infrared ($3.6 \mu\text{m}, 4.5 \mu\text{m}$) bands considered in our analysis. These objects are assigned the label “quasar” in our training sample. This label assignment could be seen as trivial, but it is fundamental because the algorithms introduced in Section 5.3.3 require labels to categorize new unlabeled data in the target sample.

The rest of the training sample that comprises non-quasar objects is compiled as follows. First, we need to consider that the target sample does not only contain $1.2 \leq z \leq 5.5$ quasars, but ones with redshifts lower than 1.2 as well. Thus, to ensure that the training sample is representative, we include quasars with redshifts $0 \leq z < 1.2$ from the Milliquas catalog. There is a total of 71830 quasars that are selected as described in Section 5.2.2, and are assigned the “non-quasar” label. Secondly, stars and galaxies are expected to be part of the target sample. We include them in the training sample by randomly drawing objects with clean photometry (see Section 5.2.2) from the SDSS database with CasJobs, and excluding sources located in the NDWFS-Boötes region. These objects have $m_{AB} < 15$ in all the bands to avoid saturated pixels, and must have been classified spectroscopically as stars or galaxies by the SDSS pipeline. If the source is matched within a radius of $2''$ to a known spectroscopic quasar it is excluded. We do not apply any morphological criteria to the sources added to the training sample. The drawing process is repeated until a total of 1098858 “non-quasar” objects are selected. This size is chosen to have, in combination with the spectroscopic quasar sample, a

Table 5.1: The performance of the classification algorithms for the quasar training sample. The statistics are indicated for each subsample.

Algorithm	Completeness	Efficiency
	%	%
Random Forest	87.83	96.69
Support Vector Machines	84.54	94.31
Bootstrap Aggregation on K-Nearest Neighbors	83.66	94.81

Notes: All experiments used the same training and test samples.

Table 5.2: Properties of the training and target samples.

Sample	Number of objects	i_{PS} magnitude
		[AB]
Training sample	5.0×10^6	$15 \leq i_{\text{PS}} \leq 24$
$1.2 \leq z \leq 5.5$ quasars	328956	$16 \leq i_{\text{PS}} \leq 23$
$0 < z < 1.2$ quasars	71830	$16 \leq i_{\text{PS}} \leq 23$
Target sample	287218	$15 \leq i_{\text{PS}} \leq 24$

total of a million and a half of objects in the training sample. The inclusion of galaxies, stars, and $z < 1.2$ quasars in the training sample is important because it helps the classification algorithms to delimit the color space of “quasars” and “non-quasars”. The details of the training sample are summarized in Table 5.2.

5.3.2 Target sample

The target sample is the Spitzer/IRAC- $3.6\mu\text{m}$ matched catalog presented in Section 5.2.1. As mentioned in Section 5.3.1, there is deep optical photometry available for the NDWFS-Boötes field in the $U_{\text{spec}}, B_w, R, I, Z_{\text{Subaru}}$ bands, however, the number of quasars with photometry in these bands is small. Fortunately, the entire NDWFS-Boötes field is located inside the SDSS/Pan-STARRS1 footprint, and thus photometry from these surveys is available for the target sample. We obtain SDSS and Pan-STARRS1 photometry for all the objects in the NDWFS-Boötes catalog following the same procedure as for the training set. We remove sources with $m_{AB} < 15$ in all the bands (SDSS, Pan-STARRS1, IRAC) regardless of their flags to avoid saturated pixels. We keep only sources with the SDSS/Pan-STARRS1 good photometry flags. We have also considered using the IRAC flags (SExtractor flags indicating possible blending issues in the source

extraction) but found that around 60% of the $z > 1.4$ spectroscopic quasars in Boötes are flagged. Considering that the removal of such a significant fraction of quasars from the analysis could affect our conclusions, we decide not to remove flagged IRAC sources from the target sample at this point. In Sections 5.3.3.5 and 5.4.3, we confirm that including sources marked by the IRAC flags does not result in a deterioration of the quality of the classification, and/or determination of the photometric redshifts. Finally, the details of the target sample are summarized in Table 5.2.

5.3.3 Classification algorithms

In supervised ML, classification algorithms rely on labeled input data (training sample) to produce an inferred function, which can be used to categorize new unlabelled data (target sample). If there is a strong correlation between the input data and the labels a robust inferred function can be obtained. This usually results in a better performance of the ML classification algorithms. In this work, our aim is to identify new quasars in the NDWFS-Boötes field in the target sample using supervised ML classification algorithms. For quasars, the obvious choice is to use their colors for classification purposes (e.g. Richards et al. 2002, 2009, 2015; Timlin et al. 2018). Therefore, we use the color indices ($u - g_{\text{PS}}$, $u - r_{\text{PS}}$, $g_{\text{PS}} - r_{\text{PS}}$, $r_{\text{PS}} - i_{\text{PS}}$, $i_{\text{PS}} - z_{\text{PS}}$, $z_{\text{PS}} - y_{\text{PS}}$, $y_{\text{PS}} - [3.6 \mu\text{m}]$, $[3.6 \mu\text{m}] - [4.5 \mu\text{m}]$) of the training and target samples as inputs to the classification algorithms. The algorithms used in our analysis (Random Forest, Support vector machine, and Bootstrap aggregation) are selected because of their extensive use in previous studies of quasars (e.g. Gao et al. 2008; Peng et al. 2012; Carrasco et al. 2015; Timlin et al. 2018; Jin et al. 2019). Each one of these algorithms is briefly explained below, and are part of the open-source scikit-learn⁸ Python library.

5.3.3.1 Random forest

A random forest (RF, Breiman 2001) ensemble is composed of random decision trees, with each one created from a random subset of the data. The outputs of the decision trees are combined to make a consensus prediction. The final RF classification of an

⁸<https://scikit-learn.org>

unlabeled instance is determined using the majority vote of all decision trees.

5.3.3.2 Support vector machines

The support vector machines (SVM, Cortes & Vapnik 1995) is a discriminative classifier for two-group problems. The basic idea is to find an optimal hyperplane in an N -dimensional space (N -the number of features) that distinctly categorizes the data points. In two dimensions, the hyperplane is a line dividing the parameter space in two parts wherein each group is located on either side. For unlabeled instances, the SVM classifier outputs an optimal hyperplane which is used to classify them.

5.3.3.3 Bootstrap aggregation on K-nearest neighbors

Bootstrap Aggregation (Breiman 1996), also called bagging, is a method for generating multiple versions from a training set, by sampling the original sample uniformly and with replacement. Subsequently, each one of these subsets is used to train the K-Nearest Neighbors algorithm (KNN, Altman 1992). In the KNN algorithm, an unlabelled object is classified by a simple majority vote of its neighbors, with the object being assigned the label most common among its k nearest neighbors (k is a positive integer). In the case $k = 1$, the label assigned is of that of the single nearest neighbor. For each bagging subset, we use a value of $k = 50$. Finally, the results of the bagging subsets are aggregated by averaging to obtain a final classification.

5.3.3.4 Performance

We assess the performance of the classification algorithms with the quasar training sample presented in Section 5.3.1, by calculating their completeness and efficiency.

The completeness C (Hatziminaoglou et al. 2000; Retana-Montenegro & Röttgering 2018) is defined as the ratio between the number of quasars correctly identified as quasars, and the total number of quasars in the sample:

$$C = \frac{\text{Number of identified quasars}}{\text{Total number of quasars}} \times 100. \quad (5.1)$$

Table 5.3: The performance of the classification algorithms for the quasar training sample. The statistics are indicated for each subsample.

Algorithm	Completeness	Efficiency
	%	%
Random Forest	93.09	90.50
Support Vector Machines	89.36	81.63
Bootstrap Aggregation on K-Nearest Neighbors	93.36	85.97

Notes: All experiments used the same training and test samples.

The efficiency E is defined as the ratio between the number of quasars correctly identified as quasars, and the total number of objects identified as quasars:

$$E = \frac{\text{Number of identified quasars}}{\text{Total no. of objects identified as quasars}} \times 100. \quad (5.2)$$

As is common practice in supervised ML, the training sample described in Section 5.3.1 is separated into validation and test samples for cross-validation (CV) purposes, in order to calculate the performance on observations of the classification algorithms. The validation set is used as the training sample, while the CV test sample, which contains spectroscopic quasars, is employed to report the completeness and efficiency of the classification algorithms. As the CV test set, we choose a random 25% subsample of the full training set. The remaining 75% of the training data is the validation sample, and is used to train the algorithms.

In Table 5.3, we present the completeness and efficiency for all the classification algorithms. While the differences are small, performance of the SVM algorithm is the worst, while RF has the highest completeness and efficiency values.

5.3.3.5 Classification results

In this section, we discuss the results of the application of the classification algorithms to our target sample. To identify the maximum number of quasars, we combined the results of the three classification algorithms. While this step increases the sample completeness, it also increases the amount of contamination on the sample despite each

algorithm having low degree of contamination (see Table 5.3). In Sections 5.3.3.6 and 5.4.4, we take additional steps to eliminate contaminants from our quasar sample. At this point, the classification algorithms identify 39160 objects as quasar candidates. Of these, 36699 lack spectroscopic observations, and 2470 sources have been classified spectroscopically. By crossmatching our sample with the Milliquas catalog, we find that 1374 are already known quasars. From these quasars, 1042 have redshifts larger than $z > 1.4$. This corresponds to a completeness of 87% for the sample of Boötes spectroscopic quasars. Additionally, we check the AGES catalog (Kochanek et al. 2012) and NED⁹ database to find that 1096 objects have been classified spectroscopically as either galaxies or galaxies. The confirmed stars and galaxies are removed from the sample.

5.3.3.6 Radio data

The contamination by the stars and galaxies in quasar samples is inevitable, as they occupy similar regions in the color space used to train the classification algorithms. As already discussed, an efficient way to eliminate stellar sources from quasar samples is to include information from radio surveys (Richards et al. 2002; Retana-Montenegro & Röttgering 2018). For this purpose, we use the LOFAR observations of the NDWFS-Boötes field by Retana-Montenegro et al. (2018a) introduced in Section 5.2.1. We crossmatch our quasar sample and the LOFAR catalog using a radius of $2''$, and inspect their stacked B_w, R, I images with radio contours overlaid to ensure that the match between optical and radio counterparts is correct, and to eliminate likely contaminants still present in our sample. To eliminate extended sources, we apply a morphology cut as a function of redshift to the photometric quasars after their photometric redshifts are calculated in Section 5.4. At this point, our RSQs sample contains 83 spectroscopic quasars and 172 photometric quasars with 5σ LOFAR detections. Figure 5.2 shows the colors of the training sample and the quasar samples in Boötes. In general, the colors of the photometric quasars agree well with those of the $z > 1.4$ quasars in the training and NDWFS-Boötes quasar samples.

⁹<https://ned.ipac.caltech.edu/>

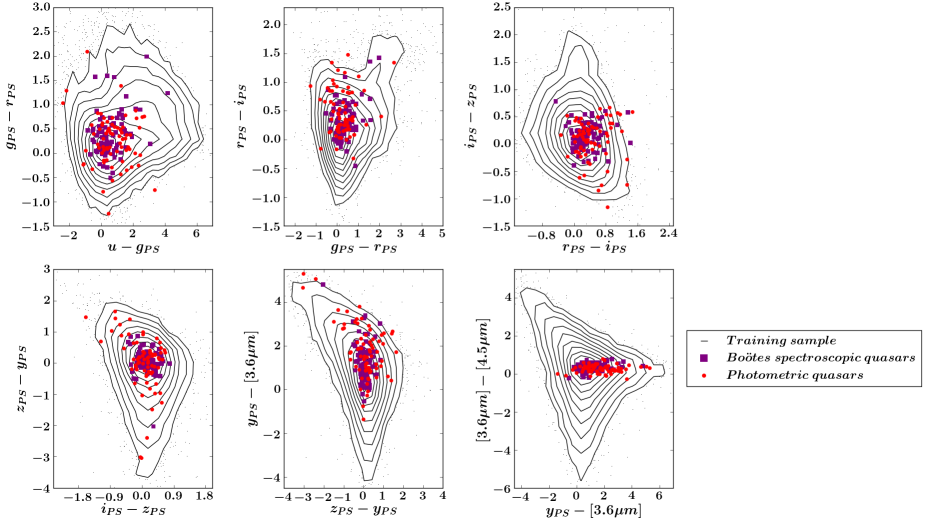


Figure 5.2: Comparison between the colors of the training and NDWFS-Boötes (photometric and spectroscopic quasars) samples. Black contours and points denote the spectroscopic quasars with $z > 1.4$ of the training sample, while purple squares indicate all the spectroscopic quasars in Boötes with $z > 1.4$. Photometric quasars in our sample are denoted by red circles. The training sample is employed to identify quasars in the target sample (see Section 5.3), and to determine their photometric redshifts with the NW kernel regression method (see Section 5.4.1).

5.4 Photometric redshifts

5.4.1 Nadaraya-Watson kernel regression

Our sample contains only 83 spectroscopic quasars. For the photometric quasars, we estimate their photometric redshifts z_{phot} using the Nadaraya-Watson (NW) kernel regression estimator (Nadaraya 1964; Watson 1964). The NW estimator is part of the family of kernel regression methods, in which the expectation of the variable Y relative to a variable X does not depend on all the X -values, as in traditional regression methods, but on sets of locally weighted values. A bandwidth scale parameter determines the amount of local averaging that is performed to obtain the estimate of Y . The NW estimator has been widely applied for nonparametric classification and regression (e.g.

Li & Racine 2011; Campbell et al. 2012; Qiu 2013), and to derive photometric redshifts (Wang et al. 2007; Timlin et al. 2018).

The NW estimate $\hat{y}$ is a weighted average of the observed variable y_i calculated utilizing nearby points around the test point x_0 . The estimate is calculated using the following equation:

$$\hat{y}(x_0) = \sum_{i=1}^N w_i(x_i, x_0) y_i, \quad (5.3)$$

where

$$w_i(x_i, x_0) = \frac{K(x_i, x_0)}{\sum_{i=1}^N K(x_i, x_0)}, \quad (5.4)$$

is the normalized kernel built using the local information from the training sample, and N is the number of objects in the training sample. The kernel weighting function $K(x_i, x_0)$ is chosen to have a Gaussian form:

$$K(x_i, x_0) = \exp\left(-\frac{1}{2h^2} \|x_i - x_0\|^2\right), \quad (5.5)$$

where h is the bandwidth scale that defines the region of parameter space in which the data is averaged, and $\|x_i - x_0\|$ is the euclidean distance between the data points from the training and test samples. A more detailed discussion about the NW estimator can be found in Härdle (1990) and Wu & Zhang (2006).

In this work, the training sample is composed of spectroscopic quasars, and the distance is calculated between the colors of the spectroscopic quasars and photometric quasars. Finally, for each photometric quasar its photometric redshift is calculated considering all the spectroscopic redshifts of the training sample by using the equation (Wang et al. 2007; Timlin et al. 2018):

$$z_{photo} = \sum_{i=1}^N w_i z_{spec,i}. \quad (5.6)$$

5.4.2 Quasar training sample

To provide a training sample on which to use the NW estimator, a spectroscopic quasar catalog is necessary. For this purpose, we use the quasar catalog introduced in Section 5.2.3. In total, our training set contains 328956 quasars with $1.2 \leq z \leq 5.5$ to determine photometric redshifts using the NW estimator (see Table 5.2).

5.4.3 Redshift estimation

The distance between the colors indices ($u-g_{PS}$, $u-r_{PS}$, $g_{PS}-r_{PS}$, $r_{PS}-i_{PS}$, $i_{PS}-z_{PS}$, $z_{PS}-y_{PS}$, $y_{PS}-[3.6\mu m]$, $[3.6\mu m]-[4.5\mu m]$) of the spectroscopic quasars from the training sample and the corresponding photometric objects are used as inputs to build the kernel matrix as given in eq. 5.5. An important decision in building the kernel is the choice of the bandwidth scale h . With smaller values of h nearby data points weight more, while larger values of h result in an increasing contribution of distant data points. This h value is slightly higher than the $h = 0.05$ used in previous estimations of photometric redshifts using the NW estimator (e.g. Wang et al. 2007; Timlin et al. 2018). Finally, the kernel functions determined using eq. 5.6 are used as weights in the computation of the photometric redshift.

The quality of the photometric redshifts determined with the NW estimator is investigated utilizing two quasar samples. The first sample is composed of 1193 quasars from the quasar training set described in Section 5.2.3 that are located in the NDWFS-Boötes field. The second sample is a subsample selected randomly from the training set. We measure the performance of the photometric redshifts in the samples using the following statistics:

- mean of the difference between photometric and spectroscopic redshifts, $\langle \Delta z \rangle = \langle (z_{phot} - z_{spec}) \rangle$, unclipped;

Table 5.4: The performance of the photometric redshifts for the quasar training sample. The statistics are given for each subsample.

Sample	Number of quasars	$\langle \Delta z \rangle$	$\sigma(\Delta z)$	$R_{0,3}$	$\langle \Delta z_{\text{norm}} \rangle$	$\sigma(\Delta z_{\text{norm}})$	$R_{0,1}^{\text{norm}}$	$R_{0,2}^{\text{norm}}$
Boötes (All)	1193	-0.060	0.35	0.76	-0.013	0.10	0.78	0.94
Boötes ($\varepsilon_{\text{IT}} \leq 0.2$)	535	-0.015	0.22	0.87	-0.0008	0.071	0.87	0.97
Boötes ($\varepsilon_{\text{IT}} \leq 0.3$)	689	-0.033	0.25	0.84	-0.006	0.080	0.84	0.97
Boötes ($\varepsilon_{\text{IT}} \leq 0.5$)	881	-0.04	0.26	0.82	-0.007	0.084	0.82	0.96
20% Training sample	65573	-0.006	0.28	0.81	0.005	0.09	0.82	0.95

Notes: The experiments all used the same training sample that does not include any spectroscopic quasars from the NDWFS-Boötes field, with the exception of the experiment where the Boötes quasars are included in the training sample.

- standard deviation of the Δz , $\sigma(\Delta z)$;
- fraction of quasars with $|\Delta z| < 0.3$, $R_{0.3}$;
- mean of normalized Δz , $\langle \Delta z_{\text{norm}} \rangle = \langle (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}}) \rangle$, unclipped;
- standard deviation of the normalized bias, $\sigma(\Delta z_{\text{norm}})$;
- fraction of quasars with $|\Delta z_{\text{norm}}| < 0.1$ ($|\Delta z_{\text{norm}}| < 0.2$), $R_{0.1}^{\text{norm}}$ ($R_{0.2}^{\text{norm}}$).

These results are summarized in Table 5.4. We tested practically all the values in the range $0.01 \leq h \leq 0.5$, and found that $h = 0.09$ gives the best performance and least scatter for the Boötes spectroscopic quasars (see Table 5.4). Figure 5.3 shows the photometric versus spectroscopic redshifts for the spectroscopic sample of Boötes quasars. At low- z , the dispersion of the photometric is slightly higher in comparison with the high- z estimations. This is expected as at low- z there are less spectral features for the NW method to exploit and predict trends on the training sample.

Figure 5.4 presents the normalized histogram of the bias $\Delta z = z_{\text{phot}} - z_{\text{spec}}$ between photometric and spectroscopic redshifts for the different experiments listed in Table 5.4. The first row lists the result for the sample of Boötes quasars. This shows that the majority of Boötes quasars have good redshift estimations. The number of outliers is small with 76% of the quasars having photometric redshifts that do not differ more than $|\Delta z| \leq 0.3$ from their spectroscopic redshifts. A good comparison for the redshift accuracy of the Boötes quasars can be obtained using a validation sample that is randomly selected from a training sample. This validation sample contains 65573 quasars, and its size is 20% of the total training sample. The result for this sample is listed in the last row of Table 5.4. The Boötes sample performs worse than the 20% of the training sample in terms of bias, scatter, and fractions of quasars with correct redshifts. In order to better understand the reasons for the difference in performance between the two samples we examine the performance of the NW method as a function of the photometric errors in the SDSS/Pan-STARRS1 bands. By restricting the Boötes sample only to quasars with photometric errors in the SDSS/Pan-STARRS1 bands smaller than $err \leq [0.2, 0.3, 0.5]$, respectively, its performance gets closer to that of the 20%

training sample, and this further improves for the case when the photometric errors are smaller or equal to 0.3 (Table 5.4, third-fourth rows).

Finally, we compare the performance of the NW regression kernel with other photometric-redshift algorithms. For instance, Richards et al. (2001) and Weinstein et al. (2004) reported that 70% and 83% of their predicted photometric redshifts are correct within $|\delta z| \leq 0.3$. These authors used empirical methods that used the color-redshift relations to derive photometric redshifts using early SDSS data. Yang et al. (2017) considered the asymmetries in the relative flux distributions of quasars to estimate quasar photometric redshifts obtaining an accuracy of $R_{0.2}^{\text{norm}} = 87\%$ for SDSS/WISE quasars. Jin et al. (2019) employed the SVM and XGBoost (Chen & Guestrin 2016) algorithms to achieve an average accuracy of $R_{0.2}^{\text{norm}} \simeq 89\%$ for the photometric redshifts of their Pan-STARRS1/WISE quasars. Schindler et al. (2017) used SDSS/WISE adjacent flux ratios to train the RF and SVM methods to obtain average results of $R_{0.2}^{\text{norm}} \simeq 93\%$ for their photometric redshifts. In summary, despite the fact that all these algorithms employed different samples and redshift intervals, their accuracy is consistent with the results obtained in this work using the NW regression kernel.

5.4.4 Final quasar sample

We have used the NW regression algorithm to assign a photometric redshift to each photometric quasar detected by LOFAR. To eliminate potential contamination by low- z galaxies, we restrict our quasar sample only to point sources. The SDSS photometry pipeline¹⁰ classifies an object as point-like (star) or extended (galaxy) source based on the difference between its PSF and cMODEL MAG magnitudes¹¹. Various methods have been proposed to perform the morphological star/galaxy separation with photometric data by adding Bayesian priors to the aforementioned magnitude differences (Scranton et al. 2002), and using decision tree classifiers (Vasconcellos et al. 2011). Here, we employ a criterion derived directly from the Boötes spectroscopic quasars by considering

¹⁰<https://www.sdss.org/dr14/algorithms/classify/>

¹¹<https://www.sdss.org/dr14/algorithms/magnitudes/>

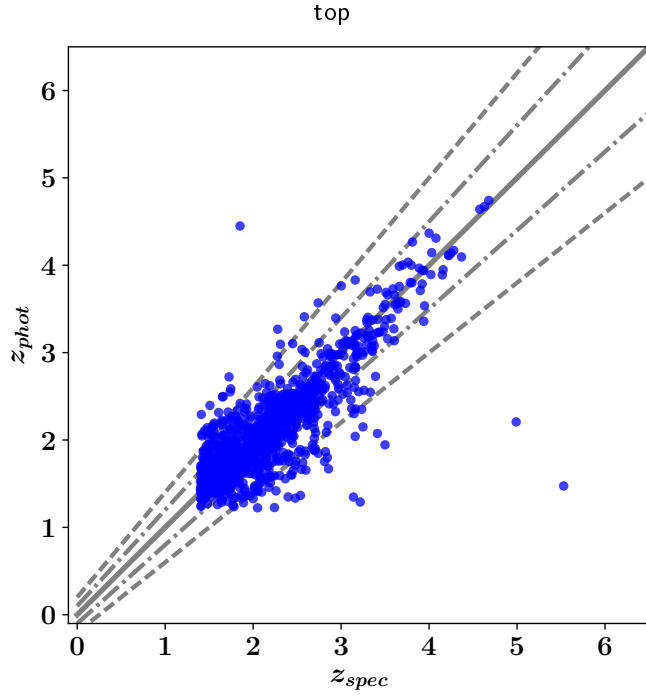


Figure 5.3: Photometric z_{photo} versus spectroscopic z_{spec} redshift for spectroscopic quasars in the NDWFS-Boötes region. The photometric redshifts are obtained using the NW kernel regression. The grey line indicates the one-to-one $z_{NW} = z_{spec}$ relation, and the dashed-dotted and dashed lines indicate the $z_{NW} - z_{spec} = \pm 0.10 \times (1 + z_{spec})$ and $z_{NW} - z_{spec} = \pm 0.20 \times (1 + z_{spec})$ relations, respectively.

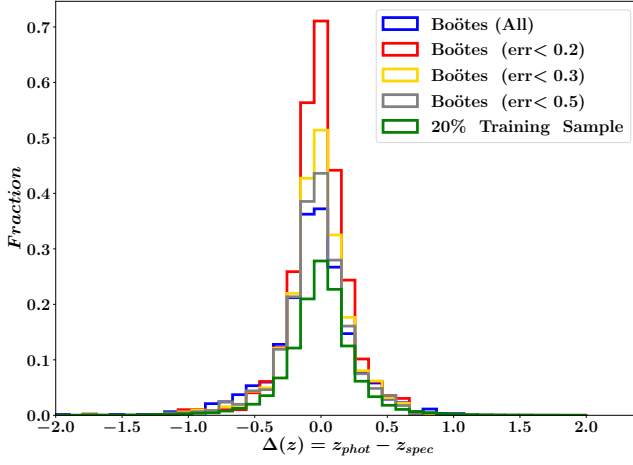


Figure 5.4: Normalized histogram of the bias $\Delta z = z_{\text{phot}} - z_{\text{spec}}$ between photometric and spectroscopic redshifts for different samples. The photometric redshifts are obtained using the NW kernel regression. Around 76% of the spectroscopic quasars in the Boötes field have photometric redshifts that are correct within $|\Delta z| = 0.3$ (see Table 5.4).

the difference Δ_{mag} between the PSF and cMODEL_{MAG} magnitudes in the SDSS-*r* band as a function of redshift. Spectroscopic quasars are binned according to their redshifts to calculate the magnitude difference as the quantile that contains 95% of the quasars in the corresponding bin. The redshift intervals match those utilized to derive the luminosity function in Section 5.5.4. Objects with magnitude differences less than the Δ_{mag} value in their corresponding redshift bin are considered to be point sources, and are included in our RSQs sample. Figure 5.6 shows the magnitude differences as a function of redshift. Finally, we restrict our RSQs sample to magnitudes $16.0 \leq i_{\text{PS}} \leq 23.0$ to avoid extrapolation beyond the range of the quasar training sample (see Table 5.2). The resulting catalog consists of 17924 objects, with 104 sources having a LOFAR counterpart within a radius of $2''$.

The final RSQs catalog resulting from our selection contains 51 photometric quasars and 83 spectroscopic quasars. In Figure 5.7, we compare the colors of the photometric quasars with those of the training sample and Boötes spectroscopic quasars as func-

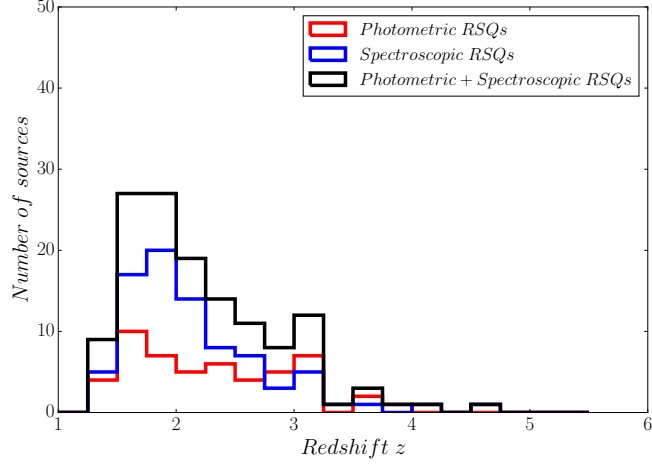


Figure 5.5: The redshift distribution of photometric (red) and spectroscopic (blue) RSQs. Also, the combined redshift distribution (black) of photometric and spectroscopic RSQs is plotted.

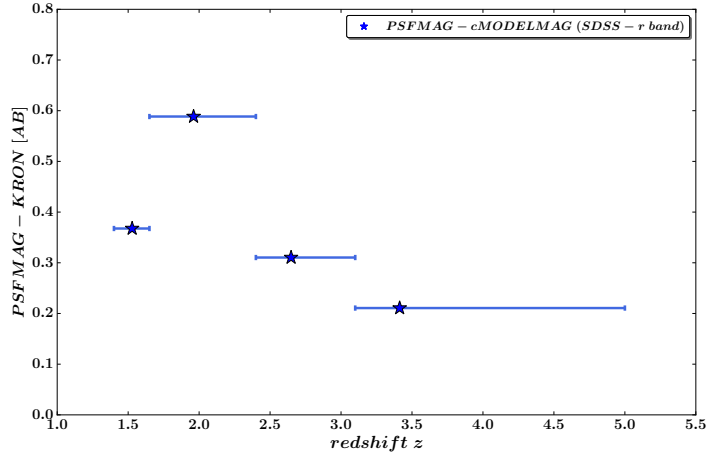


Figure 5.6: The difference between the PSF and cMODEL MAG magnitudes in the r band to separate point-like and extended sources as a function of redshift. The difference values are calculated considering the quantile that contains 95% of the quasars in the corresponding redshift bin. The redshift intervals match those used to derive the luminosity function in Section 5.5.4.

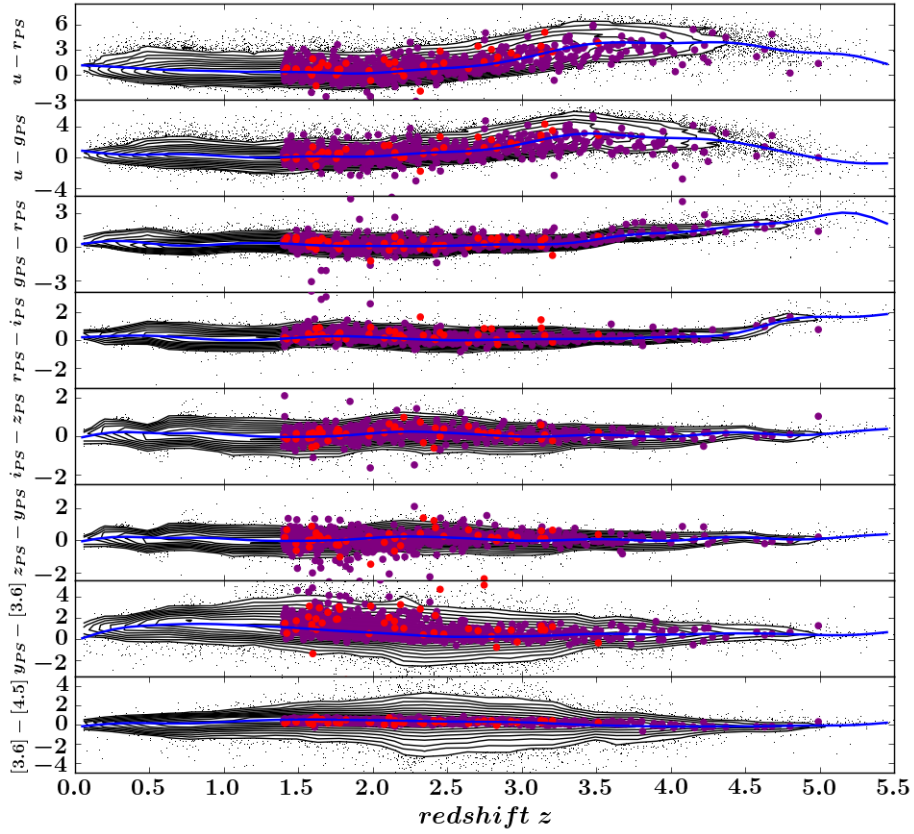


Figure 5.7: Quasar colors versus redshift for different quasar samples between $z = 1.4$ and $z = 5.5$. *Red points*: RSQs with photometric redshifts; *purple points*: Boötes spectroscopic quasars; *black contours* and points: the color distributions of the quasar training sample from Section 5.2.3; *blue lines*: mean color–redshift relations derived from the quasar training sample.

tions of redshift. The color-redshift spaces occupied by photometric RSQs are in good agreement with those of Boötes spectroscopic quasars and the quasar training sample. Figure 5.5 displays the redshift distribution of photometric and spectroscopic quasars. At $z \lesssim 2.8$, the majority of the quasars in our sample are spectroscopic. Considering that 77% of the photometric quasars have photometric errors smaller than $err \leq 0.5$, we conclude that the accuracy of their photometric redshifts is similar or slightly worse to that of the Boötes sample with photometric errors that are $err \leq 0.5$ (see Table 5.4). The i_{PS} -band magnitude and 150 MHz flux distributions of the RSQs sample are presented in Figure 5.8, while the absolute magnitude and radio luminosity are displayed in Figures 5.9 and 5.10, respectively. The absolute magnitudes are calculated using the K-correction discussed in Section 5.5.2. The majority of RSQs (134 in total) in our sample are unresolved or resolved into single-component radio sources in the LOFAR-Boötes mosaic with a resolution of $\sim 5''$. In our sample, only 11 quasars show present radio-morphologies consistent with a core-jet structure. Appendix 5.A presents a selection of false-color RGB images of spectroscopic and photometric RSQs from our final sample.

5.4.5 LOFAR and wedge-based mid-infrared selection of quasars

In this section, we compare the LOFAR and wedge-based mid-infrared selection for objects classified as quasars by the ML classification algorithms. Firstly, we use the mid-infrared color cuts proposed by Lacy et al. (2007) and Donley et al. (2012) to identify the presence of AGN-heated dust in our photometric RSQs. Figure 5.11 shows the mid-infrared colors for different quasar samples in the NDWFS-Boötes field. The mid-infrared colors of the photometric RSQs are in good agreement with those of spectroscopic quasars, and the majority of both spectroscopic and RSQs are located within the region delimited by the Donley et al. (2012) color cuts. To be exact, 89.93% (70.59%) of

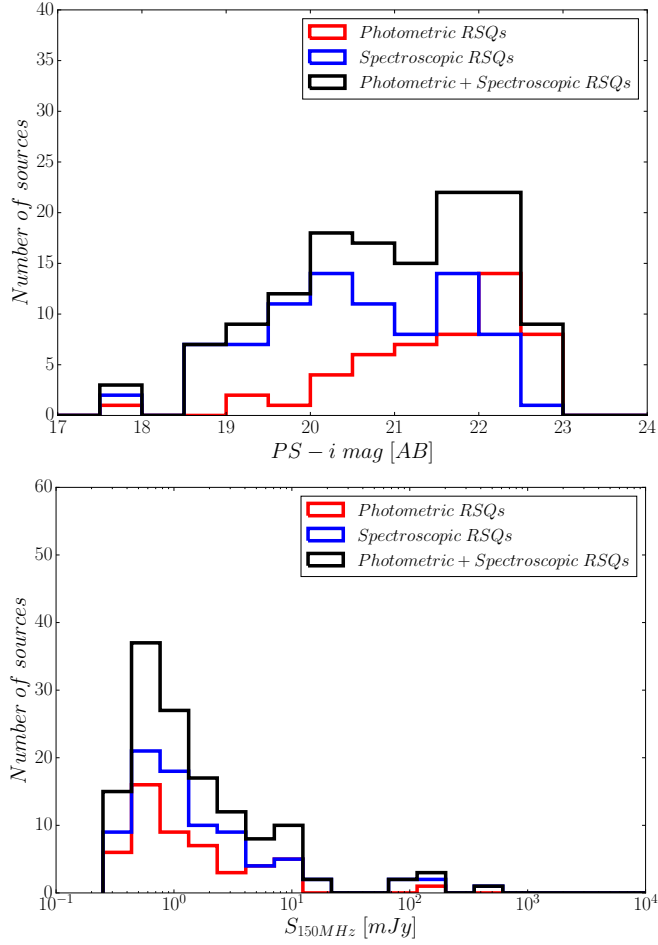


Figure 5.8: The i_{PS} (top) and total flux $S_{150\text{MHz}}$ (bottom) distributions of photometric (red) and spectroscopic (blue) RSQs. Also, the combined redshift distribution (black) of photometric and spectroscopic RSQs is plotted. The method employed to select the quasars is described in Section 5.3.3.

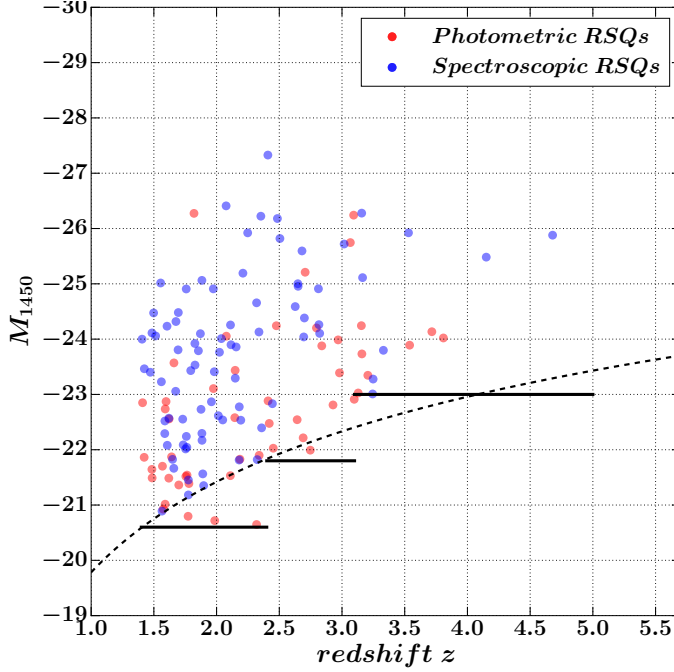


Figure 5.9: The absolute magnitude M_{1450} versus redshift for photometric (red) and spectroscopic (blue) RSQs in our sample. To minimize incompleteness due to incomplete M_{1450} bins while retaining the maximum numbers of quasars for estimating the luminosity function, we consider only RSQs with $M_{1450} \leq [-20.6, -21.8, -23.0]$ at $1.4 \leq z < 2.4$, $2.4 \leq z < 3.1$, and $3.1 \leq z < 5.0$; respectively. The dashed line denotes the magnitude limit $i_{PS} = 23.0$. This limit is calculated assuming a quasar continuum described by a power-law with slope $\alpha = -0.5$ with no emission line contribution or Ly_α forest blanketing.

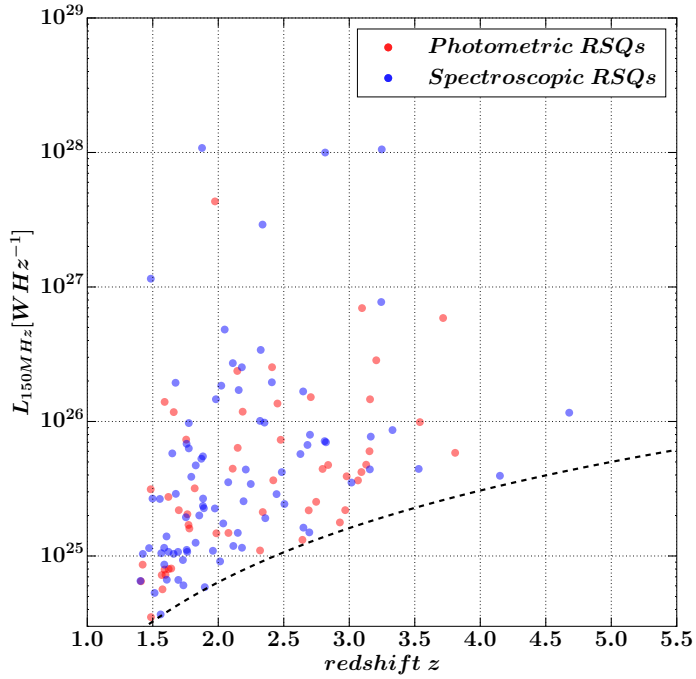


Figure 5.10: The 150 MHz rest frame absolute luminosity density versus redshift for photometric (red) and spectroscopic (blue) RSQs in our sample. The solid line denotes the 5σ flux limit ($275\mu\text{Jy}$) of the Boötes observations presented by Retana-Montenegro et al. (2018a).

the 1192 (51) spectroscopic quasars (photometric RSQs) with redshifts $z > 1.4$ are located within the region delimited by the Donley et al. (2012) color cuts, 98.15% (94.12%) reside in the region common to the Lacy et al. (2007) and Donley et al. (2012) color cuts, and only 1.85% (5.88%) are located outside the boundaries delimited by the aforementioned wedge-based mid-infrared color cuts. Considering the following points regarding our photometric RSQs: i) these objects are identified as quasar by the ML algorithms; and ii) their optical and mid-infrared colors are similar to those of spectroscopic quasars. We are confident therefore that our sample of photometric RSQs is composed mainly of real quasars and the number of contaminants is minimal. Moreover, these points show that utilizing ML algorithms trained using optical/infrared photometry and combined with LOFAR data is a very efficient and robust way to identify quasars.

To investigate the wedge-based mid-infrared selection of quasars without radio detections, we first need to establish the nature of these objects as robust quasars candidates or contaminants (i.e. stars or galaxies). For this purpose, we follow a classification method based on the goodness-of-fit, χ^2 , estimated through the fitting of the object's photometric data to a given type of spectral energy distribution (SED) template (quasar, galaxy, and stellar) (e.g. Ilbert et al. 2006; Masters et al. 2015). This fitting assigns to each object three χ^2 values: χ^2_{QSO} , χ^2_{GAL} , and χ^2_{STAR} , which corresponds to the fitting of the object photometry against the quasar, galaxy, and stellar templates, respectively. We perform these fittings using the photometric redshift code EAZY (Brammer et al. 2008), and utilize only the SDSS, Pan-STARRS1, and IRAC $3.6\mu\text{m}/4.5\mu\text{m}$ bands as this was the photometry used to classify them originally as quasars by the ML algorithms. This also represents a scenario where there is SDSS, Pan-STARRS1, and LOFAR coverage, but shallow or incomplete mid-infrared data to perform a wedge-based mid-infrared selection of faint quasars (e.g. Stern et al. 2005; Messias et al. 2012). The three template libraries used in this analysis are as follows. For the galaxy library, we use the latest version of the Brammer et al. (2008) SED templates that include nebular emission lines; while for the star library we select the Chabrier et al. (2000) SED templates. For the quasar template set, we use

a compilation of quasar templates from the literature. This includes the composite quasar templates by Cristiani & Vio (1990), Vanden Berk et al. (2001) and Gavignaud et al. (2006), respectively; and the type-1 quasar templates by Salvato et al. (2009) (pl_l22491_10_TQSO1_90, pl_QSOH, pl_QSO_DR2, pl_TQSO1). To account for dust extinction in the galaxies and quasar hosts, we modify the galaxy and quasar templates using the Calzetti et al. (2000) starburst and Pei (1992) Small Magellanic Cloud (SMC) extinction laws. Additionally, and only for quasar templates, we use the Czerny et al. (2004) quasar extinction curve. The extinction is applied to each template according to a grid $E(B - V) = [0.0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5]$. For all the SED libraries, the zero points are calculated using the standard procedure of fitting the observed SEDs of a sample of spectroscopic objects, fixing their redshift to the known spectroscopic redshift (see Ilbert et al. 2006 for more details). The spectroscopic samples used to determine the zero points are the quasars, galaxies, and stars of the training sample presented in Section 5.3.1. Due to the large number of templates in each template library, only single template fits are considered. From the χ^2 distribution of spectroscopic quasars in Boötes, we define empirical χ^2 cuts to separate quasars from stars and galaxies. Figure 5.12 displays the comparison of the χ^2 values resulting from the quasar, galaxy, and star template fitting to the photometry of spectroscopic and photometric quasars in the NDWFS-Boötes field. We find that the empirical cuts $\chi_{\text{STAR}}^2 \geq 22$, $\chi_{\text{QSO}}^2 \leq \chi_{\text{STAR}}^2 \times 0.33 + 8.33$ and $\chi_{\text{GAL}}^2 \geq 2.5$ select the majority of quasars and reject an important fraction of likely stars and galaxies. Based on the cuts described before, we select 957 total of 1193 spectroscopic quasars; and 4935 of 17829 photometric quasars. From these, 1423 photometric quasars are located in the region delimited by the Donley et al. (2012) color cuts. The analysis of the mid-infrared selection is limited to the Donley et al. (2012) wedge, as it is expected that the majority of quasars will be located within this region.

Next, we compare the number of photometric quasars selected by the Donley et al. (2012) wedge to the expected number of quasars. The expected number of quasars can be determined using the following expression:

$$N_{\text{QSO}} = A \times \iint \Phi^*(M_{1450}^*, z) V_c(z) dz dM_{1450}, \quad (5.7)$$

where A is the survey area, $\Phi^*(M_{1450}^*, z)$ is the quasar luminosity function, and $V_c(z)$ is the comoving volume. We estimate the number of total (radio detected and undetected) faint quasars using the results by Yang et al. (2018) and eq. 5.7. These authors studied the luminosity function of faint quasars between $0.5 < z < 4.5$ in a 1.0 deg^2 field within the VIMOS VLT Deep Survey (Le Fèvre et al. 2013). For redshifts $z > 3.5$, Yang et al. (2018) did not fit the luminosity function due to the low number of quasars in their sample, therefore, our analysis is limited to the redshift range $1.4 \lesssim z \lesssim 3.5$. According to Yang et al. (2018), the expected number of faint quasars at $1.4 \lesssim z \lesssim 3.5$ in a 9.3 deg^2 region like the NDWFS-Boötes field is approximately $N_{\text{QSO}} \sim 1060$. However, in the redshift range $1.4 \leq z \leq 3.5$ there are more than 1100 spectroscopic quasars, and 1380 photometric quasars fulfill the Donley et al. (2012) color cuts. Counting, only the photometric quasars and not known spectroscopic quasars, the number is higher than the expected number of quasars. This indicates some degree of contamination in the sample of photometric quasars selected using the wedge-based mid-infrared selection. Most likely these contaminants are compact galaxies or AGNs that mimic the colors of quasars, as the majority of stars are expected to be located outside the wedge define by the Donley et al. (2012) color cut. Additionally, comparing these numbers against those of the LOFAR selection, it is likely that there are more contaminants in the wedge-based mid-infrared selected sample than in the LOFAR selected sample. For a case like this, the addition of Euclid (Laureijs et al. 2011) near-infrared imaging (JH bands) in the ML classification process will be useful in eliminating some of these likely contaminants.

In summary, the results of this section demonstrate that in the cases where a lack of deep and complete mid-infrared coverage needed to perform a wedge-based mid-infrared selection of quasars, ML algorithms trained with optical/infrared photometry combined with LOFAR data provide an efficient and robust approach for obtaining samples of quasars.

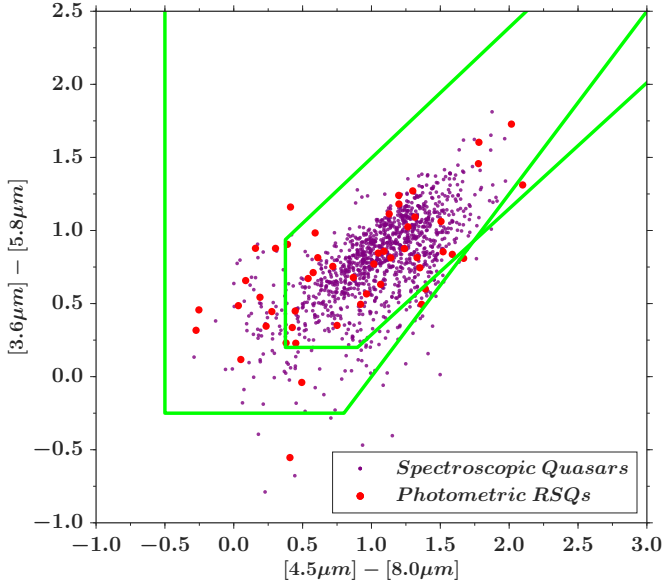


Figure 5.11: Mid-infrared colors for photometric and spectroscopic quasars in the Boötes field. The photometric RSQs are plotted as red circles, while the spectroscopic quasars are indicated by purple circles. The light green lines are the mid-infrared color cuts proposed by Lacy et al. (2007) and Donley et al. (2012).

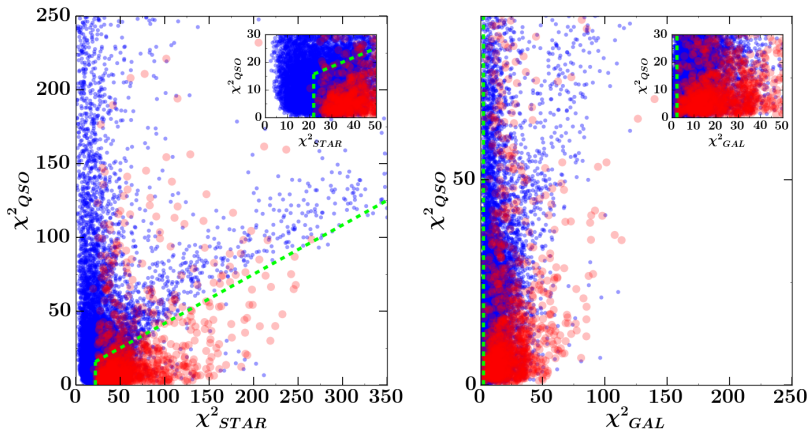


Figure 5.12: Comparison of the χ^2 values for quasar, galaxy, and star template fitting to the photometry of spectroscopic quasars and candidate quasars in the NDWFS-Boötes field.

5.5 Luminosity Function of radio-selected quasars

In the following subsections, we describe the steps required to compute the luminosity function of RSQs.

5.5.1 Selection completeness and accuracy of photometric redshifts

An important step in measuring the luminosity function of quasars is to account (and correct) for the different sources of incompleteness that could bias the quasar counts. In our analysis, we need to consider the completeness of our sample selection and the accuracy of the photometric redshifts. The selection completeness, $P_{\text{comp}}(i_{\text{PS}}, z)$, is the fraction of quasars that were successfully identified as quasars by the classification algorithms, as function of magnitude i_{PS} and redshift. $P_{\text{comp}}(i_{\text{PS}}, z)$ is derived from the data themselves as follows. First, the spectroscopic quasar sample introduced in Section 5.3.1 is binned according to magnitude i_{PS} and redshift z , in bins of size $\Delta i_{\text{PS}} = 0.5$, and $\Delta z = 0.3$, respectively. The quasars in each bin are separated into two subsamples. The first subsample is created by randomly sampling without replacement all quasars in the bin, while the second subsample includes all the quasars that were not sampled. The sizes of the first and second subsamples are 20% and 80% of quasars in the $i_{\text{PS}} - z$ bin, respectively. Having done this for all bins, the corresponding samples are combined to create final training (80%) and target (20%) samples. The final result is the uniform and randomly sampled separation of the training sample into two subsamples in the $i_{\text{PS}} - z$ plane. The main advantage of this binning scheme is that it provides an unbiased and efficient way to map locally the selection completeness obtained using the classification algorithms as a function of magnitude and redshift. The second subsample is used as the training sample for the classification algorithms, while the first subsample has the role of target sample and it is employed to derive the selection completeness.

In addition to correcting for selection incompleteness in our sample, we also need to correct for the accuracy of the photometric redshifts determined using the NW regression

method. For this purpose, we determine the expected number of spectroscopic quasars to have photometric redshifts correctly and incorrectly assigned within a redshift bin using the NW method. This is done following a similar approach to determining the selection completeness. First, the spectroscopic quasar sample introduced in Section 5.3.1 is divided into the same redshift bins used to derive the luminosity function in Section 5.5.4. In each bin, the quasars contained in that bin are randomly separated to create samples with sizes of 20% and 80% of all quasars in the bin, respectively. The corresponding samples from all the bins are combined to create final training (80%) and target (20%) samples. The training sample is used to train the NW regression method, while the target sample is utilized to determine the expected number of spectroscopic quasars with correctly and incorrectly assigned redshifts within the boundaries of the redshift bins. The ratio between the number of spectroscopic quasars with correctly and incorrectly assigned redshifts, $f_{\text{photo-z}}$, provides an estimate of the excess of photometric quasars with incorrectly assigned photometric redshifts within a redshift bin. This ratio is used as a correction factor for each photometric quasar within the corresponding redshift bin. The derived correction factors have a median factor of $f_{\text{photo-z}} \simeq 1.0$, with the $1.65 < z < 2.4$ redshift bin having the smallest value with $f_{\text{photo-z}} = 0.90$.

5.5.2 Simulated Quasar Spectra

In order to calculate the K-correction (see Section 5.5.3), we construct a synthetic quasar library that is an accurate representation of the quasar demography. The variety in the quasar spectral features (UV continuum slope, emission-line EW, and intervening HI absorbers along the line of sight) determine the range in quasar colors. It is important that these spectral features are taken into account to obtain reliable simulated quasar spectra. These spectra are later incorporated into a synthetic quasar library that allow us to compute the K-correction for a given redshift. We explain below the procedure followed to build our synthetic quasar library

Following several authors (Møller & Jakobsen 1990; Fan 1999; Richards et al. 2006; McGreer et al. 2013), the synthetic quasar spectra are built using a Monte-Carlo (MC) approach. We perform MC simulations to generate quasar spectra adopting a broken

power-law ($f_\lambda \propto \lambda^{-\alpha_\lambda}$) for the UV continuum at 1100 Å. The slope values are drawn from a Gaussian distribution, with mean values of $\langle \alpha_\lambda \rangle = -1.7$ for $\lambda < 1100$ Å (Telfer et al. 2002), and $\langle \alpha_\lambda \rangle = -0.5$ for $\lambda > 1100$ Å (Vanden Berk et al. 2001), both with standard deviations of $\sigma = 0.30$. We bin BOSS quasars by their luminosity to obtain the distribution for the parameters (wavelength, EW, FWHM) of emission lines. This allows us to recover the intrinsic emission line mean and dispersion as function of luminosity, as well as reproducing empirical trends such as the Baldwin effect (Baldwin 1977). We again assume gaussianity when the emission line features are added to the quasar continuum. For each template spectrum, the intergalactic absorption that gives rise to the Ly $_\alpha$ forest is included by creating sightlines in a MC fashion adopting the prescription of neutral absorbers by Bershady et al. (1999). The spectrum is then convolved with our filter passbands to obtain the colors for each synthetic quasar. A Gaussian error is added to the photometry of the mock quasars with a σ derived from the photometric errors of the real magnitudes that match the simulated ones. This error is combined in quadrature with the photometric calibration errors of Pan-STARRS1 (Tonry et al. 2012). In Figure 5.1, we show a synthetic spectrum from our quasar library.

5.5.3 K-correction

Usually, the luminosity functions of quasars are expressed in the absolute magnitude at rest-frame 1450 Å, M_{1450} , which provides a good measurement of the quasar continuum in a region without strong emission lines (e.g. Richards et al. 2006; Croom et al. 2009b; Masters et al. 2012). To derive M_{1450} for the Boötes quasars, we use the apparent magnitude m_X in a fiducial filter as a proxy:

$$M_{1450} = m_X - 5 \log (d_L/10) - K_X, \quad (5.8)$$

where $d_L(z)$ is the luminosity distance in parsecs, and K_X is the K-correction which allows us to convert the magnitudes of distant objects in a given bandpass filter into an equivalent measurement into their rest-frame. Using our synthetic quasar library described in Section 5.5.2, the K-correction can be determined from the difference

between apparent magnitudes m_X and m_{1450}

$$K_X = m_X - m_{1450} - 2.5 \log(1 + z), \quad (5.9)$$

with m_{1450} calculated using a top-hat filter of width $50 \text{ \AA}$. Figure 5.13 displays the K-correction obtained for five different filters, and the expected result from a quasar that has only a power-law continuum and no emission line contribution or Ly_α forest blanketing. The K-correction curves between $1.0 \leq z \leq 6.0$ are obtained by calculating their average value in redshift bins of size $\Delta z = 0.1$. At $z \gtrsim 3.7$, the difference between the r_{PS} and i_{PS} bands becomes more significant as the Ly_α line moves or exits the filters. The same situation occurs at $z \gtrsim 4.7$, but for the i_{PS} and z_{PS} bands. Therefore, we estimate M_{1450} using K-corrections selected to minimize any bias caused by the redshifting of the Ly_α emission line. For $z < 3.7$ quasars, we use a K-correction based on the r_{PS} band, while for the intervals $3.7 \leq z \leq 4.7$ and $z > 4.7$ K-corrections based on the i_{PS} and z_{PS} bands, respectively, are employed.

5.5.4 Quasar Luminosity function

We construct the luminosity functions for all quasars in our radio-matched sample using the classical $1/V_{\text{max}}$ method (Schmidt 1968) for flux limited samples. The main advantage of this method is that an assumption about the underlying model of the luminosity function is not required. The estimator adopted to compute the comoving quasar density in a certain luminosity bin is:

$$\begin{aligned} \Phi(L) = \frac{1}{\Delta L} \sum_{i=1}^n & (F(S_{150\text{MHz}}) \times P_{\text{comp}}(i_{\text{PS}}, z) \\ & \times P_{\text{comp}}(i_{\text{PS}}, z) \times f_{\text{photo-}z}(z) \times V_{\text{max},i})^{-1} \end{aligned} \quad (5.10)$$

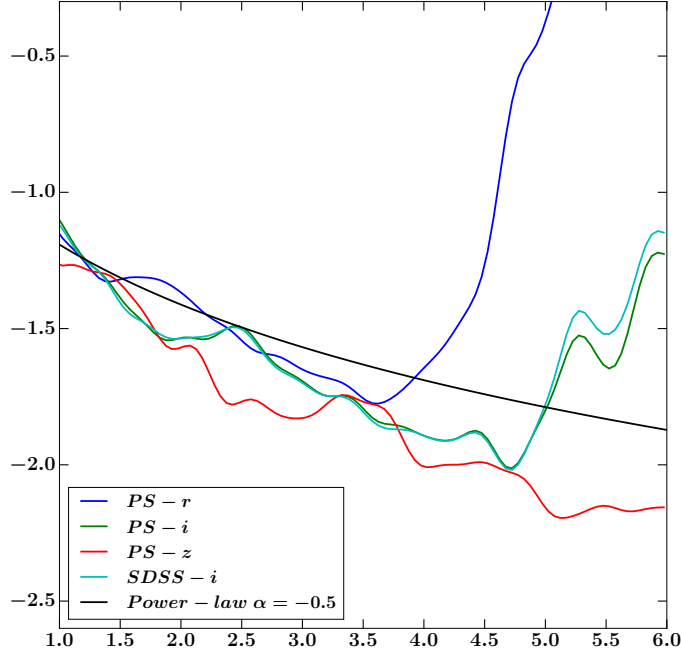


Figure 5.13: K-correction for different filters determined using our simulated quasar spectra. The Pan-STARRS1 r_{PS} and i_{PS} filters are indicated by blue and green, while the red and cyan are the expected K-corrections for the Pan-STARRS1 z_{PS} and SDSS- i bands. The solid black line is the K-correction assuming a power-law with slope $\alpha = -0.5$ with no emission line contribution or Ly_{α} forest blanketing. At $z \gtrsim 3.7$, the difference between the r_{PS} and i_{PS} bands becomes more significant as the Ly_{α} line moves in or out of the filters. The same situation occurs at $z \gtrsim 3.7$, but for the i_{PS} and z_{PS} bands.

where n is the number of quasars in the luminosity bin, $V_{\max,i}$ is the maximum comoving volume in which a quasar would be observable and included in our sample, ΔL is the luminosity bin width, $F(S_{150\text{MHz}})$ is the radio-catalog completeness of the LOFAR-Boötes mosaic (Retana-Montenegro et al. 2018a). $P_{\text{comp}}(i_{\text{PS}}, z)$ is the selection completeness, and $f_{\text{photo-z}}(z)$ is the accuracy of the NW photometric redshifts derived in Section 5.5.1. Since our quasar sample is built using a radio-optical survey, we calculate $V_{\max}$ using the maximum redshift at which the flux of a quasar with a certain luminosity lies above the corresponding flux limit (Cirasuolo et al. 2005; Tuccillo et al. 2015), $z_{\max} = \min(z_{\max}^{\text{R}}, z_{\max}^{\text{O}})$, where $z_{\max}^{\text{R}}$ and $z_{\max}^{\text{O}}$ are the maximum redshifts according to the radio and optical flux limits.

We model the quasar luminosity function as a double power-law in absolute magnitude M_{1450} (Pei 1995)

$$\Phi(M_{1450}) = \Phi^*(M_{1450}^*) \times \left(10^{0.4(\alpha+1)(M_{1450}-M^*)} + 10^{0.4(\beta+1)(M_{1450}-M^*)} \right)^{-1}, \quad (5.11)$$

where M_{1450}^* is the break magnitude, Φ^* the normalization constant, α is the faint-end slope, and β is the bright-end slope. We split the quasar sample into four different redshift intervals between $1.4 < z < 5.0$ with a M_{1450} bin size equal to $\Delta M_{1450} = 1.2 \text{ mag}$. The redshift intervals and M_{1450} bin size are selected to obtain the highest S/N in the luminosity function calculations. Due to the small number of quasars in each luminosity bin ($N_{\text{QSO}} < 50$), the error bars are calculated assuming the low-statistics limit, using the 84.13% confidence Poisson upper limits and lower limits from Gehrels (1986). Finally, we use the M_{1450} lower limits indicated in Figure 5.9 to avoid incompleteness effects in the calculation of the luminosity function.

Figure 5.14 shows the luminosity function measurements in four subpanels with one for each redshift interval. We use a total of 83 spectroscopic and 51 photometric quasars to estimate the quasar luminosity function. The resulting binned luminosity functions are tabulated in Table 5.5. The luminosity function in the range $1.65 < z < 2.4$ is plotted as a reference in all the subpanels. This reference indicates that the space

Table 5.5: Binned luminosity functions for RSQs between $1.4 < z < 5.0$.

Redshift range	z_{median}	M_{1450}^* bin center	$\log(\Phi)^a$	σ_{low}^b	σ_{upp}^b	N^c
$1.4 < z < 1.65$	1.57	-24.8	-7.03	0.42	1.81	3
		-23.6	-6.50	1.89	5.01	6
		-22.4	-6.38	2.90	5.79	11
		-21.2	-6.30	3.09	7.53	7
$1.65 < z < 2.4$	1.92	-26.0	-7.36	0.22	0.77	4
		-24.8	-6.52	1.99	4.54	8
		-23.6	-6.62	1.82	3.08	18
		-22.4	-6.33	3.58	5.95	19
		-21.2	-6.23	3.42	6.41	13
$2.4 < z < 3.1$	2.70	-27.2	-7.70	0.02	0.34	1
		-26.0	-7.14	0.55	1.47	6
		-24.8	-6.83	0.91	1.97	9
		-23.6	-7.39	0.56	1.66	5
		-22.4	-6.64	2.01	4.35	9
$3.1 < z < 5.0$	3.25	-27.2	-8.13	0.01	0.24	1
		-26.0	-7.04	0.41	1.78	3
		-24.8	-7.63	0.08	0.54	2
		-23.6	-6.99	0.70	1.44	10

Notes: ^a Φ is in units of $1 \times 10^{-7} \text{ Mpc}^{-3} \text{ mag}^{-1}$. ^b σ_{low} and σ_{upp} are in units of $1 \times 10^{-7} \text{ Mpc}^{-3} \text{ mag}^{-1}$. ^c N is the number of quasars in the corresponding M_{1450} bin.

density of RSQs is higher at $1.65 < z < 2.4$ in comparison to the other redshift intervals, i.e. the comoving space density of RSQs reaches a maximum between $1.65 < z < 2.4$. The comoving space density of RSQs is discussed with further detail in Section 5.6.3. Additionally, a good continuity between the points of the faint and bright ends is obtained in the five redshift intervals.

5.6 Results

5.6.1 Model-fitting

The observed evolution of the AGN luminosity function has been traditionally studied using luminosity (Boyle et al. 2000; Croom et al. 2009b; Richards et al. 2006), density (McGreer et al. 2013; Willott et al. 2010; Kashikawa et al. 2015), and even hybrid luminosity/density models (Ueda et al. 2003; Hasinger et al. 2005; Ross et al. 2013;

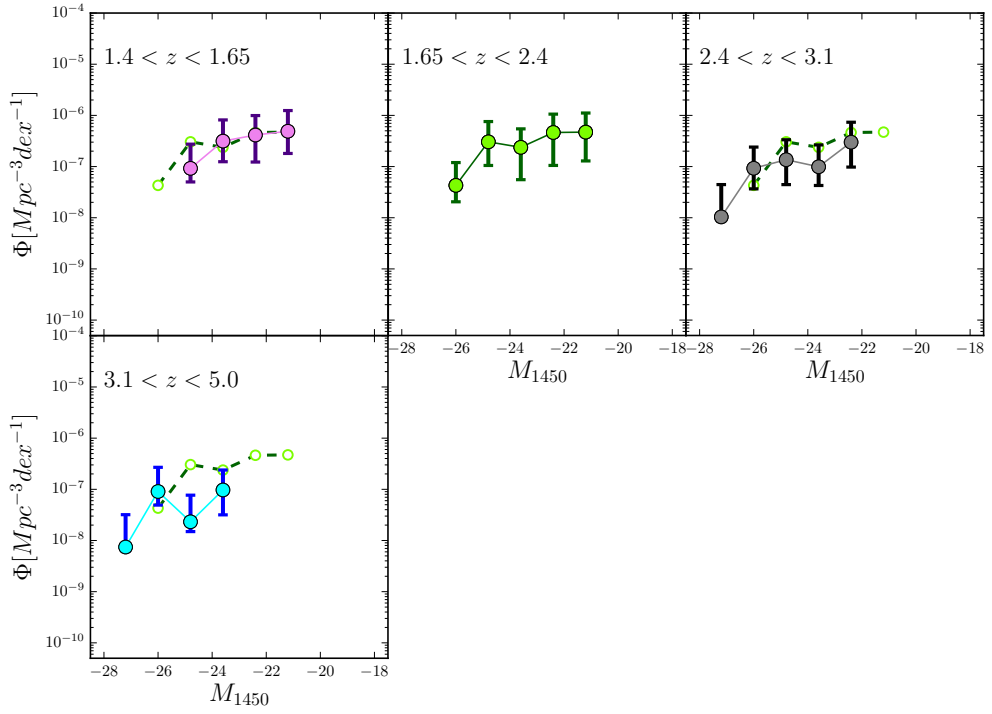


Figure 5.14: The rest-frame M_{1450} binned luminosity functions of our Boötes RSQ samples (colored circles) for five non-overlapping redshift bins between $1.4 < z < 5.0$. In each panel, we show as a reference the luminosity function at $1.65 < z < 2.4$.

Palanque-Delabrouille et al. 2016). These studies have found that the evolution of the luminosity function of quasars can be described by a pure luminosity evolution (PLE) model at $z \lesssim 2.2$ (Croom et al. 2009b), while a combined luminosity evolution and density evolution (LEDE) model can describe its evolution at $z \gtrsim 2.2$ (Croom et al. 2009b; Ross et al. 2013; Palanque-Delabrouille et al. 2016). The PLE introduces the redshift-dependence of the break magnitude using the following second-order polynomial (Croom et al. 2009b)

$$M_{1450}^*(z) = M_{1450}^*(z=0) - 2.5 (k_1 z + k_2 z^2), \quad (5.12)$$

while the LEDE model introduces the redshift-dependence in the normalization and break magnitude using the following log-linear ansatz:

$$\log(\Phi^*) = \log[\Phi^*(z=z_p)] + c_1 (z - z_p), \quad (5.13)$$

$$M_{1450}^*(z) = M_{1450}^*(z=z_p) + c_2 (z - z_p), \quad (5.14)$$

where z_p is the pivot redshift. Following previous works (e.g. Ross et al. 2013), we employ the PLE model to fit our binned luminosity function for redshift intervals $z < 2.4$, while at $z > 2.4$ we use the LEDE model with $z_p = 2.4$.

We use the χ^2 minimization to fit the luminosity function data points in each redshift bin to the corresponding models described above. Because of the relatively small area ($\sim 9.3 \text{ deg}^2$) of the Boötes field, there are only a few bright quasars in our sample. This implies that the bright-end slope β will be determined with high-uncertainty due to small number statistics. Therefore, we fix the bright-end slope β to the values reported by Ross et al. (2013) in their study of the quasar luminosity function using SDSS-DR9/BOSS data (Ahn et al. 2012). Additionally, we fix the parameters (k_1, k_2) and (c_1, c_2) in the PLE and LEDE models, respectively, to the values obtained by Ross et al. (2013). The parameter values from Ross et al. (2013) are chosen to match our redshift intervals. For the LEDE model, we use the parameter values corresponding to their S82 sample. Using the parameters from their DR9 sample produces similar results. Finally,

the best-fit parameters and their associated uncertainties are summarized in Table 5.6. The corresponding best-fit model is shown with a colored line in each subpanel of Figure 5.15. The models have a good agreement with the binned luminosity function.

5.6.2 Comparison to previous works

In Figure 5.15, we also compare our best-fit models with previous works. The SDSS-III/BOSS luminosity function (Ross et al. 2013) was estimated at $2.2 < z < 3.5$ employing a uniform sample of 22301 quasars over an area of 2236 deg^2 . Additionally, Ross et al. (2013) investigated the evolution of the QLF using a combination of SDSS (Richards et al. 2006), boss21+MMT and BOSS Stripe 82 datasets to over a redshift range of $0.3 < z < 4.75$. Ross et al. (2013) fitted their QLF data using a PLE model at $z < 2.2$, while their fittings at $z > 2.2$ were carried out employing a LEDE model. Furthermore, we compare our results with previous surveys of faint quasars ($M_{1450} < -22.0$) (Siana et al. 2008; Glikman et al. 2011; Giallongo et al. 2015; Masters et al. 2015; Akiyama et al. 2018; McGreer et al. 2018). The survey area of these studies ranges from 170 arcmin^2 (Giallongo et al. 2015) to 339.8 deg^2 (Akiyama et al. 2018). Finally, we also consider the results of SDSS and 2QZ (Croom et al. 2001, 2004) radio-selected samples (Cirasuolo et al. 2005; McGreer et al. 2009; Tuccillo et al. 2015) using FIRST (Becker et al. 1995). It is clear that the number density of our RSQs at all redshifts considered is lower in comparison with that of samples of optically bright and faint quasars, which are composed of both radio-detected and radio-undetected objects. Naturally, the number density of RSQs is higher than the density of SDSS/2QZ FIRST-selected samples (Cirasuolo et al. 2005; McGreer et al. 2009; Tuccillo et al. 2015), which are composed mainly of radio-loud quasars with fluxes $S_{1.4\text{GHz}} > 1.0 \text{ mJy}$. This flux limit corresponds to a LOFAR flux of $S_{150\text{MHz}} > 4.80 \text{ mJy}$, assuming a spectral index of $\alpha = -0.7$. As can be seen in the bottom panel of Figure 5.8, the fraction of quasars fluxes with $S_{150\text{MHz}} > 4.80 \text{ mJy}$ is just 18% of the total number of RSQs in our sample. Therefore, the lower number densities presented by SDSS/2QZ FIRST-selected samples

Table 5.6: Parametric model best-fit parameters and uncertainties. See Section 5.5.4 for more details about the models used.

Model	Redshift range	α	β	$M_{1450}^*(z=0)$	k_1	k_2	Φ^*
		[faint-end]	[bright-end]	[mag]			[Mpc ⁻³ mag ⁻¹]
PLE	$1.4 < z < 1.65$	-1.13 ± 0.01	-3.55	-20.99 ± 0.01	1.293	-0.268	$3.28 \times 10^{-7} \pm 4.03 \times 10^{-9}$
PLE	$1.65 < z < 2.4$	-1.18 ± 0.15	-3.55	-21.87 ± 0.43	1.293	-0.268	$2.25 \times 10^{-7} \pm 8.80 \times 10^{-8}$
PLE	$1.4 < z < 2.4$	-1.19 ± 0.16	-3.55	-20.99 ± 0.25	1.293	-0.268	$3.61 \times 10^{-7} \pm 1.15 \times 10^{-7}$
				$M_{1450}^*(z=2.4)$	c_1	c_2	
LEDE	$2.4 < z < 3.1$	-1.40 ± 0.16	-3.51	-26.98 ± 1.38	-0.689	-0.809	$6.17 \times 10^{-8} \pm 6.20 \times 10^{-8}$
LEDE	$3.1 < z < 5.0$	-1.13 ± 0.65	-3.51	-26.41 ± 2.06	-0.689	-0.809	$1.40 \times 10^{-7} \pm 2.41 \times 10^{-7}$
LEDE	$2.4 < z < 5.0$	-1.29 ± 0.18	-3.51	-26.12 ± 0.58	-0.689	-0.809	$1.63 \times 10^{-7} \pm 1.07 \times 10^{-7}$

Note: Parameters without errors are kept fixed during the fitting.

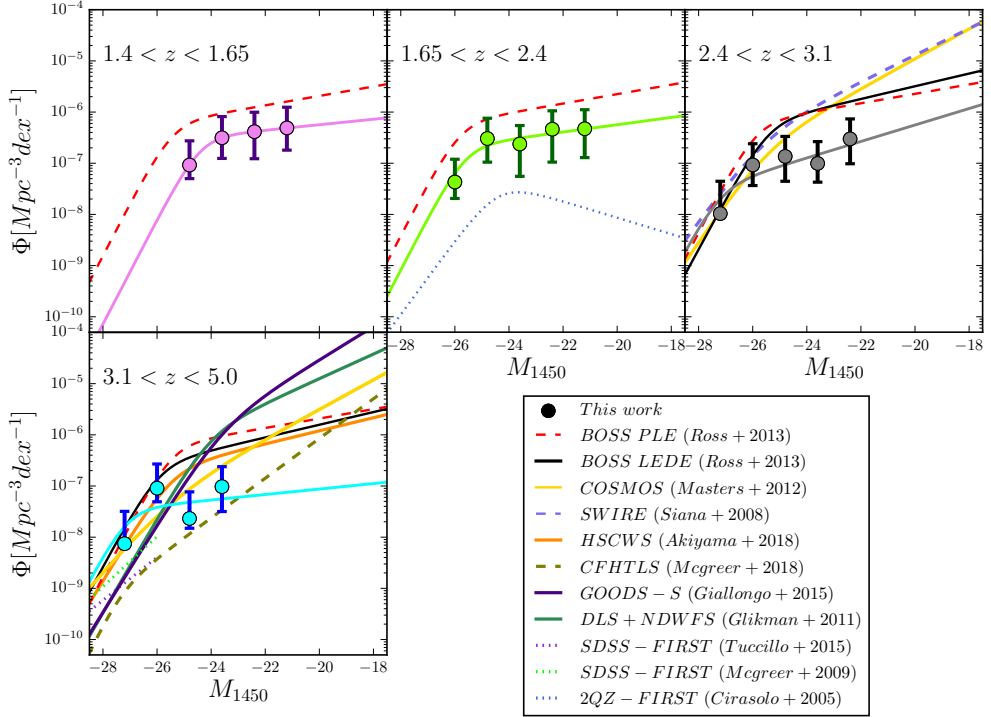


Figure 5.15: The rest-frame M_{1450} binned luminosity functions of our Boötes RSQ samples (colored circles) for four non-overlapping redshift intervals between $1.4 < z < 5.0$. The lines show the corresponding best-fit models in each redshift bin. The best-fitting parameters for each fit are presented in Table 5.6. For comparison, we show the QLFs from previous works (Cirasuolo et al. 2005; Siana et al. 2008; McGreer et al. 2009; Glikman et al. 2011; Masters et al. 2012; Ross et al. 2013; Giallongo et al. 2015; Tuccillo et al. 2015; Akiyama et al. 2018; McGreer et al. 2018) measured over similar redshift intervals. The single-power law fits by McGreer et al. (2009) and Tuccillo et al. (2015) are plotted in the range $-29 \leq M_{1450} \leq -26$, which is the original range where they were measured.

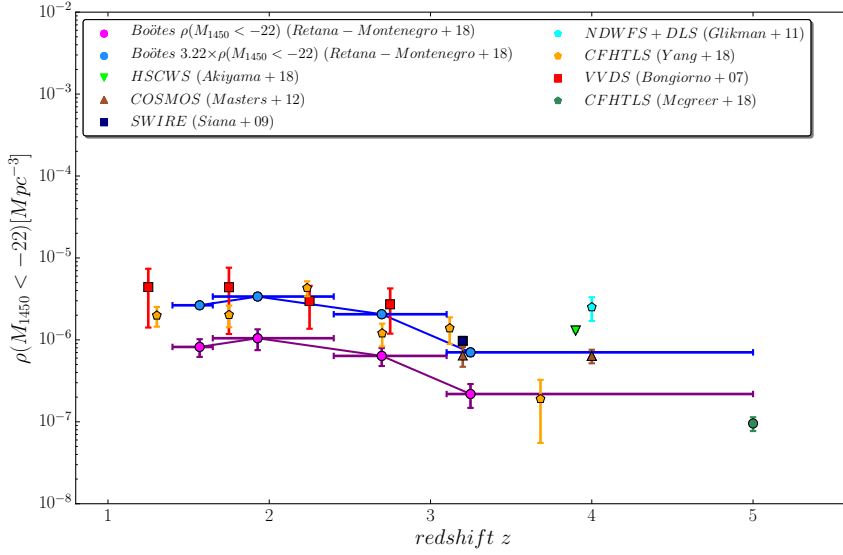


Figure 5.16: The spatial density of RSQs with $M_{1450} < -22$ as a function of redshift compared to the space density of faint quasar samples ($M_{1450} < -22$) from the literature. The spatial density of RSQs is indicated by purple circles, while estimates from the literature (Bongiorno et al. 2007; Siana et al. 2008; Glikman et al. 2011; Masters et al. 2012; Akiyama et al. 2018; Yang et al. 2018; McGreer et al. 2018) are represented by the corresponding symbols in the legend box. We also plot the spatial density of RSQs scaled by a factor of 3.22 ($1/0.31$) (blue circles).

are expected.

We compare our best-fit parameters with previous studies at different redshifts to constrain the evolution of the luminosity function of RSQs. In Figure 5.17, we compare the normalization constant $\log(\Phi^*)$, the break magnitude M_{1450}^* , and the faint-end slope α with previous values reported for faint quasars as a function of redshift. We also plot the PLE and LEDE models by Ross et al. (2013), as well as our PLE and LEDE models. From Figure 5.17, we find the following trends:

1. Our $\log(\Phi^*)$ values are lower in comparison with those of other samples of faint quasars (Siana et al. 2008; Glikman et al. 2011; Masters et al. 2012; Niida et al. 2016; Yang et al. 2018; Akiyama et al. 2018). For redshifts $z < 2.4$, the normalization constant within uncertainties is consistent with a PLE evolutionary trend. At $z > 2.4$, $\log(\Phi^*)$ seems to decrease following a linear-log trend reminiscent of a LEDE evolution, and similar to that of the faint quasars.
2. The break magnitude M_{1450}^* seems to get brighter with increasing redshift, a trend that is consistent with previously estimates (Siana et al. 2008; Glikman et al. 2011; Masters et al. 2015; Niida et al. 2016).
3. The faint-end slope α does become steeper with increasing redshift with a mean value of $\alpha = -1.15$ at $z < 2.4$, while at $z > 2.4$ the mean value is $\alpha = -1.26$. Our α values are consistent within error bars with those reported previously by several authors (Siana et al. 2008; Glikman et al. 2011; Niida et al. 2016; Akiyama et al. 2018; Yang et al. 2018).

5.6.3 Density evolution of RSQs

In the last 20 years, the evolution of quasar activity has been studied in the UV/optical (Fan et al. 2001; Wolf et al. 2003; Richards et al. 2006; Fontanot et al. 2007; Bongiorno et al. 2007; Croom et al. 2009b; Glikman et al. 2011; Ross et al. 2013; Masters et al. 2015; Jiang et al. 2016; McGreer et al. 2018; Yang et al. 2018), X-ray (Hasinger et al. 2005; Silverman et al. 2005; Aird et al. 2015; Georgakakis et al. 2015; Miyaji et al.

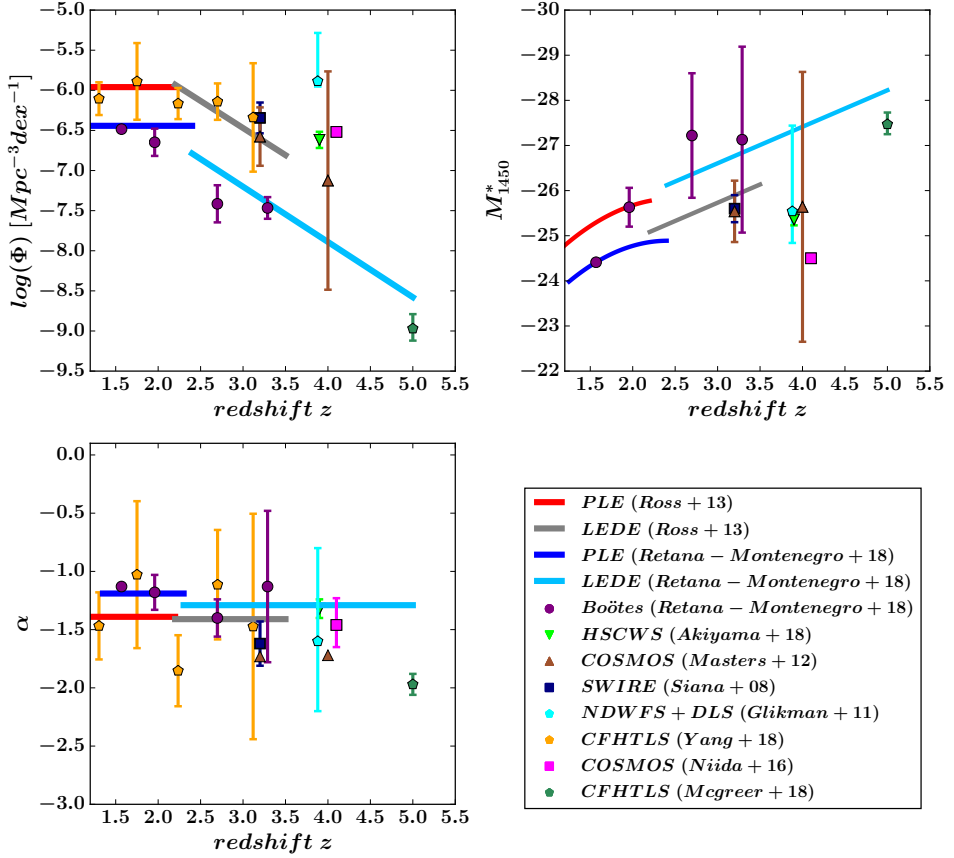


Figure 5.17: The best-fit quasar luminosity function parameters as a function of redshift. Our results are indicated by purple circles, while estimates from the literature (Siana et al. 2008; Glikman et al. 2011; Masters et al. 2012; Niida et al. 2016; Akiyama et al. 2018; Yang et al. 2018; McGreer et al. 2018) are represented by the corresponding symbols in the legend box. The red and gray lines represent the PLE and LEDE models from Ross et al. (2013), and the blue and dark cyan lines our PLE ($1.4 < z < 2.4$) and LEDE ($2.4 < z < 5.0$) models listed on Table 5.6. For clarity, we shift vertically the PLE and LEDE models from Ross et al. (2013) by a factor of +0.1 in the third panel.

2015), infrared (Brown et al. 2006; Siana et al. 2008; Assef et al. 2011), and radio (Vigotti et al. 2003; Carballo et al. 2006; Cirasuolo et al. 2006; McGreer et al. 2009; Tuccillo et al. 2015). Here, we study this evolution using the spatial density of quasars (Fan et al. 2001; McGreer et al. 2013; Tuccillo et al. 2015)

$$\rho(< M_{1450}, z) = \int_{-\infty}^{M_{1450}} \Phi(M_{1450}, z) dM, \quad (5.15)$$

where $\Phi(M_{1450}, z)$ is the luminosity function of quasars, and it is integrated over all quasars more luminous than M_{1450} . The integration is performed using the binned luminosity functions instead of using the best-fit luminosity functions, as it avoids uncertainties related to the model fitting and the extrapolation of the models. An upper limit of $M_{1450} = -22.0$ is selected for the integration, as it is the lowest luminosity limit that is common between our work and other samples of faint quasars, to compare their spatial density.

Figure 5.16 displays the space density of RSQs ($M_{1450} < -22.0$) from our sample as a function of redshift, along with other faint quasar samples from the literature. Our spatial density estimates have lower values in comparison with those of other samples. However, our spatial density values show a very similar redshift evolution to that of faint quasars, and this trend continues towards $z = 5.0$. It is clear that a peak in quasar activity occurs around $z \approx 2.0$ according to our results and the surveys carried out by Bongiorno et al. (2007) and Yang et al. (2018). This agrees with the picture provided by measurements of the luminosity function using samples of bright quasars ($M_{1450} < -24.0$) (e.g. Richards et al. 2006; Croom et al. 2009b; Ross et al. 2013). However, the quasar samples by Akiyama et al. (2018) and Glikman et al. (2011) display space densities much higher than expected in comparison with the trend suggested by our results and other samples of faint quasars (Bongiorno et al. 2007; Masters et al. 2015; Yang et al. 2018; Siana et al. 2008; McGreer et al. 2018). These higher spatial densities could be attributed to sample contamination (Akiyama et al. 2018) or cosmic variance (Glikman et al. 2011). Finally, and as expected, our values of the space density of RSQs are significantly higher in comparison to those previously estimated for RLQs. Samples of RLQs present space density values of just a few Gpc^{-3} (e.g. Vigotti et al.

2003; Tuccillo et al. 2015). This highlights the fact that deep LOFAR observations allow us to detect the radio-emission of quasars that otherwise would be classified as radio-quiet (Retana-Montenegro & Röttgering 2018).

Having computed the spatial density for the relevant samples of faint quasars and our RSQ sample, we calculate their normalized to $z \sim 2$ spatial densities as a function of redshift. In the left panel of Figure 5.18, we show the normalized spatial density of RSQs and faint quasars with $M_{1450} < -22$. It is clear that the space density of faint quasars and RSQs decreases rapidly with redshift. From its maximum at $z \sim 2$, the space density of faint quasars (RSQs) declines between $z \simeq 2$ and $z \simeq 3$ by a factor of 1.74 ± 0.84 (1.64 ± 0.41), while it reduces further from $z \simeq 2$ to $z \simeq 5.0$ by a factor of 5.23 ± 1.41 (4.80 ± 1.55). At $z \sim 1.5$, the normalized space-density ratio is 1.07 ± 0.59 (1.28 ± 0.31). Note the agreement within error bars between the evolution of the space density of RSQs and that of other faint quasar samples (Bongiorno et al. 2007; Siana et al. 2008; Masters et al. 2015; Yang et al. 2018). Considering the works by Akiyama et al. (2018) and Glikman et al. (2011), the space density of faint quasars decreases from $z \simeq 2$ to $z \simeq 5.0$ by a factor of 3.54 ± 0.87 .

Only a small fraction, less than 10%, of the quasars are classified as radio-loud (Kellermann et al. 1989; Ivezić et al. 2002; Jiang et al. 2007a). However, RLQs are often associated with massive host galaxies (Shen et al. 2009; Retana-Montenegro & Röttgering 2017, and references therein), whose radio emission is produced by large and powerful radio jets (Bridle et al. 1994; Mullin et al. 2008). However, the dependency of the radio-loud fraction (RLF) of quasars on redshift and luminosity is still a matter of debate. Some authors have found that the RLF is a strong function of luminosity (e.g. Padovani 1993; La Franca et al. 1994) and redshift (e.g. Peacock et al. 1986; Miller et al. 1990; Visnovsky et al. 1992), while others have found that it does not depend significantly on either redshift or luminosity (e.g. Stern et al. 2000; Cirasuolo et al. 2003). Kratzer & Richards (2015) found that selection effects could be biasing the conclusions about the evolution of the RLF. Using the LoTSS survey (Shimwell et al. 2019), Gurkan (2019) showed that quasars exhibit a wide continuum of radio properties, with no clear bimodality in the radio-loudness parameter.

In the context of this work, we compute the relative fraction of RSQs with respect to the spatial density of faint quasars as a function of redshift by dividing the spatial density of RSQs, $\rho_{\text{RSQs}}(z)$, by the spatial density of faint quasars (radio-detected plus radio-undetected), $\rho_{\text{QSO}}(z)$. Figure 5.18 displays the relative fraction of RSQs, $\rho_{\text{RSQs}}(z)/\rho_{\text{QSO}}(z)$, as a function of redshift. The relative fraction of RSQs considering the error bars and excluding the results of Glikman et al. (2011) and Akiyama et al. (2018) is roughly independent of redshift, with a median value of 0.31 ± 0.22 . This fraction is of 0.28 ± 0.20 considering the results of Glikman et al. (2011) and Akiyama et al. (2018). In Figure 5.18, the spatial density of RSQs is multiplied by a factor of 3.22 ($1/0.31$) to compare it with the spatial density of faint quasars. With the multiplicative factor applied, the agreement between the two spatial densities is good. Moreover, it highlights the similarity in the redshift evolution of RSQs and faint quasars up to $z \sim 5$. A fraction of ~ 0.30 of RSQs with respect to faint quasars is significantly higher than the fractions of $\sim 0.10 - 0.15$ of RLQs with respect to the whole quasar population previously estimated (e.g. Goldschmidt et al. (1999); Stern et al. (2000); Jiang et al. (2007a)). However, this is not unexpected; as previously mentioned, our deep LOFAR observations allow us to detect the radio-emission of a considerable number of quasars that otherwise would be identified as radio-quiet.

Finally, our results for the spatial density of RSQs demonstrate that the selection of quasars utilizing ML algorithms that combines optical/infrared with LOFAR observations (see Section 5.3.3.5) is very efficient and robust.

5.6.4 Contribution of RSQs to IGM Photoionization

To compute the RSQs contribution to the IGM photoionization, we consider the total quasar emissivity of hydrogen-ionizing photons, $\dot{\eta} \text{ [s}^{-1} \text{ Mpc}^{-3}]$,

$$\dot{\eta} = \epsilon_{1450} \xi_{ion}, \quad (5.16)$$

where ϵ_{1450} is the quasar emissivity at 1450 Å, and ξ_{ion} is the number of ionizing

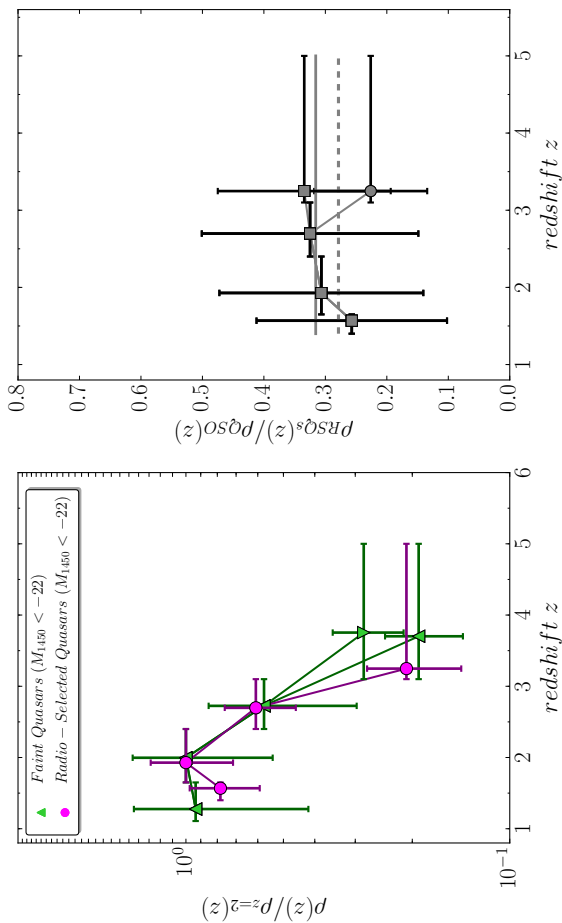


Figure 5.18: *Left panel:* Spatial densities, normalized to $z \sim 2$, as a function of redshift for optically faint quasars and RSQs. Our results are indicated by fuchsia circles, while the spatial density for faint quasars is determined from results reported in the literature (Bongiorno et al. 2007; Siana et al. 2008; Glikman et al. 2011; Masters et al. 2012; Akiyama et al. 2018; Yang et al. 2018) and is denoted by green triangles. The upward triangle indicates the spatial density excluding the results by Glikman et al. (2011) and Akiyama et al. (2018), while the value denoted the downward triangle includes them. *Right panel:* Relative fraction of RSQs with respect to the spatial density of faint quasars as a function of redshift. The ratio is calculated between overlapping redshift bins. The gray solid line indicates the mean ratio of 0.31 excluding the works by Glikman et al. (2011) and Akiyama et al. (2018), while the dashed line denotes a mean ratio of 0.28 including these works.

photons for a source with a luminosity of $1 \text{ erg s}^{-1} \text{ Hz}^{-1}$ at $1450 \text{ \AA}$. The quasar emissivity (Madau et al. 1999) can be computed using the observed quasar luminosity function $\Phi(L_\nu, z)$:

$$\epsilon_\nu = \int L_\nu \Phi(L_\nu, z) dL_\nu, \quad (5.17)$$

where the integration limits are the relevant survey limiting magnitudes, and L_ν is the typical quasar SED at frequency ν . This is approximated by a broken power-law SED with slope $\alpha_\lambda = -1.7$ for $\lambda < 1100 \text{ \AA}$ (Telfer et al. 2002), and $\alpha_\lambda = -0.5$ at $1100 < \lambda < 2500 \text{ \AA}$ (Vanden Berk et al. 2001). To calculate ξ_{ion} , we integrate the quasar SED in the energy range 1-4 ryd (photons with a energy larger than 4 ryd are absorbed by Hell). The luminosity function is integrated over the luminosity range $-30 \leq M_{1450} \leq -18$. The upper limit of $M_{1450} = -18$ is chosen as it is a typical value in the literature. Additionally, a value of unity is assumed for the photon escape fraction as suggested by Grazian et al. (2018), who recently studied the Lyman Continuum escape fraction for a large sample of $z \sim 4$ faint AGNs.

Using the derived best-fit models of the luminosity function from Table 5.6, we find a total quasar emission rate of hydrogen-ionizing photons of $\dot{n} = 9.60 \times 10^{49} \text{ s}^{-1} \text{ Mpc}^{-3}$ at $z = 3.29$. We compare this value with the total photon emissivity per unit comoving volume required to ionize the universe at a given redshift (Madau et al. 1999; Fan et al. 2001)

$$\dot{\mathcal{N}} = 10^{51.2} \left(\frac{C}{30} \right) \left(\frac{1+z}{6} \right)^3 \left(\frac{\Omega_b h^2}{0.02} \right)^2, \quad (5.18)$$

where $\Omega_b h^2 = 0.02$ is the baryon density (Spergel et al. 2007), and C is the clumping factor of the IGM. We assume that the clumping factor evolves with redshift according to the relation $C(z) = 1 + 43^{-1.71z}$ (Pawlik et al. 2009). At $z = 3.25$, the total photon emissivity required to ionize the IGM is $\dot{\mathcal{N}} = 1.26 \times 10^{50} \text{ s}^{-1} \text{ Mpc}^{-3}$. Our quasar emission rate is approximately 76% of the rate needed to reionize the IGM at $z = 3.25$. However, taking into account the systematic uncertainties involved in the determination of the total quasar emissivity such as the escape fraction, clumping

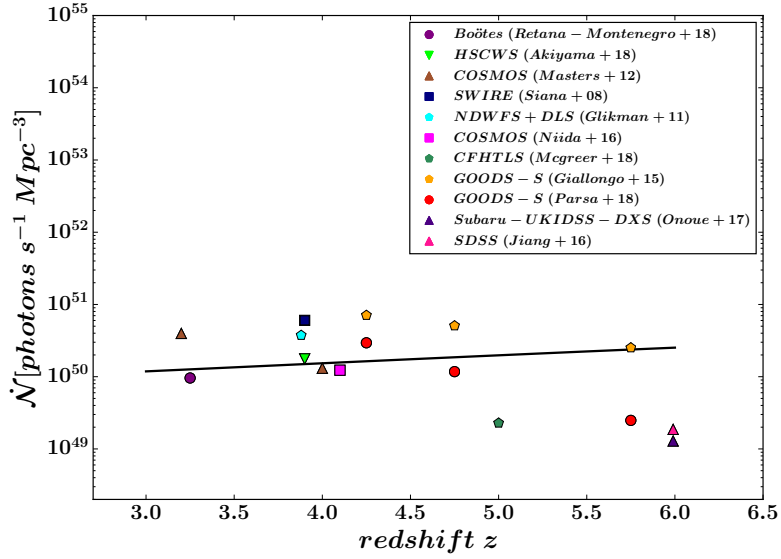


Figure 5.19: The hydrogen ionizing rate as a function of redshift. Results from this work are indicated by purple circles, while the results from the literature (Siana et al. 2008; Glikman et al. 2011; Masters et al. 2012; Niida et al. 2016; Akiyama et al. 2018; Yang et al. 2018; McGreer et al. 2018) are represented by the corresponding symbols in the legend box. The solid line represents the hydrogen ionizing rate required to ionize the IGM at any given redshift.

factor (Jiang et al. 2016), quasar SED, and the lower magnitude limit of the luminosity function, we conclude that RSQs produce only a few per cent of the photons required to ionize the IGM at $z > 3$. Additionally, as determined in Section 5.6.3, RSQs are only a fraction of the whole quasar population, therefore their spatial density is too low to be major contributors to the IGM ionization. This scenario is supported by the comparison shown in Figure 5.19 between our quasar emission rates and previous works (Siana et al. 2008; Glikman et al. 2011; Giallongo et al. 2015; Masters et al. 2015; Niida et al. 2016; Akiyama et al. 2018; Yang et al. 2018; McGreer et al. 2018). Our results agree with the majority of the previous quasar emission rates estimated in the literature and support the aforementioned conclusion.

5.7 Discussion

In this section, we discuss several aspects related to RSQs such as: the origins of their radio-emission, the environments where these objects reside, and their location in spectroscopic parameter spaces.

5.7.1 The origins of radio-emission in RSQs

An important piece of information needed to understand the origins of the radio-emission in RSQs could come from their observed radio-morphologies. Although, some of the brightest RLQs have double-lobed radio morphologies, the majority of intermediate-luminosity RLQs show core dominated radio-morphologies (de Vries et al. 2006; Lu et al. 2007; Coziol et al. 2017). Approximately, 92% of RSQs in our sample present compact radio-morphology at the resolution of the LOFAR-Boötes mosaic. A possible explanation for the origin of radio-emission in these objects lies in the interaction between outflows and the IGM, as first suggested by Stocke et al. (1992) to explain the low radio-emission in broad absorption-line quasars (BALQSOs). In this mechanism, radio-emission originates from particles accelerated on the shock fronts caused by the collision of uncollimated central outflows with the IGM of the host galaxy (Zakamska & Greene 2014; Zakamska et al. 2016). This scenario is supported by the observations of BALQSOs, as these objects have intermediate radio-luminosities and the majority present core dominated radio-morphologies (Becker et al. 2000; Liu et al. 2008; DiPompeo et al. 2011; Morabito et al. 2018), along with core-jet structures (Kunert-Bajraszewska et al. 2015) and lobes (Welling et al. 2014) in some instances. Assuming that the absorbing troughs observed in the spectra of BALQSOs are caused by uncollimated central outflows loaded into the broad emission-line region (BELR), the low-numbers of BAL systems in RLQs with double-lobed radio-morphology (DiPompeo et al. 2011; Pu 2013; Welling et al. 2014) can be attributed to the fact that in these quasars the central outflows form collimated jets which are physically separated from the BELR. In this scenario, the lack of radio-emission in quasars, traditionally classified as radio-quiet, can be explained considering that in a majority of cases the outflowing material is slowed

down by a dense interstellar clump and the formation of shock fronts is hindered. In our LOFAR-Boötes mosaic, we detect the radio-emission of many quasars that otherwise would have remained radio-undetected. The radio-emission in these quasars could have originated from the interaction between quasar outflows and the IGM. However, deeper optical and low-frequency radio surveys, in addition to LOFAR sub-arcsecond resolution observations (Varenius et al. 2015; Morabito et al. 2016), are needed to explore this mechanism in detail.

5.7.2 The environment of RSQs

A fraction of 92% of RSQs in our sample could be classified as compact steep-spectrum sources (CSS) according to their radio properties. CSS sources are usually a fraction of $\sim 10 - 30$ per cent in previous radio surveys (Peacock & Wall 1982; Fanti et al. 1990; O’Dea 1998), and they are characterized by their small projected linear sizes and median steep radio spectrum ($\alpha < -0.77$, O’Dea 1998). The brighter RSQs in our sample have a steep spectral index distribution with a median value of $\alpha \simeq -0.70$ (Retana-Montenegro & Röttgering 2018), and only 8% of RSQs in our sample present morphologies consistent with core-jet structures.

It has been suggested that CSS may be small either because they are young and still in an early stage of their evolutionary path, eventually developing into Fanaroff-Riley type-I/II (Fanaroff & Riley 1974) radio sources (Fanti et al. 1995; Alexander 2000; Snellen et al. 2000; Collier et al. 2016), or because they are embedded in a very dense environment that frustrates the propagation of the radio jets (van Breugel et al. 1984; Fanti et al. 1986, 1989; Orienti et al. 2007). The compactness of their radio-morphologies suggests that RSQs may reside in host-galaxies with a large supply of gas to fuel the early stages of quasar activity. Ultimately, these scenarios will have to be tested against high-resolution observations with submillimeter and radio interferometers that can spatially resolve the host-galaxies of RSQs and their synchrotron-dominated core/jets, respectively. These observations will, in turn, help us to shed light on the complex interplay between RSQs and their host-galaxies, and how quasar activity is triggered in these systems.

5.7.3 RSQs and their location in spectroscopic parameter spaces

The most striking features of Figure 5.18, excluding the results by Glikman et al. (2011) and Akiyama et al. (2018) from the analysis are: i) RSQs show evolutionary trends and declining factors that are similar to those presented by faint quasars ($M_{1450} < -22.0$) (see Figs. 5.16, 5.17, 5.18), and ii) the fact that RSQs may compose to up $31 \pm 22\%$ of the total faint quasar population, a fraction that within uncertainties is independent of redshift (see Fig. 5.18). Interestingly, similar decline factors in the space density of low- (Warren et al. 1994; Croom et al. 2009b; Palanque-Delabrouille et al. 2016) and high- (Schmidt et al. 1995; Kenefick et al. 1995; Palanque-Delabrouille et al. 2016) optical luminosity quasars, respectively, had been reported before. In these works, high-luminosity quasars have declining factors of $\simeq 2 - 3$ between $z \approx 2$ and $z \approx 4$, while low-luminosity quasars present steeper declining factors of $\simeq 6 - 8$ between the same redshift intervals. These factors are consistent with a downsizing evolutionary scenario. In this scenario, high-luminosity quasars evolve first at earlier epochs and reach their maximum space density at high- z , while low-luminosity quasars predominantly evolve at later epochs reaching their maximum space density at low- z (e.g. Hasinger et al. 2005; Silverman et al. 2005; Croom et al. 2009b). Similar declining factors to those of high-luminosity quasars had been reported for RLQs samples (Hook et al. 1998; Vigotti et al. 2003).

Since the early 2000's there has been substantial advances in our understanding of quasars using the broad emission line properties and their correlations. Probably, the most widely used broad emission line correlations are the eigenvector 1 (E1) and C_{IV} parameter spaces. The E1 parameter space started as the primary eigenvector in the Principal Component Analysis performed by Boroson & Green (1992), where Fe_{II} and H_{β} emission are related to line width. The generalization of this concept led to the 4D Eigenvector 1 (4DE1) parameter space, with the addition of the properties of the C_{IV} and the soft X-ray photon index (Sulentic et al. 2000a,b). The 4DE1 parameter space serves as a 4D equivalent of the 2D Hertzsprung-Russell diagrams (Hertzsprung 1909; Russell 1914). This parameter space has revealed a principal sequence of quasars char-

acterized by the Eddington ratio, and since its introduction has become an important tool for depicting the diversity of quasars and their evolutionary states (see Sulentic & Marziani 2015, and references therein). The C_{IV} parameter space (C_{IV} EW versus C_{IV} blueshift) has been used to study different quasar properties at high- z (e.g. Brotherton & Francis 1999; Sulentic et al. 2007; Richards et al. 2011; Kratzer & Richards 2015; Coatman et al. 2016). In the context of the 4DE1, E1, and C_{IV} parameter spaces, several authors (Sulentic et al. 2003, Sulentic et al. 2007, Zamfir et al. 2008, and Richards et al. 2011) have determined that RLQs and RQQs are clustered at different locations in their corresponding parameter spaces (see Fig. 14 in Richards et al. 2011 and Fig. 3 in Sulentic et al. 2003). In particular, Richards et al. (2011) and Kratzer & Richards (2015) demonstrated using the E1 and C_{IV} spaces, which may trace the relative power of radiation line-driven accretion disk winds (Richards et al. 2011), that on average RLQs present weaker radiation line-driven winds in comparison with RQQs. These authors suggest that RLQs and RQQs are two parallel evolutionary sequences, and possibly a series of spin/mergers events (Sikora et al. 2007; Sikora 2009; Schulze et al. 2017) are responsible for the triggering of radio-jets, and turning RQQs into radio-loud.

Considering that RSQs present evolutionary trends similar to those of faint quasars, and bright quasars, it is possible that faint (radio and optically) RSQs could share properties of both RLQs and RQQs. Thus, RSQs would occupy intermediate locations between RQQs and RLQs in their corresponding E1, 4DE1 and C_{IV} parameter spaces. Future spectroscopic studies of RSQs would be a major step forward towards understanding radio-loudness.

5.8 Conclusions

In this work, we train three ML algorithms: RF, SVM, and Bootstrap aggregation with optical/infrared imaging to compile a sample of quasars in the 9.3 deg^2 Boötes field. We eliminate stellar and likely galaxy contaminants from our sample by requiring a 5σ detection in deep LOFAR imaging, and applying a morphological criterium, respectively. The final sample consists of 134 spectroscopic quasars and photometric quasars in

the redshift range of $1.4 \leq z \leq 5.0$. We estimate the photometric redshifts of the photometric quasars using the NW kernel regression estimator (Nadaraya 1964; Watson 1964). When comparing the predictions of this method to the spectroscopic redshifts of 1193 Boötes spectroscopic quasars, we find that 76% of the quasars have photometric redshifts that are within $|\delta z| \leq 0.3$ of their spectroscopic redshifts. We demonstrate that in cases of lack of deep and complete mid-infrared coverage needed to perform a wedge-based mid-infrared selection of AGNs, the selection of quasars using ML algorithms trained with optical/infrared photometry in combination with LOFAR data is an effective approach for obtaining samples of quasars. We compute the fraction of quasars missed due to our selection (i.e. selection function) using a library of simulated quasar spectra. The binned luminosity function of RSQs is computed using the $1/V_{\text{max}}$ method (Schmidt 1968) in five different redshift bins between $1.4 \leq z \leq 5.0$. These luminosity functions are corrected for incompleteness due to the radio observations and selection method employed. The parametric fits to the binned luminosity function of RSQs are consistent with a PLE evolution model at $z < 2.4$, and a LEDE evolution at $z > 2.4$.

We have studied the luminosity function of RSQs down to faint luminosities of $M_{1450} = -22$. Previous studies were mostly limited to bright RLQs with luminosities $M_{1450} > -26$ (Cirasuolo et al. 2005; Carballo et al. 2006; McGreer et al. 2009; Tuccillo et al. 2015). We find evidence that suggests that the faint-end slope α is becoming steeper with increasing redshift, as found by previous measurements of the luminosity function of faint quasars (Glikman et al. 2011; Giallongo et al. 2015). Our mean values of the faint-end slope are $\alpha = -1.15$ at $z < 2.4$, while at $z > 2.4$ the mean value is $\alpha = -1.26$. We calculate the space density of RSQs over $1.4 \leq z \leq 5.0$, and find an evolutionary trend below and above the peak of their space density that is comparable to that of faint quasars. By integrating the luminosity function to $M_{1450} = -18$, we find evidence that supports a scenario where RSQs produce only a few per cent of the photons required to ionize the IGM at $z > 3$. By comparing the spatial density of RSQs with that of faint quasars at similar redshifts, we find that RSQs may compose to up $31 \pm 22\%$ of the total faint quasar population. This fraction, within uncertainties, seems to remain constant with redshift. We argue that considering the similarities in

evolutionary trends and declining factors between RSQs and faint quasars, the fainter (optically and radio) RSQs may have properties of RLQs and RQQs. Finally, we discuss several aspects of RSQs such as: the origins of their radio-emission, the environments where these objects reside, and their location in $E1$, $4DE1$ and C_{IV} parameter spaces.

Our work has demonstrated the feasibility of studying the evolution of RSQs using quasar samples compiled using the selection method proposed by Retana-Montenegro & Röttgering (2018). Future studies of the luminosity function of RSQs will benefit from the advent of the new generation of wide-field radio (LOTSS: Röttgering et al. 2011; Shimwell et al. 2017b, 2019; EMU: Norris et al. 2011), optical (LSST: Tyson 2002; LSST Science Collaboration et al. 2009; DES: Flaugher 2005), infrared (WFIRST: Spergel et al. 2013; Euclid: Laureijs et al. 2011), and spectroscopic surveys (EBOSS: Dawson et al. 2016; DESI: DESI Collaboration et al. 2016).

5.A Appendix: A sample of false color RGB ($R=B_W$, $G=R$, $B=I$) images

In this appendix, we present a sample of false color RGB ($R=B_W$, $G=R$, $B=I$) images centered on spectroscopic and photometric quasars. Each image covers $70'' \times 70''$, and the inset size is $7'' \times 7''$. The contours are $[3, 5, 7, 9, 11, 13] \times \sigma$ times the local noise level in the LOFAR (white) and FIRST (purple) images.

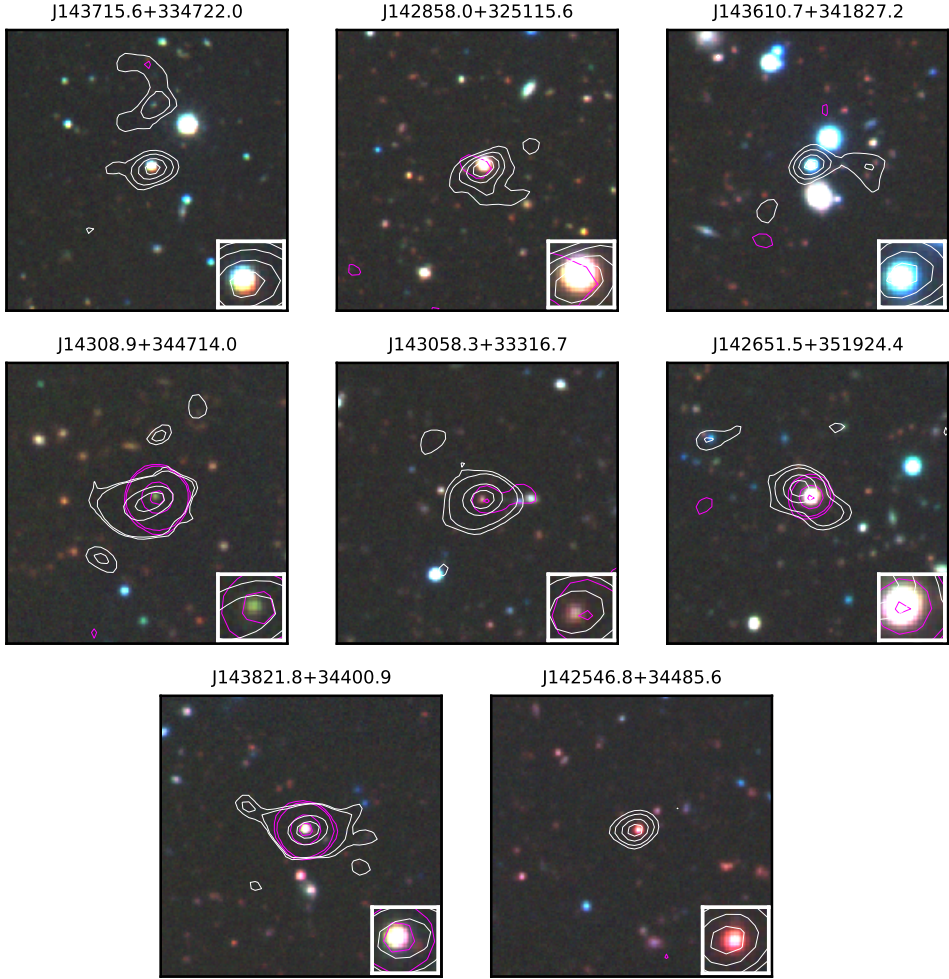


Figure 5.20: A sample of false color RGB ($R=B_W$, $G=R$, $B=I$) images centered on spectroscopic quasars. Each image covers $70'' \times 70''$, and the inset size is $7'' \times 7''$. The contours are $[3, 5, 7, 9, 11, 13] \times \sigma$ times the local noise level in the LOFAR (white) and FIRST (purple) images.

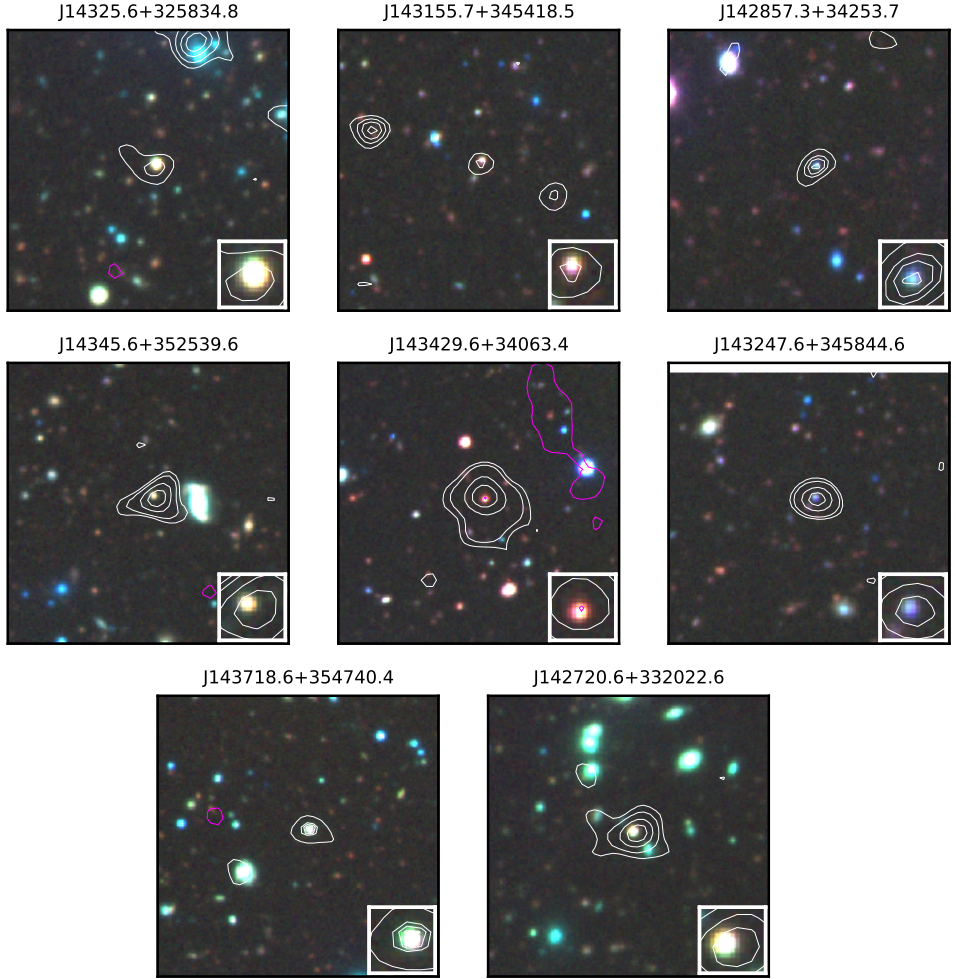


Figure 5.21: A sample of false color RGB ($R=B_W$, $G=R$, $B=I$) images centered on photometric quasars. Each image covers $70'' \times 70''$, and the inset size is $7'' \times 7''$. The contours are $[3, 5, 7, 9, 11, 13] \times \sigma$ times the local noise level in the LOFAR (white) and FIRST (purple) images.