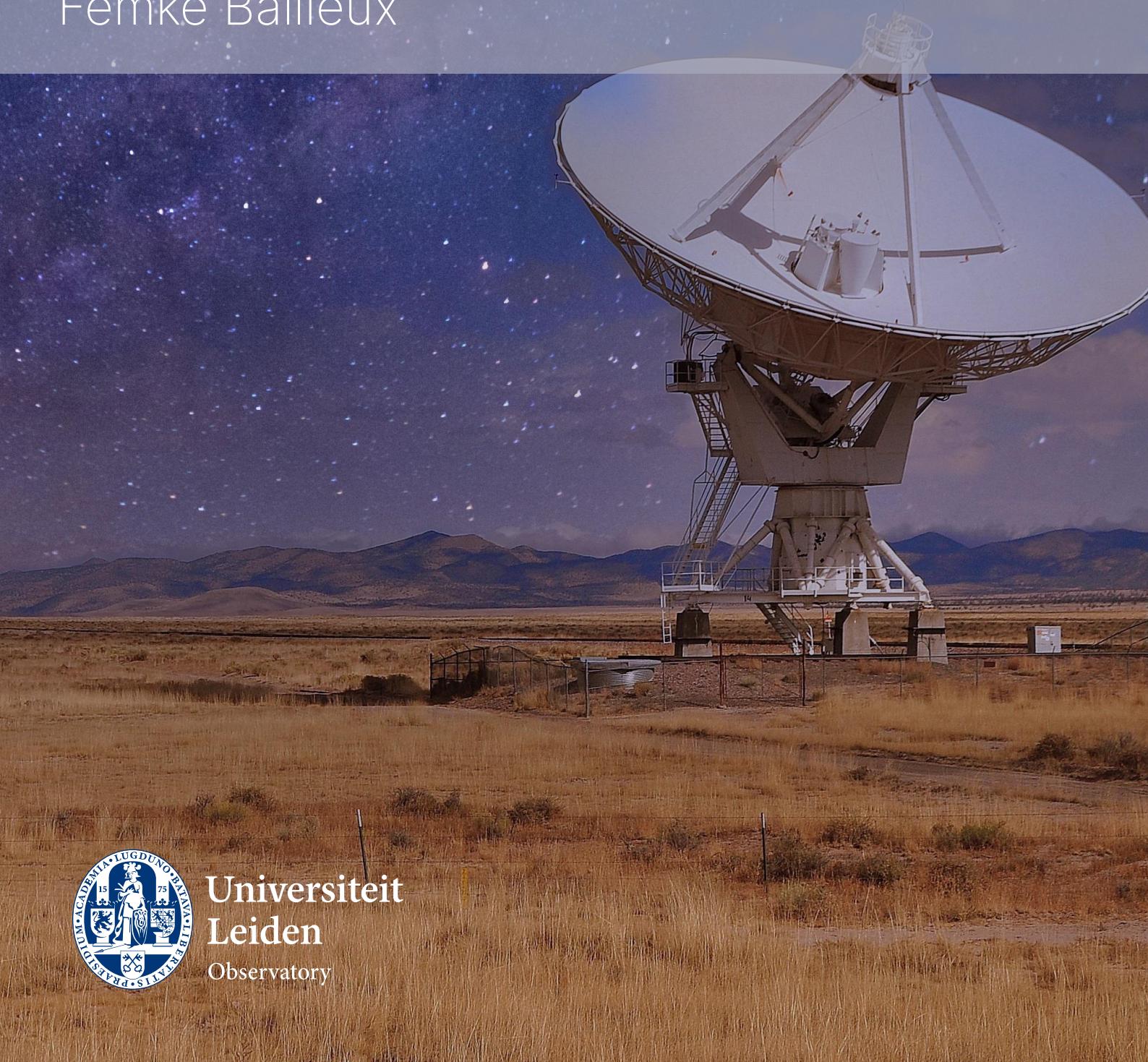


# The evolutionary path of MPS and GPS sources

Introducing the largest sample of peaked-spectrum sources to date

MSc thesis

Femke Ballieux



Universiteit  
Leiden  
Observatory

# The evolutionary path of MPS and GPS sources

Introducing the largest sample of peaked-spectrum sources to date

A thesis submitted in partial fulfillment  
of the requirements for the degree of

Master of Science  
in  
Astronomy

Author:  
Student ID:  
Supervisor:  
Second corrector:  
Project duration:

Femke Ballieux  
s2124149  
Dr. J.R. Callingham  
Prof.dr. H.J.A. Rottgering  
September 2022 – July 2023

Cover image:  
Template style:  
Template licence:

Pierluigi D'Amelio  
Thesis style by Richelle F. van Capelleveen  
Licenced under CC BY-NC-SA 4.0



Universiteit  
Leiden  
Observatory

P.O. Box 9513, 2300 RA Leiden, The Netherlands

# Abstract

With new sensitive wide-field radio surveys being produced, such as the LOFAR Two Meter Sky Survey (LoTSS), the LOFAR LBA Sky Survey (LoLSS), and the Very Large Array Sky Survey (VLASS), the selection of statistically large samples of peaked-spectrum (PS) sources are now possible. These sources are believed to be precursors to large-scale radio galaxies. The increased sample size allows us to robustly compare samples of Gigahertz peaked-spectrum (GPS) sources and Megahertz peaked-spectrum (MPS) sources against one another. By comparing these samples, we can investigate the transition and evolution between GPS and MPS sources, as well as study their lifetimes and occurrence in the universe through redshift space.

We present a sample of 8,975 GPS sources with spectral turnovers near 1400 MHz, and a sample of 522 MPS sources with turnovers near 144 MHz. Our GPS sample is over six times larger than the previously known largest sample of PS sources. To understand the environment of these GPS and MPS sources, we identified their accretion modes using emission line criteria. We find that the fraction of high-excitation radio galaxies is greater in our PS samples relative to the general active galactic nuclei (AGN) population. This suggests that our PS samples represent a population undergoing rapid evolution, agreeing with the commonly proposed youth hypothesis.

We further calculated the Euclidean normalized source counts for both PS samples and found that their shapes match those of general AGN samples, scaled down by a factor  $\sim 45$  for the MPS sample, and a factor  $\sim 42$  for the GPS sample. These offsets imply that both MPS and GPS sources have shorter lifetimes than general AGN, with MPS sources tending to have slightly shorter lifetimes compared to GPS sources.

We also demonstrate that the 144 MHz and 1400 MHz luminosity functions of our MPS and GPS samples follow similar trends as the underlying radio-loud AGN population, but with significant offsets. The GPS sample is scaled down by approximately a factor of  $\sim 14$ , which remains relatively constant over redshift, up until  $z = 3$ . The MPS sample has an average offset of  $\sim 17$ , which could potentially be evolving with redshift. The evolution of the MPS sample can be explained by the youth model, where sources first evolve out of their GPS phase before becoming MPS sources. The offsets also imply that MPS sources generally have slightly shorter lifetimes compared to GPS sources. We find that these shorter lifetimes are likely caused by the MPS phase coinciding with the jet break-out phase of PS sources, which increases the speed at which MPS sources evolve compared to GPS sources.

# Acknowledgements

I would like to extend my heartfelt thanks to all the individuals who supported me throughout this project. Special thanks to my primary supervisor Joe Callingham, who has been a mentor to me during this research. His lessons have been invaluable in becoming a better scientist. I would also like to thank Huub Röttgering for the useful discussions, which have helped me stay on the right track with this work. I further express my gratitude to Martje Slob, who has answered many of my questions during the year, and who has also been a great inspiration and a good friend to me.

I would also like to thank my fellow master students in the HL4.26 office. Shared moments of joy and relaxation, infinite coffee breaks and celebrations made it so much fun to be here. Moments of stress, winter blues and difficult months of grief became bearable with you around.

This research has made use of "Aladin sky atlas" developed at CDS, Strasbourg Observatory, France. This research made use of Astroquery (Ginsburg et al. (2019)) This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency. LOFAR data products were provided by the LOFAR Surveys Key Science project (LSKSP; <https://lofar-surveys.org/>) and were derived from observations with the International LOFAR Telescope (ILT). LOFAR van Haarlem et al. (2013) is the Low Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources), and which are collectively operated by the ILT foundation under a joint scientific policy. The efforts of the LSKSP have benefited from funding from the European Research Council, NOVA, NWO, CNRS-INSU, the SURF Co-operative, the UK Science and Technology Funding Council and the Jülich Supercomputing Centre. This research made use of matplotlib, a Python library for publication quality graphics (Hunter, 2007). The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. This research has made use of the NASA/IPAC Extragalactic Database (NED) operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This research made use of TOPCAT, an interactive graphical viewer and editor for tabular data (Taylor, 2005).

# Contents

<b>Abstract .....</b>	<b>iii</b>
<b>Acknowledgements .....</b>	<b>iv</b>
<b>List of Tables .....</b>	<b>viii</b>
<b>List of Figures .....</b>	<b>ix</b>
<b>1 Introduction .....</b>	<b>1</b>
1.1 Peaked spectrum sources	1
1.2 Youth hypothesis	2
1.3 Frustration hypothesis	3
1.4 Revolution in radio astronomy	4
1.5 PS samples from literature	4
1.5.1 Snellen et al. (1998) .....	4
1.5.2 Callingham et al. (2017) .....	5
1.5.3 Slob et al. (2022) .....	5
1.6 Outline of this work	5
<b>2 Surveys .....</b>	<b>6</b>
2.1 LOFAR surveys	6
2.1.1 LotSS .....	6
2.1.2 LoLSS .....	7
2.2 VLA surveys	7
2.2.1 NVSS .....	7
2.2.2 VLASS .....	7
2.3 Additional wide-field radio surveys	8
2.4 Systematic uncertainties	8
2.5 Sensitivity of samples of peaked-spectrum sources	8

---

<b>3</b>	<b>Sample selection .....</b>	<b>11</b>
3.1	Selecting the LoTSS+NVSS sample	11
3.1.1	Isolated in LoTSS .....	11
3.1.2	Unresolved in LoTSS .....	12
3.1.3	S-code in LoTSS .....	13
3.1.4	Crossmatching LoTSS to NVSS .....	13
3.2	Low-Frequency sample	13
3.3	High-Frequency master sample	13
3.3.1	Unresolved in VLASS .....	14
3.3.2	Selecting the VLASS sample .....	14
3.3.3	Crossmatching VLASS to LoTSS+NVSS .....	16
3.3.4	VLASS completeness .....	16
3.4	Crossmatching to additional radio surveys	19
<b>4</b>	<b>Peaked Spectrum Sources .....</b>	<b>20</b>
4.1	Spectral classification and selection of peaked spectrum sources	20
4.2	Megahertz peaked-spectrum sample	21
4.2.1	Discrepancies among LoLSS data releases .....	21
4.3	Gigahertz peaked-spectrum sample	23
4.4	Extremely steep-spectrum sources	24
4.5	PS candidates in color-color plots	27
<b>5</b>	<b>Accretion mode classification .....</b>	<b>28</b>
5.1	Accretion modes	28
5.2	Accretion mode classification	28
5.2.1	Best & Heckman (2012) classification .....	29
5.2.2	Excitation index classification .....	30
5.2.3	Equivalent width of the [OIII] emission line classification .....	30
5.2.4	[OIII] emission line luminosity .....	31
5.3	The occurrence of HERGs and LERGs in our samples	31
5.4	Statistics of the HERG/LERG populations	33
<b>6</b>	<b>Source counts .....</b>	<b>34</b>
6.1	Source counts of the MPS sample	34
6.2	Source counts of the GPS sample	35
6.3	Comparing to source counts of literature PS samples	36
6.4	Lifetimes derived from source counts	37
<b>7</b>	<b>Luminosity functions .....</b>	<b>39</b>
7.1	Redshift information	39
7.2	The $1/V_{max}$ method	40

7.2.1	Luminosity functions of the MPS sample .....	41
7.2.2	Luminosity functions of the GPS sample .....	44
7.3	<b>Luminosity function offsets</b>	46
7.4	Redshift evolution of the luminosity function offsets	47
<b>8</b>	<b>Discussion</b> .....	<b>51</b>
8.1	Luminosity function offset evolution	51
8.2	Accretion mode classifications	51
8.3	Frequency selection window	52
8.4	PS selection criteria	52
<b>9</b>	<b>Conclusions &amp; future outlook</b> .....	<b>54</b>
9.1	Conclusions	54
9.2	Future outlook	55
9.2.1	Completion of wide-field radio survey used in this research .....	55
9.2.2	Future wide-field radio surveys .....	56
9.2.3	SED modeling .....	56
9.2.4	Luminosity function modeling .....	57
9.2.5	Potential high-frequency peaked spectrum sources .....	57
	<b>Bibliography</b> .....	<b>58</b>
<b>A</b>	<b>1400 MHz source counts</b> .....	<b>64</b>
<b>B</b>	<b>Luminosity functions</b> .....	<b>65</b>

# List of Tables

<b>1.1</b>	An overview of the sub-classes of compact radio sources with spectral turnovers. ....	<b>2</b>
<b>3.1</b>	A summary of the sample selection process. ....	<b>12</b>
<b>5.1</b>	The different classification methods of accretion modes, as well as the number of sources that were identified during each step. ....	<b>29</b>
<b>6.1</b>	The 144 MHz source counts for our samples. ....	<b>38</b>
<b>7.1</b>	The number of sources with reported redshifts in our master sample and PS samples. ....	<b>39</b>
<b>7.2</b>	The mean luminosity function offsets as a function of redshift. ....	<b>48</b>
<b>A.1</b>	The 1400 MHz source counts for our GPS sample. ....	<b>64</b>
<b>B.1</b>	144 MHz luminosity functions for the LF and MPS sample. ....	<b>66</b>
<b>B.2</b>	1400 MHz luminosity functions for the HF and GPS sample. ....	<b>68</b>

# List of Figures

1.1	The relation between the projected linear size and intrinsic turnover frequency for CSS and GPS sources. ....	2
2.1	The observational limits of relevant radio surveys and samples of PS sources. ....	10
3.1	The angular separation to the nearest neighbor of each source in the VLASS QL component catalog. ....	15
3.2	The median RMS noise in VLASS. ....	18
3.3	LOTSS images of some sources that are extended at 144 MHz, and fall below the detection limit at 3 GHz. ....	19
4.1	Color-color plot for the 13,891 sources in our MPS sample. ....	22
4.2	The distribution of $\frac{S_{total,DR1}}{S_{total,PDR}}$ , as a function of SNR in LoLSS. ....	23
4.3	Color-color plot for the 120,041 sources in our GPS sample. ....	24
4.4	The SEDs of sources with extremely steep spectral indices. ....	26
5.1	The [OIII] luminosity against the 1400 MHz luminosities, for all sources in the LF and HF master samples that were classified as either HERG or LERG. ....	32
6.1	The Euclidean normalized differential source counts for our GPS and MPS samples. ....	35
7.1	The distribution of redshifts for our MPS and GPS samples. ....	40
7.2	The 144 MHz luminosity functions for our LF and MPS samples. ....	43
7.3	The 1400 MHz luminosity functions for our HF and GPS samples. ....	45
7.4	The offset between the LF and MPS samples luminosity functions. ....	46
7.5	The offset between the HF and GPS samples luminosity functions. ....	47
7.6	The mean luminosity function offsets as a function of redshift. ....	49
7.7	The mid-plane density at various stages of time, in a simulation by Sutherland & Bicknell (2007). ....	50

# 1. Introduction

## 1.1 Peaked spectrum sources

The study of Active Galactic Nuclei (AGN) poses a number of interesting questions, such as what triggers AGN activity, how jets form, how AGN evolve, and the effect an AGN has on its host (O'Dea & Saikia, 2021). Peaked spectrum (PS) sources are compact sources that show a peak in their radio spectrum and have been hypothesized to be the young progenitors of radio-loud AGN.

PS sources can be found with a range of intrinsic turnover frequencies  $\nu_{\text{rest,turn}}$ . A correlation between the projected linear size  $l$  in kpc, and the turnover frequency in GHz of these sources has been identified by O'Dea (1998), as shown in Figure 1.1. The relation is given by

$$\log \nu_{\text{rest,turn}} \approx -0.21(\pm 0.05) - 0.65(\pm 0.05) \log l, \quad (1.1)$$

or  $\nu_{\text{rest,turn}} \propto l^{-0.65}$ . This indicates that PS sources with the smallest linear sizes peak at higher frequencies.

PS sources are typically classified into the sub-classes Gigahertz Peaked Spectrum (GPS) sources, High-Frequency Peaked (HFP) sources, and Megahertz Peaked Spectrum (MPS) sources, which differ in terms of linear size and turnover frequency. Compact Steep Spectrum (CSS) sources are also thought to be associated with PS sources, with their spectral turnover expected to be below the common observing frequency of  $\sim 1$  GHz (O'Dea & Saikia, 2021).

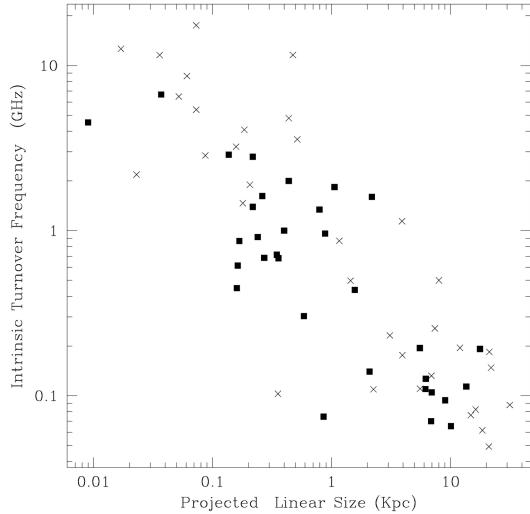
HFP sources are sources that peak above 5 GHz (O'Dea & Saikia, 2021). They are the most compact PS sources (Dallacasa et al., 2001), as follows from Equation 1.1.

GPS radio sources have spectral peaks between 500 MHz and 5 GHz (Gopal-Krishna et al., 1983). They have typical linear sizes  $\lesssim 1$  kpc, and are powerful, with  $\log P_{1.4\text{GHz}} \gtrsim 25\text{Whz}^{-1}$  (O'Dea, 1998).

CSS sources are about as powerful as GPS sources, with larger linear sizes (1-20 kpc; Fanti et al., 1990), and spectral peaks at  $< 500$  MHz. Turnovers below  $\sim 30$  MHz cannot be measured robustly due to the ionospheric cutoff at 10 MHz (O'Dea, 1998).

MPS is a term sometimes used to refer to PS sources that peak below 1 GHz in the observer's frame. This population is likely to be a combination of relatively nearby CSS and GPS sources, or compact high-frequency peaked sources further away, whose turnover frequency has been shifted to low frequencies due to the cosmological redshift (O'Dea & Saikia, 2021).

In this work, we consider two samples of PS sources, which we will refer to as our GPS and MPS samples. These names refer to the observer's frame turnover frequencies



**Figure 1.1:** This figure, from O'Dea (1998), shows the relation between the projected linear size and intrinsic turnover frequency for the Fanti et al. (1990) CSS and Stanghellini et al. (1998) GPS sample. Galaxies are represented by solid squares, and quasars are represented by crosses. The formula describing the fitted relation is given by Equation 1.1.

used for the selection of these samples. A general overview of the different types of PS sources as adapted in this work are indicated in Table 1.1.

## 1.2 Youth hypothesis

There are two main hypotheses as to what causes the spectral turnovers and small linear sizes of PS sources. One of these mechanisms is that PS sources are the young precursors of Fanaroff-Riley I and II (FRI and FRII) galaxies. PS sources have shown evidence of compact double-lobed structures. Furthermore, a relationship between radio power and linear size has been found for PS sources (e.g. Kunert-Bajraszewska et al. 2010, An & Baan 2012).

Acronym	Text	Definition
HFP	High-Frequency Peaked	$\nu_p > 5\text{GHz}$
GPS	Gigahertz Peaked Spectrum	$0.5 \lesssim \nu_p \lesssim 5\text{GHz}$
MPS	Megahertz Peaked Spectrum	$\nu_p < 1\text{GHz}$
PS	Peaked-Spectrum	All GPS, MPS, and HFP
CSS	Compact Steep Spectrum	$1 \leq l \leq 20\text{kpc}$

**Table 1.1:** An overview of the sub-classes of compact radio sources with spectral turnovers. The categorization can be somewhat arbitrary as these sources are hypothesized to be part of a continuum of sources. The classification is mostly based on observational properties such as peak frequency  $\nu_p$ , and linear size  $l$ . In this work, we study MPS and GPS sources, which we will generally refer to as PS sources.

From Equation 1.1, we find that if this youth model is generally true, we would expect HFP sources to be the youngest PS sources, evolving into GPS sources, and then into MPS sources. After the MPS phase, PS sources would grow into FRI and FRII sources. (e.g. Carvalho 1985, Kunert-Bajraszewska et al. 2010). Evidence from spectral break modeling (Callingham et al., 2015) and the motions of lobe hot spots (Owsianik & Conway 1998, Kaiser & Best 2007) supports this theory.

In the youth hypothesis, synchrotron self-absorption (SSA; Snellen et al. 2000, de Vries et al. 2009) is likely the reason for the spectral turnover. In this absorption mechanism, relativistic electrons internal to the source re-absorb emission. The turnover is caused because the source cannot have a brightness temperature exceeding the plasma temperature of the non-thermal electrons. The turnover frequency is the frequency at which relativistic electrons and emitted synchrotron photons have the highest probability of scattering (Tingay & de Kool, 2003). SSA is successful in explaining the relation between turnover frequencies and linear sizes (O'Dea, 1998).

If we assume all PS sources grow into an FRI or FRII radio galaxies, we would expect to find an offset in luminosity functions relative to the difference in the lifetime of the evolutionary phase. We would also expect that PS sources are quickly evolving sources, and therefore generally have a high-excitation accretion mode.

### 1.3 Frustration hypothesis

An inconsistency in the youth hypothesis is that there is an overabundance of PS sources relative to large AGN (e.g. Kapahi 1981, Peacock & Wall 1982, An & Baan 2012). Furthermore, studies of the emission measure of ionized gas implied by line emission from CSS sources do not appear consistent with the SSA mechanism (van Breugel et al., 1984).

This leads us to the second possible hypothesis, which states that the small linear sizes and spectral turnovers are caused by sources being in extremely dense environments. Due to the dense environments, sources get 'frustrated', and they are not able to grow to larger spatial scales. The absorption mechanism associated with the frustration hypothesis is Free-Free Absorption (FFA; Bicknell et al. 1997, Peck 1999, Callingham et al. 2015). The absorption of radio emission is caused by an external ionized plasma, where radio photons get absorbed by free electrons in the presence of positive ions (Longair, 2011). Even though standard homogeneous FFA cannot reproduce the relation between linear size and turnover frequency, adapted versions of FFA can. The successful inhomogeneous FFA model by Bicknell et al. (1997) reproduces the observed relation between linear size and turnover frequency.

Evidence for the frustration hypothesis is that the radio morphologies of CSS sources indicate strong interactions between the jets of a source and their environments (Wilkinson et al. 1984, van Breugel et al. 1984, Kunert-Bajraszewska et al. 2010). Furthermore, extended emission around PS sources has been observed, indicating multiple epochs of activity (Baum et al. 1990, Stanghellini et al. 1990). In individual PS sources, unusually high densities have been found, which also indicates FFA emission would be dominant (Peck 1999, Callingham et al. 2015, Sobolewska et al. 2019).

It is always possible that a source is both young and frustrated, and individual PS sources could be associated with either or both mechanisms. We are investigating whether one mechanism dominates over the other in a general sample of PS sources. Differentiating between the two mechanisms can be done by low-frequency spectral modeling, where spectral measurements below the turnover frequency are required. This

allows us to specify the absorption mechanism associated with a source. In the most extreme case, emission where SSA dominates cannot reach a spectral index larger than 2.5 (O'Dea, 1998), therefore any spectral index steeper than this indicates the source must be FFA dominated. Such a characteristic allows us to unambiguously identify frustrated sources. Furthermore, low-frequency luminosity functions would be invaluable in studying the evolution of the luminosity of PS sources, relative to that of a general sample of AGN.

## 1.4 Revolution in radio astronomy

In the past decades, a revolution has taken place in wide-field radio astronomy. Surveys of unparalleled sensitivity, resolution, and sky area have been made. In this work, we will make use of all-sky surveys from the LOw Frequency ARray (LOFAR; van Haarlem et al., 2013) and the Karl G. Jansky Very Large Array (VLA).

We use the LOFAR Two-meter Sky Survey (LoTSS; Shimwell et al., 2022), which is a wide-field radio survey at 120–168 MHz. Due to the high sensitivity of this survey, it has been predicted that we could observe PS sources with radio powers  $< 10^{25} \text{ W Hz}^{-1}$ . This is a massive improvement compared to earlier samples at 5 GHz with median radio luminosities  $10^{26} - 10^{27} \text{ W Hz}^{-1}$ . The improved sensitivity of wide-field radio surveys also allows us to probe higher redshifts, which is vital to study the evolution of PS sources.

We will also make use of the LOFAR LBA Sky Survey (LoLSS; de Gasperin et al., 2023), which probes the sky at 42–66 MHz. Such low-frequency observations are crucial in identifying and studying for example MPS sources, which are an important phase in the youth hypothesis.

Other important radio surveys used in this research are the NRAO VLA Sky Survey (NVSS; Condon et al., 1998) and the Very Large Array Sky Survey (VLASS; Lacy et al., 2020), where VLASS is still in progress. It has a resolution of 2.5'' at 3 GHz, and a median noise level of  $\sim 0.2 \text{ mJy}$ . In Chapter 2 we will describe these surveys and their sensitivity more in depth.

## 1.5 PS samples from literature

Recent work on large samples of PS sources has been done by Snellen et al. (1998), Callingham et al. (2017), and Slob et al. (2022). We will compare the results obtained in our work to these samples, and briefly introduce them below.

### 1.5.1 Snellen et al. (1998)

We will be comparing our samples of PS sources to the work of Snellen et al. (1998). In this work, the Westerbork Northern Sky Survey (WENSS; Rengelink et al., 1997), was used to identify 47 GPS sources, with peak frequencies between 500MHz and 15GHz. The peak flux densities in this sample, ranging from 40 to 900 mJy, are one to two orders of magnitude lower than in earlier samples. They find an offset in source counts of  $\sim 250$ , which would indicate that PS sources have lifetimes 250 times shorter than that of a typical large-scale radio source. However, they warn the reader that source counts should always be interpreted cautiously since they do not show any redshift evolution between samples.

### 1.5.2 Callingham et al. (2017)

From Callingham et al. (2017), we obtain a sample of 1,483 PS sources, that were selected from the GLEAM (Wayth et al., 2015) survey. These PS sources peak between 72 MHz and 1.4 GHz. At the time of publishing, the GLEAM survey was the widest fractional band-width wide-field sky survey, which made it ideal for identifying PS sources at low frequencies. At the time of publishing, this sample doubled the number of known PS candidates. They highlight that some of the GPS in their sample show multiple epochs of activity. They find no relation between observed peak frequency with redshift, which is consistent with the PS sample consisting of both low-redshift CSS sources and high-redshift GPS sources.

### 1.5.3 Slob et al. (2022)

Slob et al. (2022) have produced a sample of 373 PS sources, selected using the second data release (DR2) of LoTSS, the preliminary data release (PDR) of LoLSS, and NVSS. Due to the sensitivity of the LOFAR surveys, their work has increased the number of PS sources found in their survey area by a factor of 50, and at the time of publishing the sample contained PS sources with the lowest luminosity by an order of magnitude. Using optical line emission criteria, it was found that this sample contained mostly high-excitation galaxies, hinting at a quickly evolving population. From source counts it was found that they are scaled down a factor of 40 compared to radio-loud AGN, which would imply that the lifetimes of PS sources are 40 times lower compared to massive radio galaxies if they have similar luminosity functions. The luminosity functions were computed, and it was found that the PS luminosity functions have the same shape as those of unresolved radio-loud AGN, but have been shifted down by a factor of 10. This was seen as strong evidence for the youth hypothesis, where the PS sources evolve into radio-loud AGN. Locally there is a surplus of PS sources, which could be caused by frustrated PS sources that do not evolve into radio-loud AGN.

## 1.6 Outline of this work

To identify a spectral turnover, radio flux density measurements at, at least, three different frequencies are required. In this work, we select two master samples and find all PS sources in these samples. This gives us a sample of MPS and a sample of GPS sources. We use these samples to study the evolution between the MPS and GPS samples, which helps us test the youth hypothesis. We discuss the surveys used in this work in Chapter 2. In Chapter 3, we discuss the selection of the master samples, and in Chapter 4 we define our samples of PS sources. In Chapter 5 we study the accretion modes of our samples to investigate whether our samples of PS sources evolve faster than a general sample of AGN. In Chapter 6 we construct Euclidean normalized source counts for our samples, and in Chapter 7 we study the luminosity functions. These two chapters help us understand the number density of PS sources and the evolution thereof. In this paper we adopt  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.27$  and  $\Omega_\Lambda = 0.73$  for a standard Lambda cold dark matter cosmological model (Hinshaw et al., 2013).

# 2. Surveys

Our research makes use of wide-field radio surveys, many of which are recent surveys at low-frequencies. Radio surveys in the past decades have seen an enormous increase in sensitivity, which allows us to select a sample of PS sources larger and more complete than previous studies. The surveys we use for the bulk of this work are two LOFAR surveys, (LoLSS at 54 MHz and LoTSS at 144 MHz) and two VLA surveys (NVSS at 1400 MHz and VLASS at 3 GHz). By combining these four surveys, we are able to construct two samples of PS sources, with different turnover frequency windows. This allows us to study the evolution from GPS to MPS sources. A quick overview of the wide-field radio surveys used in our work is detailed below.

## 2.1 LOFAR surveys

LOFAR (van Haarlem et al., 2013) is a large radio telescope operating at low frequencies. It consists of 52 dipole antenna stations located around the village of Exloo, spread out in the Netherlands, and spread out across Europe. The stations consist of high-band (HBA, 110–250 MHz) and low-band (LBA, 10–90 MHz) antennae. Its dense core has baselines of 68 m, and the international baselines have baselines up to 1000 km, which gives LOFAR good sensitivity and angular resolution at low frequencies. LOFAR is also an important pathfinder for the Square Kilometer Array (SKA; Dewdney et al., 2009) since LOFAR is one of the first to deal with the transport, processing, and storage of very large amounts of radio data.

### 2.1.1 LoTSS

LoTSS (Shimwell et al., 2022) is a deep wide-field radio survey, made with the LOFAR HBA at 120–168 MHz, and has a central frequency of 144 MHz. It aims to observe the whole northern sky. Data from the second data release (DR2) was used in this work, which consists of observations from 27 % of the northern sky. The coverage is split into two regions centered at approximately 12h45m +44° 30' and 1h00m +2° 00', spanning  $4178 \text{ deg}^2$  and  $1457 \text{ deg}^2$ , respectively. The DR2 catalog contains 4,396,228 radio sources, derived from the total intensity (Stokes I) maps. It reaches a median rms sensitivity of  $83 \mu\text{Jy beam}^{-1}$ , a resolution of 6'', and is 90 % complete at  $0.8 \text{ mJy beam}^{-1}$ . The data reduction process is explained in detail in Shimwell et al. (2022).

In-band spectra are also available for the sources in LoTSS, at central frequencies 128, 144, and 160 MHz. The in-band spectra provide information about spectral properties along the 48 MHz wide band. However, these in-band spectra are not reliable for most

sources, due to the narrow frequency range, combined with non-negligible uncertainties in the alignment of the flux density scale of the in-band images. Therefore, the in-band spectra were only used in this work as a visual guide to confirm spectral fits and were not used to select PS sources.

Optical associations and redshifts are available for a part of the LoTSS catalog (Williams et al. (2019), Duncan et al. (2019), Hardcastle et al. in prep.). More information about the optical associations can be found in Chapter 7.

### 2.1.2 LoLSS

LoLSS (de Gasperin et al., 2023) is a sensitive wide-area survey at 41–66 MHz, centered at 54 MHz, which eventually will cover the whole sky above declination 24°. In this work, data from the first data release (DR1) was used, which has a typical sensitivity of  $1.55 \text{ mJy beam}^{-1}$ , which is four times higher relative to the preliminary data release (PDR; de Gasperin et al., 2021). So far,  $650 \text{ deg}^2$  sky area has been released, centered around the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX; Hill et al., 2008) Spring Field with  $11 < \text{RA} < 16\text{h}$  and  $45 < \text{Dec} < 62^\circ$ . The LoLSS DR1 catalog consists of 42,463 sources, with an angular resolution of 15'', and 95% completeness at 11 mJy. Separate in-band spectra, at central frequencies 44, 48, 52, 56, 60, and 64 MHz, are also available. The in-band spectra were only used for visual confirmation that spectral fits seem reliable, and were not used for selecting PS sources.

## 2.2 VLA surveys

The VLA is a large multi-dish radio telescope, consisting of 28 dish antennas, each 25 m in diameter.

### 2.2.1 NVSS

NVSS (Condon et al., 1998) is a 1.4 GHz sky survey, covering the sky north of a declination  $-40^\circ$  (82% of the celestial sphere). Observations were completed between September 1993 and October 1996. The NVSS catalog consists of 1,773,484 sources, has a resolution 45'', and is 99% complete at 3.4 mJy.

### 2.2.2 VLASS

VLASS (Lacy et al., 2020) is a wide-field radio survey at 2–4 GHz, with an angular resolution of 2.5'', covering the whole sky above a declination of  $-40^\circ$  ( $33,885 \text{ deg}^2$ ). It has a  $1\sigma$  goal of  $70 \mu\text{Jy beam}^{-1}$ . The B configuration of the VLA was used, with a maximum antenna separation of 10 km. These long baselines are very suitable for obtaining a high resolution and are mainly useful for observing point sources and compact structures. The lack of short baselines makes it more difficult to detect diffuse emission. Since PS sources are compact objects, the B configuration of the VLA is ideal for our research.

VLASS will observe the entire sky three times, which allows for studies of variable and transient sources. By the time of writing, the first and second epochs have been observed.

Awaiting high-quality single-epoch images, quick look (QL) images for the first two epochs have been released. The rapid CLEANing of these images limits the quality of these images, but are reliable for our science since the majority of the VLASS sources we

are interested in are high signal-to-noise point sources (Gordon, 2023). In our research, the component catalog of epochs 2.1 and 2.2 was used, and 1.1 and 1.2 were left out. We only use the second epoch because between 1.1 and 1.2, the rapid CLEAN algorithm used was updated. There are a number of issues with epoch 1.1, key issues being a systematic under-measurement of flux density value in the QL images, issues with astrometry, and ghost images of bright sources. These issues have been solved for the second epoch, so we disregard epoch 1 (Gordon, 2023). The VLASS QL catalog used in this research consists of 2,446,020 components.

## 2.3 Additional wide-field radio surveys

After the initial selection of our master samples, as we will describe in Chapter 3, more radio surveys were crossmatched to the sample for further data analysis. The additional radio surveys are the TIFR GMRT Sky Survey Alternative Data Release 1 (TGSS; Intema et al., 2017) at 150 MHz, the Faint Images of the Radio Sky at Twenty-centimeters (FIRST; Becker et al., 1995) at 1.4 GHz, the Westerbork Northern Sky Survey (WENSS; Rengelink et al., 1997) at 326 MHz, and the Very Large Array Low-Frequency Sky Survey Redux (VLSSr; Lane et al., 2014) at 74 MHz. These surveys were used for visual inspection of the spectral index fit, not for the selection of PS sources, or the corresponding source counts and luminosity functions.

We obtained optical photometry and emission lines from the Sloan Digital Sky Survey (SDSS). SDSS is a large optical sky survey, taken with a 2.5 m telescope at Apache Point Observatory (APO) in Sunspot, New Mexico. We used SDSS data to calculate luminosity functions and classify the accretion modes of the sources in our samples. More information about how this data was obtained will be given in Chapter 7.

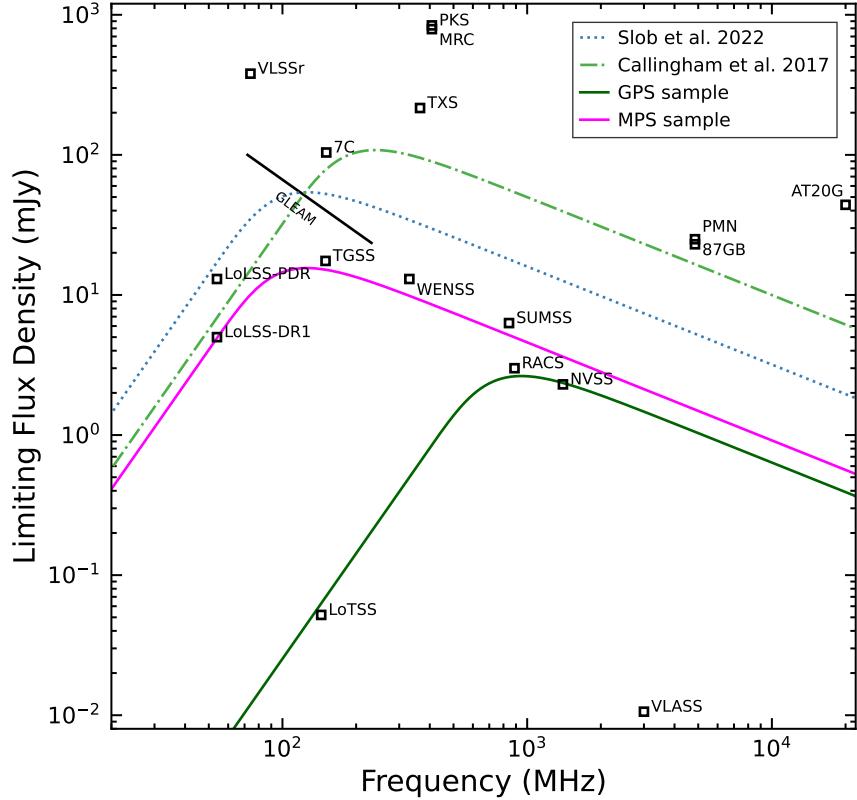
## 2.4 Systematic uncertainties

In this work, uncertainties are dominated by systematic uncertainties between surveys, rather than individual uncertainties on sources. Therefore, we assumed the uncertainty on flux measurements is always 10% of the flux density in LoTSS, NVSS, and VLASS. This allowed us to account for the systematics between surveys. For sources in LoLSS, we are studying a lower SNR regime, so we used the reported flux density uncertainties for LoLSS. The precision of LoLSS is 6% (de Gasperin et al., 2023).

## 2.5 Sensitivity of samples of peaked-spectrum sources

To select our PS sources, we used the most sensitive wide-field radio surveys to date, which allowed us to observe fainter PS sources than ever before. The sensitivity of radio surveys and PS samples is illustrated in Figure 2.1. We indicate the lowest reported flux density for many wide-field radio surveys from the last three decades. We also plot the SEDs of the faintest typical SSA PS sources that could be observed in different PS samples from literature. The solid lines correspond to our GPS and MPS samples. The parameter space of this work reaches much lower flux densities than other samples of PS sources. LoTSS, NVSS, and VLASS are the three most sensitive surveys known to date, and LoLSS is the lowest frequency wide-field survey and is also among the most sensitive in the sample.

From Figure 2.1, we can infer which survey determines the flux density limit of each sample. We find that for our MPS sample, LoLSS determines the flux density limit and for the GPS sample, NVSS limits the flux density.



**Figure 2.1:** The observational limits of different radio surveys and PS samples. The lowest catalogued component flux density reported for each survey used in this research, and other relevant large radio surveys, is indicated in the plot. The GaLactic and Extragalactic All-Sky MWA Survey (GLEAM; Wayth et al., 2015) is represented by a line since it has variable limiting flux densities at different observing frequencies. The curves represent the observational limit for different PS samples, that have been introduced in Section 1.5. The Callingham et al. (2017) sample is represented by an SSA model with a faintest peak flux density of 160 mJy at a turnover frequency of 190 MHz. The Slob et al. (2022) sample is represented by an SSA model with a faintest peak flux density 80 mJy at a turnover frequency of 100MHz. We indicate the observational limits of our MPS and GPS samples with SSA models with a faintest flux density of 23 mJy at a turnover frequency of 100 MHz, respectively 3.9 mJy at 750 MHz. It should be noted that these curves merely have an illustrative purpose. Since PS sources are identified by at least three flux density measurements at different frequencies, we cannot fit the exact turnover frequency. The sources in our sample have a wide range of possible turnover frequencies and corresponding flux densities. In this plot, we indicate the faintest PS sources that could theoretically be identified in our master samples, if the turnover frequency was approximately at the central radio survey of each sample. In the figure, multiple wide-field radio surveys were indicated that have not been mentioned before; Parker (PKS; Wright & Otrupcek, 1990), Molonglo Reference Catalogue (MRC; Large et al., 1991), MIT-Green Bank 5 GHz (87GB; Gregory & Condon, 1991), Parker-MIT-NRA (PKS; Wright et al., 1994), Texas Survey (TXS; Douglas et al., 1996), Sydney University Molonglo Sky Survey (SUMSS; Mauch et al., 2003), Cambridge 7C (Hales et al., 2007), Australia Telescope 20 GHz (AT20G; Murphy et al., 2010), and the RAPID ASKAP Continuum Survey (RACS; Hale et al., 2021).

# 3. Sample selection

In this work, we selected two different master samples, from which we defined two different samples of PS sources. We start by selecting all isolated, unresolved sources in the LoTSS catalog. We then crossmatch this sample to the NVSS catalog. We call this combined sample the LoTSS+NVSS sample. By crossmatching the LoTSS+NVSS sample to other radio surveys, we identify two different master samples. The Low-Frequency master sample (LF) was identified by crossmatching the LoTSS+NVSS sample to the LoLSS catalog. The High-Frequency master sample (HF) was identified by crossmatching the LoTSS+NVSS sample to all VLASS sources that were isolated. Both master samples were then crossmatched to additional wide-field radio surveys, which were only used for visual comparisons.

In Chapter 4 we identified PS sources in the LF and HF master samples using the spectral indexes, which defined our MPS and GPS samples, respectively. An overview of each step in the sample selection can be found in Table 3.1, which includes the number of sources in our samples after each selection step. In the following sections, we provide more details on the sample selection that was outlined above.

## 3.1 Selecting the LoTSS+NVSS sample

### 3.1.1 Isolated in LoTSS

The selection of the master samples started with the LoTSS catalog, as this survey is included in both master samples and has a higher resolution than NVSS. The LoTSS DR2 catalog contains 4,396,228 sources.

To ensure that source confusion did not impact the derived spectra, we made a selection based on whether the sources are isolated. We deemed a source to be isolated in LoTSS where there is no other source within  $47''$ , corresponding approximately to the resolution of NVSS. Of all surveys used in this research, NVSS has the lowest resolution. Therefore, if a source is isolated in a radius of  $47''$  in LoTSS, it will be a single source in NVSS or other surveys with a lower resolution than LoTSS.

Deconvolution errors around very bright ( $\gtrsim 200\text{ mJy}$ ) sources can produce artifacts that can cause it to appear as if the source is not isolated. In order to include these bright sources that are surrounded by complex noise structures, we included a source in the master sample if the neighboring source had a total flux density less than 10 % of the brightest source. Even if a faint component near a bright source is not an artifact but part of the bright source that is resolved, we assumed that the bulk of the flux density was captured in the bright component. The LoTSS catalog has 2,676,735 sources that

LoTSS+NVSS sample selection			
Step	Number of sources		
Total LoTSS catalogue	4,396,228		
Isolated 47'' in LoTSS	2,676,735		
Unresolved in LoTSS	2,493,574		
S-code 'S' or 'M' in LoTSS	2,493,565		
Has a NVSS counterpart	157,029		

LF sample selection		HF sample selection	
Step	Number of sources	Step	Number of sources
Total LoLSS catalog	42463	Total VLASS catalog	2,995,271
		Recommended VLASS catalog	2,446,020
		Isolated 47'' in VLASS	1,397,173
		S-code 'S' or 'M' in VLASS	1,396,989
LF master sample	13,891	HF master sample	120,041
MPS sample	522	GPS sample	8,975

**Table 3.1:** A summary of the sample selection process. The number of sources in our sample that remain after each selection step is indicated.

are isolated within 47''.

### 3.1.2 Unresolved in LoTSS

We only considered sources that are unresolved in LoTSS because we expected almost all PS sources to be unresolved in LoTSS. At  $z = 0.1$ , PS sources that peak around 150 MHz are expected to be smaller than 8 kpc from equation 1.1, which corresponds to an angular size of  $\sim 4''$ . Such sources are unresolved in LoTSS, which has a resolution of 6''. Therefore, we were able to remove any source that is resolved in LoTSS from our sample without biasing against selection PS sources.

We used the criteria for identifying resolved sources in LoTSS that were described by Shimwell et al. (2022). If the data was perfectly calibrated and deconvolved, and there was no noise, you would expect for any perfectly unresolved source that  $\frac{S_I}{S_P} = 1$ . Here  $S_I$  is the integrated flux density, and  $S_P$  is the peak flux density. In reality, noise and deconvolution issues exist, and there is a correlation between  $\sigma_I$  and  $\sigma_P$ , which are the standard deviations of the fitted integrated and peak flux densities. Therefore the distribution of  $R = \ln(\frac{S_I}{S_P})$  is no longer a simple Gaussian distribution.

Shimwell et al. (2022) defined an envelope that encompasses the 99.9 percentile of the  $R$  distribution. They define a source as unresolved if  $R$  is larger than or equal to

$$R_{99.9} = 0.42 + \left( \frac{1.08}{1 + (\frac{SNR}{96.57})^{2.49}} \right). \quad (3.1)$$

Here, the SNR is defined as  $\frac{S_I}{\sigma_I}$ . All sources that have  $R < R_{99.9}$  are deemed unresolved, so we removed them from the sample. While this cut will potentially leave some marginally resolved sources in the sample, a balance must always be found between completeness and reliability. Sources where  $R \approx R_{99}$  cannot be definitively classified as resolved or

unresolved. Fortunately, slightly resolved sources in LoTSS are still reliably measured by our source-finding algorithm, so they can be left in the sample.

Selecting only unresolved sources in LoTSS removed 6.8% of the sources from our sample, leaving us with 2,493,574 sources.

### 3.1.3 S-code in LoTSS

The above criterion has excluded the majority of the resolved sources, but it is based on the 99.9th percentile of the distribution. Some resolved sources might still be present in the sample. In order to exclude as many unresolved sources as possible from our sample, we removed all sources with *PyBDSF* (Mohan & Rafferty, 2015) S-code 'C'. These correspond to multiple single-Gaussian sources in a source-fitting island. By excluding these sources, we are left only with sources that have S-code 'S', which are isolated sources fitted by a single Gaussian distribution, or S-code 'M', which are sources fit by multiple Gaussian distributions. The latter are kept in the sample because the fitting to multiple Gaussian distributions might be because of deconvolution errors around the brightest sources. We removed 9 'C' sources from our sample, leaving 2,493,565 sources.

### 3.1.4 Crossmatching LoTSS to NVSS

The isolated, unresolved sources in LoTSS are crossmatched to NVSS, with a cross-matching radius of 15''. No selection had to be made for NVSS, since it has a lower resolution than LoTSS. All isolated and unresolved sources in LoTSS will be single unresolved sources in NVSS. The LoTSS+NVSS sample we have now defined contains 157,029 sources and forms the basis for both master samples defined in this work. Below, the LF and HF master samples are introduced separately.

## 3.2 Low-Frequency sample

We crossmatched the LoTSS+NVSS sample defined above to LoLSS DR1 with a cross-matching radius of 15'', which forms the LF sample. It contains 13,891 sources. We did not make any selections to the LoLSS survey, since it has a lower resolution than LoTSS. Because all sources in our LF sample are isolated and unresolved in LoTSS, we considered only a single unresolved source in LoLSS.

LoLSS is the survey with the lowest sensitivity of all surveys in the LF sample, therefore not all sources in the LoTSS+NVSS sample will have a counterpart in LoLSS. Furthermore, LoLSS has the limiting sky coverage, which reduces the number of sources in LoTSS+NVSS with a match in LoLSS even further.

## 3.3 High-Frequency master sample

We defined the HF sample by crossmatching the LoTSS+NVSS sample to the VLASS QL component catalog. Since VLASS has a higher resolution compared to LoTSS, unresolved or slightly resolved sources in LoTSS might be cataloged as multiple components in VLASS. If we then crossmatch a source in LoTSS+NVSS to a single component in VLASS, this might cause us to miss a significant portion of the flux density. Crossmatching is most reliable if only unresolved sources in VLASS are considered, as then we will be sure all of the flux is captured. We will show that the number of PS sources that get

resolved in VLASS is negligible. We are then able to remove all sources that are resolved in VLASS, without removing any PS sources, ensuring our sample will be as complete as possible. After we have confirmed that we do not remove any PS sources if we select only unresolved sources in VLASS, we define when sources in VLASS are unresolved. Only then we crossmatch the unresolved VLASS sources to the LoTSS+NVSS sample, which forms the HF master sample.

### 3.3.1 Unresolved in VLASS

PS sources in our HF sample be only be classified as PS if they peak roughly between 250 MHz and 2 GHz in the observer's frame,  $\nu_{\text{obs,turn}}$ . Using the relation between the rest frame frequency and the observer's frame frequency, as defined in Equation 3.2, we calculate which turnover frequency the sources identifiable in our sample could have in the rest frame:

$$\nu_{\text{rest}} = \nu_{\text{obs}}(1 + z). \quad (3.2)$$

At a typical redshift of  $z = 1$ , where the number of AGN peaks in the Universe, the range of intrinsic turnover frequencies we can identify is  $500\text{MHz} \leq \nu_{\text{rest,turn}} \leq 4000\text{MHz}$ . In the nearby universe, at a redshift of  $z = 0.1$ , a source will be identifiable as PS if the intrinsic turnover frequency is  $275\text{MHz} \leq \nu_{\text{rest,turn}} \leq 2200\text{MHz}$ .

We determine whether the largest PS sources that the HF sample is sensitive to are unresolved. Using Equation 1.1, we find that the largest linear sizes correspond to the lowest intrinsic turnover frequencies. For  $z = 1$ , the largest PS sources identifiable have a turnover frequency of  $\nu_{\text{rest,turn}} = 500\text{MHz}$ , which corresponds to  $I = 1.38\text{kpc}$ . For the more extreme redshift of  $z = 0.1$ , the PS sources with the largest linear sizes are found for  $\nu_{\text{rest,turn}} = 275\text{MHz}$ , which corresponds to  $I = 3.46\text{kpc}$ .

We convert this physical size to an angular size using the cosmological model outlined in the introduction. For  $z = 1$ , the scale between angular and linear distance is  $8.156\text{kpc}/''$ . Therefore the maximum angular size of a PS source identifiable in our sample, at this redshift, is  $\theta = 0.0466''$ . These sources are thus impossible to resolve with VLASS. For  $z = 0.1$ , the scale is  $1.848\text{kpc}/''$ , so we find a maximum angular size of  $\theta = 1.87''$ , which is also unresolved in VLASS.

The smallest redshift at which the largest GPS sources identifiable in our sample would be resolvable in VLASS is at  $z \approx 0.075$ . This corresponds to a very small volume of space, such that we would only miss a negligible number of PS sources.

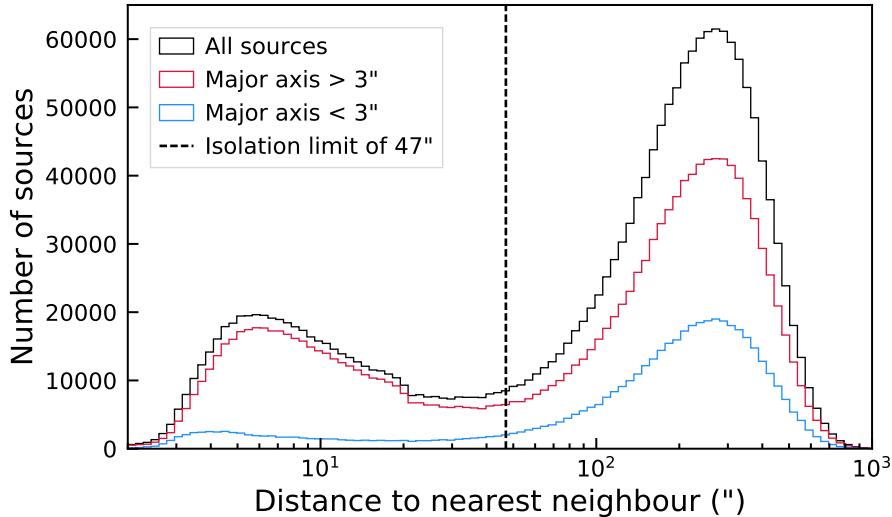
We thus find that even for nearby sources, the largest PS sources identifiable in our sample could not be resolved in VLASS. We are therefore free to remove any resolved source in VLASS, with a negligible chance of removing any source we could identify as PS.

### 3.3.2 Selecting the VLASS sample

The entire VLASS QL epoch 2 catalog contains 2,995,271 components and still contains duplicates due to the overlap of the QL images. As recommended by the user guide (Gordon, 2023), we selected only the sources without a duplicate, or sources with the best signal-to-noise ratio among multiple duplicate sources. We also select only sources with quality flags 0 or 4. These flags correspond to sources that have been fit without any quality flags, respectively to sources where the total flux density is less than the peak flux density. Sources with any other flags were removed from the sample. Examples of such

flags are sources that have a high probability of being a sidelobe, a signal-to-noise ratio less than five, or a bad peak-to-ring ratio. The peak-to-ring ratio is a measure that was designed to capture false-positive components that have been fit to artifacts (Gordon, 2023). After making these cuts to obtain reliable flux density measurements, we were left with a VLASS sample containing 2,446,020 components.

Slightly resolved sources in VLASS still fit reliably, but completely resolved sources are problematic. If PyBDSF fits such sources into multiple components, the crossmatching of these sources to other surveys is unreliable. As shown above, we are free to remove any resolved source from our sample, as the chance is negligible that such a source is a PS source.



**Figure 3.1:** The angular separation to the nearest neighbor of each source in the VLASS QL component catalog. The plot was made over the full component catalog as recommended by the user guide (Gordon, 2023). The distribution is plotted for the entire sample, as well as separately for sources with major axes larger and smaller than  $3''$ . All populations show a bimodal distribution, where the largest peak is dominated by unrelated radio components, and the lower peak is most likely caused by resolved radio sources (Gordon et al., 2023). The black dotted line shows the cutoff at  $47''$ .

In Figure 3.1, the distribution of the nearest neighbor distance in the VLASS component catalog is plotted, for all sources, sources with a major axis larger than  $3''$ , and a major axis smaller than  $3''$ . All populations show a bimodal distribution, for which the peak at large angular separations is dominated by random associations, and the peak at smaller angular separations most likely results from genuinely associated radio components (Gordon et al., 2023).

As illustrated by this bimodal distribution, selecting only the VLASS sources that are isolated within a radius of  $47''$  is a reliable measure of selecting unresolved sources. Any source in VLASS that has a neighbor within  $47''$ , that does have a match in the LoTSS+NVSS sample, is most likely resolved in VLASS. It is not likely to be multiple independent sources since we already defined our LoTSS+NVSS sample to be isolated within  $47''$ . Such a resolved source is then fit as multiple components in VLASS.

Sources with a neighbor within  $47''$ , where the total flux density of this source is at least ten times brighter than their neighbor, were included in the sample since these

neighbors are likely spuriously identified sources caused by sidelobe structures. Even if they are not sidelobes, but different components of a resolved source, the majority of the flux density is still captured in the bright source. By removing these clustered sources that are likely resolved, we removed 1,048,847 sources from the HF master sample, leaving us with 1,397,173 sources.

After removing the clustered sources, we also removed any source with a PyBDSF S-code 'C', which removed 184 sources from our sample, leaving us with only 'S' or 'M' S-code sources. Our final VLASS catalog contains 1,396,989 sources.

### 3.3.3 Crossmatching VLASS to LoTSS+NVSS

After selecting the VLASS sample as defined above, we crossmatched it to the LoTSS+NVSS sample with a crossmatching radius of  $6''$ . This crossmatching radius is larger than required by the positional accuracy and resolution of VLASS ( $2.5''$  resolution with  $0.5 - 1''$  positional accuracy; Gordon, 2023).

The reason we used this larger crossmatching radius is that sources in LoTSS that have S-code 'M' in VLASS might be positioned at slightly different positions in the sky when comparing the different surveys. VLASS is more sensitive to compact core components, while LoTSS is more sensitive to lobe emission. This is because the core of a radio galaxy generally has a flatter spectrum than its lobes, so at a higher frequency, the core will dominate over the lobes. The lack of short baselines of the VLA configuration used for VLASS also implies it is less sensitive to extended, diffuse structures than LoTSS. When a source has S-code 'M' in VLASS, the lobes and core get fit as multiple Gaussians, which gives the combined component a slightly different position in VLASS than in LoTSS. This shift in position can be larger than the astrometric inaccuracy. Hence, we used a larger crossmatching radius. Since we defined all sources in our VLASS catalog and in the LoTSS+NVSS sample to be isolated within  $47''$ , the probability of chance associations due to the larger crossmatching radius is negligible. The  $6''$  crossmatching radius corresponds to the LoTSS resolution.

There are 120,041 sources in this crossmatched sample, which we refer to as the HF master sample.

### 3.3.4 VLASS completeness

The LoTSS+NVSS sky coverage is completely covered by VLASS, which has a higher sensitivity than LoTSS or NVSS. Therefore, it was expected that almost all sources in the LoTSS+NVSS sample will have a detection in VLASS. However, we found that a fraction of sources in the LoTSS+NVSS sample do not have a match in VLASS. Below we explain why these sources are not detected in VLASS.

When crossmatching to the full VLASS catalog, 135,461 (86%) sources in LoTSS+NVSS had a crossmatch in VLASS. Isolating all VLASS sources within  $47''$  brought the number of sources in LoTSS+NVSS with a match in VLASS down to 120,041 (76%). There were thus 10% of sources that did have a match in VLASS when we do not select on isolation, that do not have a match when we do select on isolation. The decrease in matched sources was caused by sources being unresolved at  $6''$  in LoTSS, and resolved in VLASS at  $2.5''$ . If these sources are completely resolved in VLASS such that multiple components are fit, the distance between the nearest neighbor in the catalog is smaller than  $47''$  and these sources will be removed from the catalog. Therefore we can account for 10% of missing sources by them becoming multiple resolved sources in VLASS.

We found that 86% of sources in the LoTSS+NVSS sample had a match in the full VLASS catalog. The other 14% of sources in LoTSS+NVSS that do not have a VLASS detection is problematic. There are three possible explanations: 1) there is no VLASS coverage around the respective sources; 2) the VLASS QL image has a higher than normal RMS noise or the source has an extreme spectrum that makes it too faint to find a  $5\sigma$  detection in VLASS, or; 3) the source gets resolved out in VLASS.

### 1) Missing sky coverage

The VLASS sky area roughly covers the combined coverage of LoTSS+NVSS. However, there are some missing fields in the VLASS sky coverage. We found that 0.4% of sources have missing VLASS sky coverage. This thus only contributes a small percentage to the missing VLASS detections.

### 2) RMS noise analysis

It is possible sources in LoTSS+NVSS do not have a VLASS counterpart because some sources do not reach a signal-to-noise ratio of 5 in VLASS, either because they are too faint or because the RMS noise in specific fields is too high.

We obtained cutouts of the VLASS QL images in a radius of  $25''$  for all sources in the LoTSS+NVSS sample. To test the noise levels in VLASS, we compute the median RMS of each cutout that we obtained. We use the source finder Aegean and the Background and Noise Estimation tool (BANE; Hancock et al. 2012, Hancock et al. 2018) to compute the RMS maps for the VLASS cutouts around all sources in the LoTSS+NVSS sample. We calculated the median RMS noise for each cutout. The median RMS noise value for all the cutouts is shown in Figure 3.2.

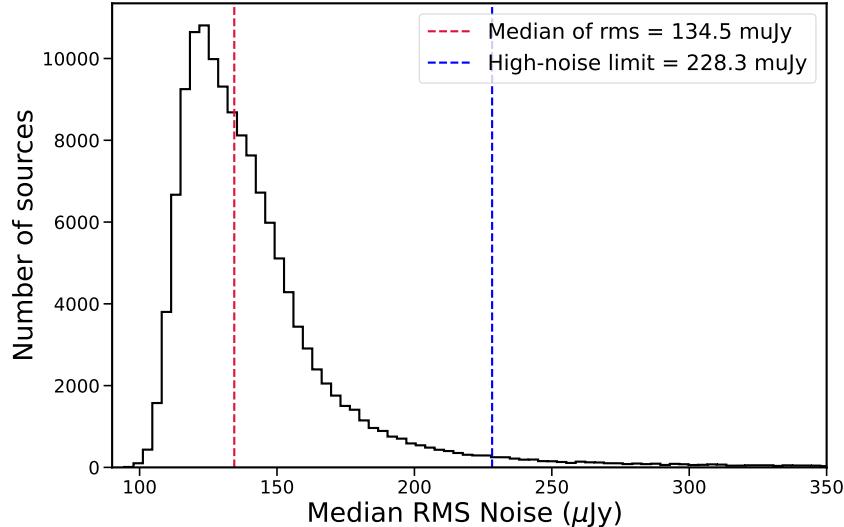
The sources with no VLASS counterpart are generally fainter in NVSS than sources with a VLASS counterpart. 80% of sources without a VLASS counterpart have a flux density in NVSS lower than 10mJy. For sources with a VLASS counterpart,  $\sim 68\%$  of sources have an NVSS flux density lower than 10mJy.

The faintest NVSS flux densities reported are approximately 2.1mJy, which is taken as a lower limit. Brighter sources in NVSS are obviously more likely to be detected in VLASS. We assume a typical source with a spectral index of  $-0.8$ , which is a typical value for the optically thick spectral index (Snellen et al., 1998). We found that the faintest sources in NVSS would have a typical flux density in VLASS of  $S_{VLASS} = S_{NVSS} \left(\frac{\nu_{VLASS}}{\nu_{NVSS}}\right)^{-0.8} = 1.14\text{mJy}$ . VLASS requires a source to have a signal-to-noise ratio of 5 for detection, so faint sources in NVSS will generally not be detected in VLASS if the median noise is higher than 0.228mJy. We refer to this noise level as the "high-noise limit". It is indicated in Figure 3.1. 3.6% of the VLASS QL cutouts have RMS noise higher than the high-noise limit, so we expect 3.6% of the sources in our LoTSS+NVSS sample to not have a detection in VLASS.

We were thus missing  $\sim 4\%$  of sources due to high noise fields (3.6 %) and missing fields (0.4 %), which leaves us to explain the remaining  $\sim 10\%$  of sources that should be observed in VLASS but are not. In the next paragraph, we investigate how the spatial sensitivity of VLASS might explain these missing detections.

### 3) Spatial sensitivity of VLASS

Another way a source could become undetected in VLASS is if it becomes resolved out. The flux density for unresolved sources is accumulated completely into a point convolved with the beam. If we then observe the same source with a higher resolution, we could



**Figure 3.2:** The median RMS noise in VLASS for all cutouts of  $25''$  by  $25''$  around the sources in LoTSS+NVSS. This figure still contains some duplicates since there is an overlap between the VLASS QL images. These duplicates are uniformly distributed over the sky so they have no effect on the outcomes of the RMS analysis. The RMS of each cutout was found using the BANE software. The median RMS of this distribution is indicated in red. The high-noise limit of 228.3 mJy, as described in the text, is indicated as well. Fields with an RMS higher than this have a large probability of missing the faintest NVSS sources if we assume a typical spectral index of  $-0.8$ .

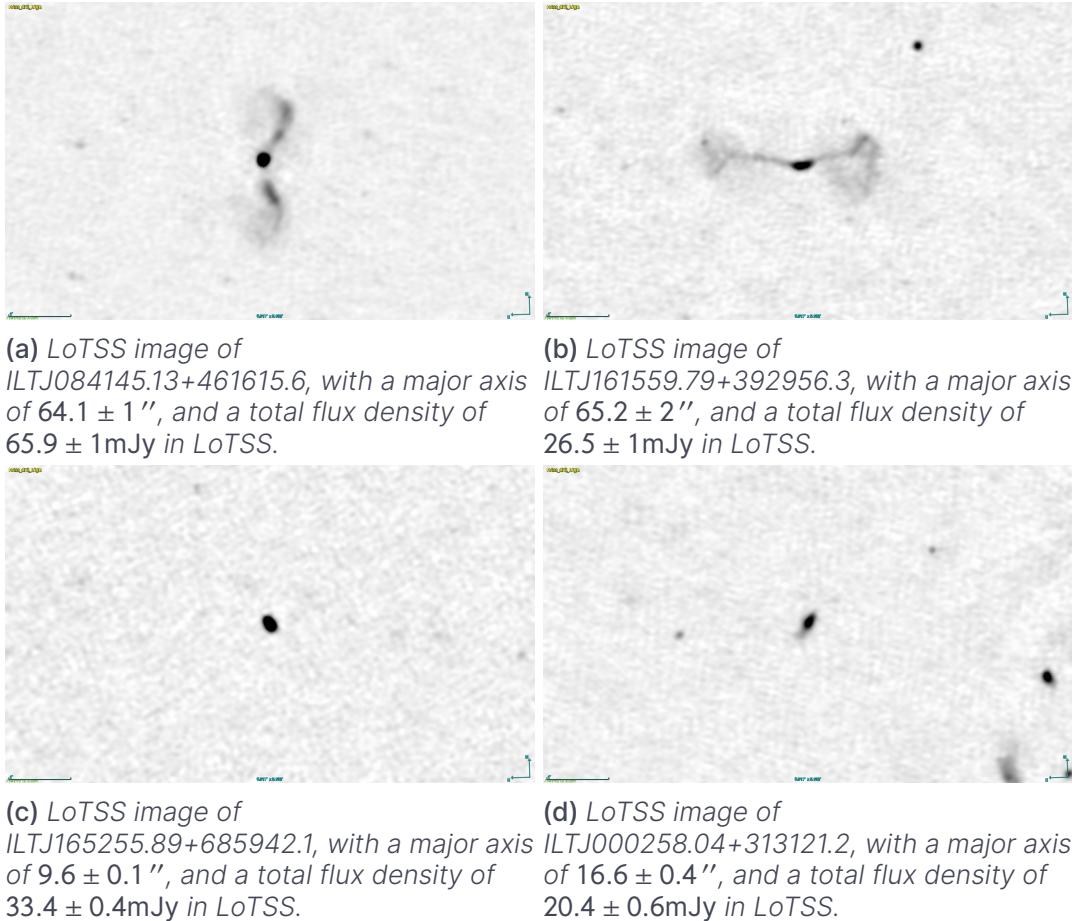
resolve the source structure. In this case, the flux density is then spread out over a larger solid angle, causing the surface brightness to fall below the detection limit, and the source to remain undetected.

We also need to consider that VLASS observes at a higher frequency than LoTSS and NVSS. Generally, the lobes of a double-lobed radio source are spectrally steeper than its cores. Therefore at higher frequencies, the core is much more dominant, and the lobes tend to be much fainter and more difficult to observe. VLASS has few short baselines compared to longer baselines. This makes it suited for capturing compact structures in the sky, but it is more difficult to capture extended emission, such as lobes, in this configuration. Therefore, the already fainter lobes have a high probability of being resolved out.

To verify if the missing VLASS detections are indeed caused by sources that are resolved out at  $2.5''$ , we carefully inspected the images of the sources without a VLASS detection in LoTSS at  $6''$ .

As can be seen in Figure 3.3, we found that these sources look like standard double-lobed radio sources. This is unexpected since we only included sources that we classify as unresolved in LoTSS. These sources must then have been incorrectly labeled as unresolved at  $6''$ .

In the LoTSS catalog, the 2 lobes get fit as separate components with S-code 'M' or 'S', and they are further than  $47''$  apart. They also happen to satisfy Equation 3.1. All these sources are point sources that are embedded in diffuse media. The source finder has failed to associate them and that is why they pass the unresolved criteria. Therefore, in our selection they get incorrectly classified as single unresolved sources in LoTSS+NVSS,



**Figure 3.3:** LoTSS images of some sources that are extended at 144 MHz, and fall below the detection limit at 3 GHz. These sources get miscategorized as unresolved in LoTSS. The top 2 images show sources with large major axis and seem to be simple double-lobes radio sources. The bottom 2 images are not selected on having a large major axis in LoTSS, but here it is also visible that they are slightly extended, causing them to be able to get resolved out in VLASS.

while they are actually resolved sources. Then at the VLASS resolution, they get resolved out and we end up without a match in VLASS.

We thus concluded that the last 10% of missing sources were likely caused by the misclassification of sources as unresolved in LoTSS, causing them to be resolved out in VLASS. Sources being missed in VLASS is actually preferential for this work since we have already shown that PS sources will be unresolved in VLASS. The missing sources in VLASS help to make sure our HF master sample contains only truly unresolved sources.

### 3.4 Crossmatching to additional radio surveys

We crossmatched TGSS, VLSSr, FIRST, the LoTSS and LoLSS inband spectra, and WENSS to the LF and HF master samples, using a crossmatching radius of  $6''$ . These surveys are not used for spectral index fitting but are used as visual guides in the SED plots.

# 4. Peaked Spectrum Sources

From the LF and HF master samples defined above, we identified PS sources by identifying the shape of the spectral energy distribution (SED). When the SED is concave, we can identify a source as a PS source. We will refer to the PS sources in the LF sample as the MPS sample, while the subsample of PS sources from the HF master sample will be referred to as our GPS sample.

## 4.1 Spectral classification and selection of peaked spectrum sources

The SED between two flux density measurements can be described by a power law. Using least squares fitting to the power law equation,

$$S = av^\alpha, \quad (4.1)$$

we found the spectral index  $\alpha$ , and the normalization parameter  $a$  between two flux density measurements. We used the spectral indices to select when a source is a PS source.

We fitted two variables to two flux density measurements, which will always give a perfect fit, implying the covariance matrix of the fit can not provide an uncertainty on the fit. To find the uncertainty  $\sigma$  in the spectral index, we calculated the spectral indices between  $S_i + \sigma_{S,i}$  and  $S_j - \sigma_{S,j}$ , as well as the spectral indices between  $S_i - \sigma_{S,i}$  and  $S_j + \sigma_{S,j}$ , where  $i$  and  $j$  correspond to different frequencies. This method gives a lower and an upper limit to the spectral index, and subtracting the actual spectral index from these values provides us with a lower and upper uncertainty on the spectral index.

In the LF master sample, we defined  $\alpha_{low}$  to be the spectral index between LoLSS and LoTSS, and  $\alpha_{high}$  to be the spectral index between LoTSS and NVSS. In the HF master sample, we defined  $\alpha_{low}$  to be the spectral index between LoTSS and NVSS, and  $\alpha_{high}$  to be the spectral index between NVSS and VLASS. From these spectral indices, we classified when sources are PS, and used this classification to define two sub-samples.

We defined a source to be PS if  $\alpha_{low} > \sigma_{\alpha\_low}$  and  $\alpha_{high} < 0$ . These criteria correspond to sources that have a peak in their spectrum, with at least a  $1\sigma$  certainty for  $\alpha_{low}$ . We are less strict for  $\alpha_{high}$ , since any source with a positive  $\alpha_{low}$  must have a turnover at higher frequencies. Then if  $\alpha_{high} < 0$ , we are sure that the turnover does not occur at higher frequencies than the upper survey frequency.

By using the uncertainty in the lower spectral index we were sure to remove the majority of flat spectrum sources that might contaminate our sample, while not excluding actual PS sources. Since the uncertainty in the spectral index is symmetric, any PS source

that we did not classify as PS will statistically counter any non-PS source contaminating our PS samples.

## 4.2 Megahertz peaked-spectrum sample

We present the color-color plot for the LF sample in Figure 4.1. This plot provides insight into the shape of the SEDs of sources in our sample. Each quadrant shows a mock SED, which is an approximate model of the shape of this SED. The lower right quadrant contains PS sources, plotted in red. In the LF master sample of 13,891 sources, we find 522 PS sources (3.8%), which we will from now on refer to as the MPS sample. This is the largest sample of MPS sources identified to date. These are sources that have a spectral peak approximately between 50 MHz and 1 GHz. The MPS sources are indicated in red in the color-color plot. We also plot the median of the uncertainties of the spectral indices, which indicate the median of the selection limit.

### 4.2.1 Discrepancies among LoLSS data releases

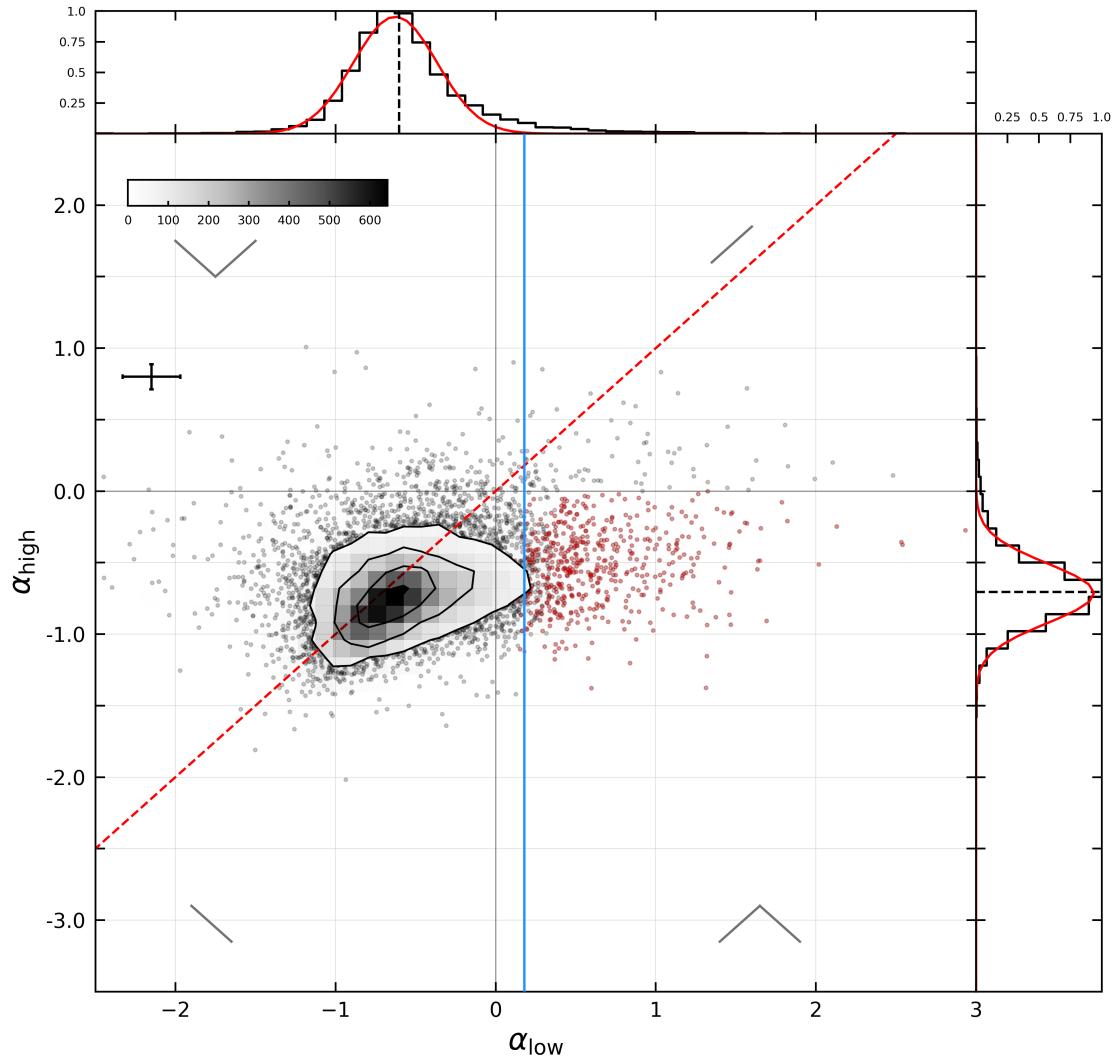
The method used for selecting the LF master sample is similar to that employed by Slob et al. (2022), whose sample is introduced in Section 1.5. They also crossmatch LoLSS, LoTSS, and NVSS to find a sample of MPS sources. However, the main difference between our work and Slob et al. is that they used the LoLSS preliminary data release (PDR), while in this work we used Data Release 1 (DR1). DR1 is more sensitive and has a higher resolution (15'') than the PDR (47'') since direction-dependent effects have been corrected. We do not expect that the change of resolution has a significant effect on the classification of PS sources since the LoTSS+NVSS sample used in both of our works only includes isolated sources in 47''.

In Slob et al., 373 sources were identified as PS sources, while we found 522 MPS sources, which already is an indication of the improved sensitivity between the PDR and DR1. Furthermore, Slob et al. use less strict criteria to define when a source is a PS source. They defined a source to be a PS source when  $\alpha_{low} \geq 0.1$  and  $\alpha_{high} \leq 0$ . When using this criteria, we found 767 MPS sources instead of 503, which is roughly twice as many sources as Slob et al.

We compared the sample by Slob et al. to our MPS sample. One would expect to find that most, if not all, of the 373 PS sources in the sample by Slob et al. are also present in the sample of 767 MPS we found by using their criteria. However, we found that there are 148 sources present in the PS sample from Slob et al., that are not present in our sample. 37 of these sources can be explained by being outside of the footprint of DR1, however that leaves 111 sources that need to be explained. Conversely, there are 75 sources that we classified as PS in our sample, that have not been classified as PS in the sample by Slob et al.

In order to explain this discrepancy, we visually inspected the SEDs of these divergent sources. We found that the total LoLSS flux density differs significantly between PDR and DR1 for many sources. This discrepancy was large enough to change the classification of a source.

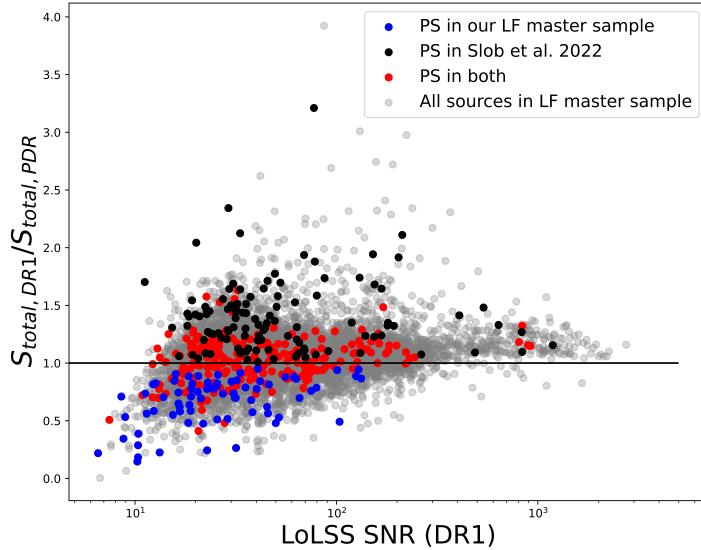
In order to study why the flux densities differ significantly between the PDR and DR1, we plot these flux densities against each other. In Figure 4.2 we plot the ratio between the total flux densities in the PDR and DR1, against the SNR in DR1. The SNR is defined here as the ratio between the total flux and the uncertainty in the total flux. This distribution



**Figure 4.1:** Color-color plot for the 13,891 sources in our MPS sample, where  $\alpha_{low}$  is the spectral index between LoLSS (54 MHz) and LoTSS(144 MHz), and  $\alpha_{high}$  is the spectral index between LoTSS and NVSS(1400 MHz). PS sources exist in the lower right quadrant and are indicated by the red color. The blue line represents the median of the error of  $\alpha_{low}$ . The contours represent 36, 125, 346, and 608 sources. It should be noted that there might be PS that are not indicated in red because they are hidden by the contours. The red dashed line represents the 1:1 ratio of the spectral indices. The median error bars are plotted instead of the individual errors, for readability. The normalized distributions of  $\alpha_{low}$  and  $\alpha_{high}$  are plotted, with a median and standard deviation of  $-0.6 \pm 0.4$  respectively  $-0.7 \pm 0.3$ . Mock SEDs are shown, to indicate the rough shape the SED of a source in each quadrant would have.

is centered at 1.04 and has a standard deviation of 0.23. Fluxes in DR1 thus tend to be generally brighter compared to the flux in the PDR, and a significant fraction of sources in the LF master sample has a large deviation in the LoLSS flux density between the PDR and DR1.

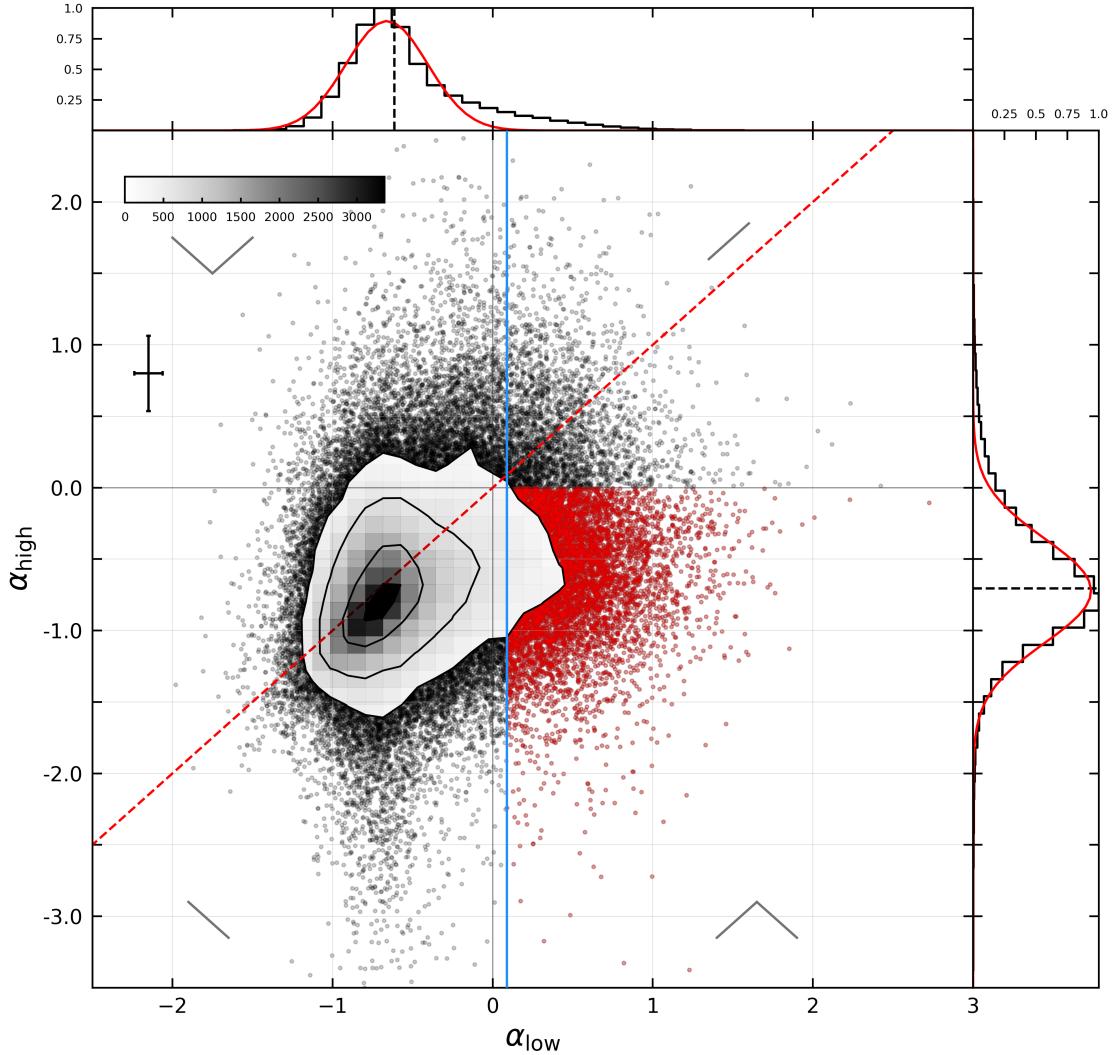
In private communication with F. De Gasperin, we conclude that there are systematic effects dominating over the noise in the PDR, which causes the wide wings in the flux density ratio. These wide wings cause sources to be mislabeled in the Slob et al. sample. The PDR should thus be used with care. We assume DR1 is more accurate and will continue working with this survey in the rest of the thesis.



**Figure 4.2:** The distribution of  $\frac{S_{total,DR1}}{S_{total,PDR}}$ , as a function of SNR in LoLSS. Ideally, we would see a tight distribution around 1, so the broad wings of this distribution are concerning. In different colors, we indicate the sources that are PS in our LF master sample, in the sample by Slob et al. (2022) and sources that are PS in both samples. The purple distribution has a median of 1.04, and has a standard deviation of 0.23.

### 4.3 Gigahertz peaked-spectrum sample

In Figure 4.3 we plot the color-color plot for the HF master sample. We used the same selection criteria as for the MPS sample to define the Gigahertz Peaked Spectrum (GPS) sample. In the HF master sample of 120,041 sources, we identified 8975 PS sources, which corresponds to 7.5 % of the sources in the HF sample. The GPS is the largest sample of GPS sources to date, over 6 times larger than the sample isolated by Callingham et al. (2017).



**Figure 4.3:** Color-color plot for the 120,041 sources in our GPS sample, where  $\alpha_{low}$  is the spectral index between LoTSS (144 MHz) and NVSS (1400 MHz), and  $\alpha_{high}$  is the spectral index between NVSS and VLASS (3 GHz). PS sources exist in the lower right quadrant and are indicated in red. The blue line represents the median of the error of  $\alpha_{low}$ , which corresponds to the median of the selection limit of PS sources. The contours represent 149, 456, 1474, and 2952 sources. It should be noted that there are also PS that are not indicated in red because they are hidden by the contours. The red dashed line represents the 1:1 ratio of the spectral indices. The median error bars are plotted instead of the individual errors, for readability. The normalized distributions of  $\alpha_{low}$  and  $\alpha_{high}$  are plotted, with a median and standard deviation of  $-0.61 \pm 0.4$  respectively and  $-0.71 \pm 0.5$ . Mock SEDs are shown, to indicate the rough shape the SED of a source in each quadrant would have.

## 4.4 Extremely steep-spectrum sources

In our LF sample, we find three PS sources with  $\alpha_{low} > 2.5$ , and in the HF sample, we find one such source. These sources are interesting because a spectral index of 2.5 is

the ultimate limit that SSA can achieve (O'Dea, 1998). Any spectral index above this value cannot be associated with SSA, implying the turnover in the spectra must be caused by FFA. This would indicate that these sources have dense environments, though it does not exclude the possibility that they are the young progenitors of FRI and FRII radio galaxies. The SEDs of these steep-spectrum sources can be found in Figure 4.4. We indicate the exact spectral indices in the caption. Since the shape of the SED is strongly dependent on the reliability of the low-frequency measurements, we visually inspected the images corresponding to these low-frequency measurements, which were deemed reliable.

To all flux measurements available we fit Equation 4.2, as taken from Snellen et al. (1998),

$$S(v) = \frac{S_{peak}}{(1 - e^{-1})} \cdot \left(\frac{v}{v_{peak}}\right)^{\alpha_{thick}} \cdot (1 - \exp(-v/v_{peak}))^{(\alpha_{thin} - \alpha_{thick})}. \quad (4.2)$$

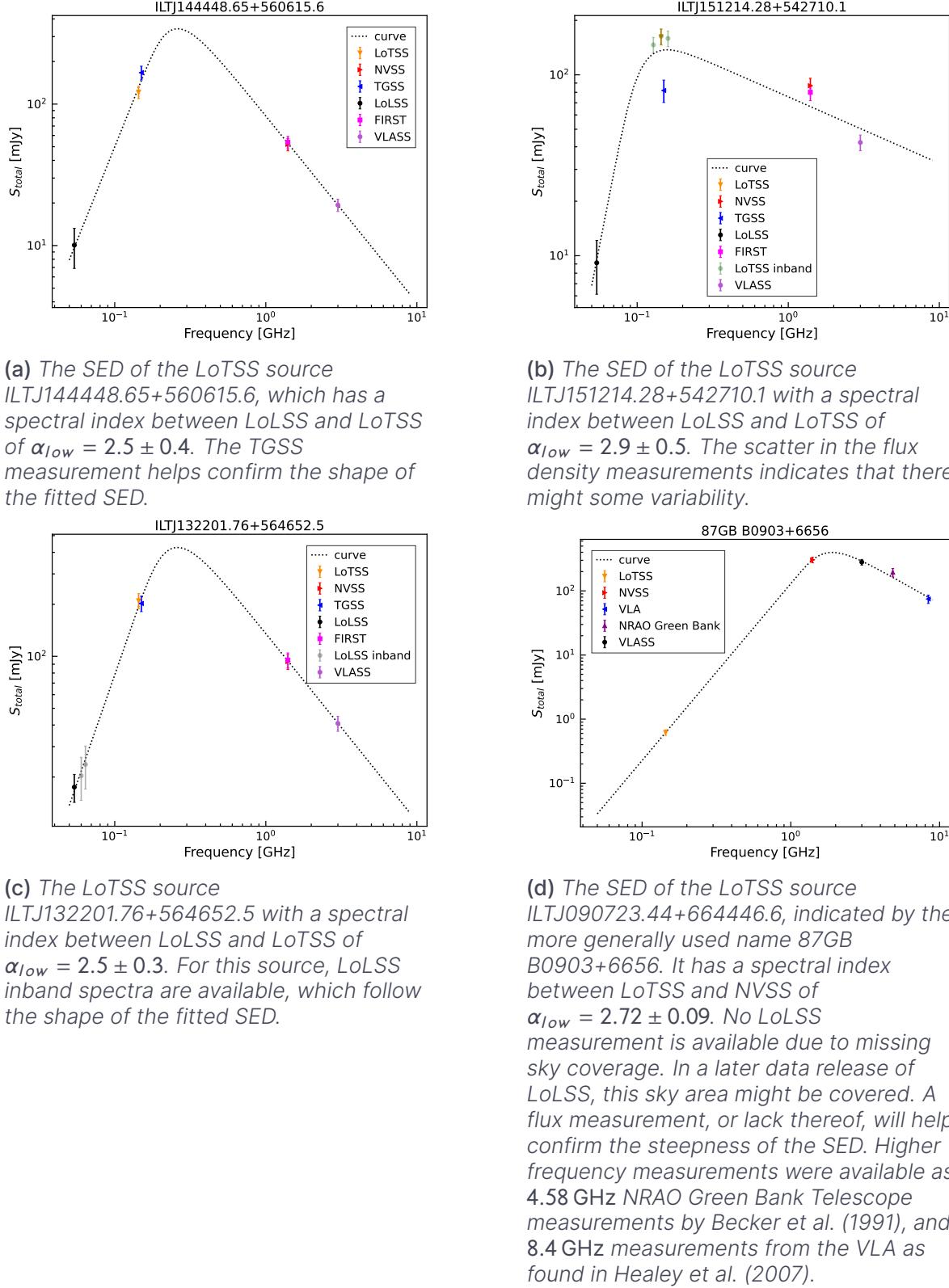
This generic curve is for illustrative purposes only.  $S(v)$  is the flux density as function of frequency  $v$ .  $\alpha_{thin}$  and  $\alpha_{thick}$  indicate the spectral indices in the optically thin respectively optically thick regions of the SED.  $v_{peak}$  indicates the peak frequency. Below, we will further discuss each of these sources.

For ILTJ144448.65+560615.6, we find that the TGSS measurement helps confirm the validity of the steep spectral index. The SED of this source can be found in Figure 4.4a. It has been identified as a quasar in the sixteenth data release of SDSS (Ahumada et al., 2020).

The SED of ILTJ151214.28+542710.1 is plotted in Figure 4.4b. This galaxy is part of the Abell 2029 cluster (e.g. Sohn et al. 2017, Sohn et al. 2019). The discrepancy between TGSS and LoTSS measurements indicates this might be a variable source. This variability could indicate that the source is a blazar, which is a type of AGN where the relativistic jets are pointed at us (Urry & Padovani, 1995). These jets cause relativistic beaming. Blazars are some of the most luminous and variable sources in the sky. The extreme variability of a blazar could cause it to appear as if the spectral index is larger than 2.5. More low-frequency measurements would need to be taken to confirm the SED of this sample.

The SED of ILTJ132201.76+564652.5 is plotted in Figure 4.4c. This source has LoLSS inband measurements available, which confirms the shape of the fitted SED.

The SED of 87GB B0903+6656 is plotted in Figure 4.4d. Higher frequency flux density measurements from literature were available for this source. We plot a 4.58 GHz observation taken by the NRAO Green Bank telescope (Becker et al., 1991), and an 8.4 GHz observation taken by the VLA (Healey et al., 2007). It has been identified as a flat-spectrum radio source by the latter, however, the lower frequency observations from LoTSS indicate that it is more likely a GPS source. In 1999, an attempt was made to find more low-frequency measurements of this source, however, not enough sky coverage was available at the time to indeed identify this source as a GPS source (Marecki et al., 1999). The host of this source has been identified as a quasar by e.g. Souchay et al. (2009).



**Figure 4.4:** The SEDs of the four sources in our sample with extremely steep spectral indices. The curve from Equation 4.2 was fitted to all radio flux density measurements available for these sources.

## 4.5 PS candidates in color-color plots

In this study, we focus our analysis on sources located in the lower right quadrants of the color-color plots shown in Figure 4.1 and Figure 4.3. These sources must have a turnover occurring below approximately 3 GHz for our GPS sample and around 1400 MHz for our MPS sample. However, it is important to note that the other quadrants of the color-color plot also contain interesting source candidates worth exploring.

The lower left quadrant comprises sources with a negative spectral index across the entire frequency range, indicating general power-law radio sources, while the upper right quadrant includes sources with a positive spectral index across the entire frequency range. For the latter category of sources, we know that a spectral turnover must occur at some frequency higher than approximately 3 GHz for our GPS sample and around 1400 MHz for our MPS sample.

There are 98 sources in the upper right quadrant of Figure 4.1 of the LF sample, which could potentially be either GPS or HFP sources. Similarly, the upper right quadrant of Figure 4.3 contains 3150 sources in the HF sample, which are most likely HFP sources. Out of the 98 sources in the upper right quadrant of the LF sample, 78 are indeed classified as GPS sources, leaving 20 sources that are likely to be HFP sources. As for the 3150 sources in the upper right quadrant of the HF sample, we only know that their spectral turnover occurs above 3 GHz, making them interesting candidates for further research.

In the upper left quadrants of the color-color plots, the sources exhibit a convex spectrum, which might be caused by multiple epochs of activity (Callingham et al., 2017). These sources are likely composite sources, with a steep-spectrum power-law component at low frequencies and an inverted component at high frequencies. However, further flux density measurements would be required to thoroughly examine these spectral energy distributions, making them a promising subject for future research.

In summary, while our primary focus is on the lower right quadrants of the color-color plots, other quadrants also hold intriguing source candidates. The sources in the upper right quadrant of the LF sample, as well as the upper left quadrant of the color-color plots, warrant further investigation.

# 5. Accretion mode classification

## 5.1 Accretion modes

We study the physical circumstances of the PS sources in our MPS and GPS samples by identifying their accretion mode. Supermassive black holes are found in essentially all massive galaxies (e.g. Magorrian et al., 1998). AGN are associated with massive amounts of material accreting onto these supermassive black holes. Two widely recognized modes exist for the accretion of material onto a black hole, each of which has a different feedback effect on the host galaxy. AGN are often classed relative to these accretion modes as low-excitation radio galaxies (LERG) or high-excitation radio galaxies (HERG) (Best & Heckman, 2012).

HERGs accrete in a radiatively efficient way, where the material is accreted to the central black hole through an optically thick, geometrically thin accretion disk. These AGN radiate across the entire electromagnetic spectrum, and are often associated with a highly evolving population, across all radio luminosities (Best & Heckman 2012, Pracy et al. 2016). HERGs are also referred to as 'quasar-mode', 'strong-lined' or 'radiative mode'.

LERGs accrete material through optically thin, geometrically thick accretion flows, also referred to as advection-dominated accretion flows (ADAFs; Narayan & Yi, 1995). LERGs produce little radiated energy, but the accretion mode of these AGN can lead to powerful radio jets. LERGs show little cosmic evolution (Best & Heckman 2012, Pracy et al. 2016). LERGs are also referred to as 'radio-mode', 'weak-lined', or 'radiatively inefficient'.

In a sample of young radio galaxies, it is expected to find relatively more HERGs, since a young population will evolve strongly over time. Therefore if PS sources are generally young radio galaxies, we would expect to find a surplus of HERGs in our sample compared to samples of general radio sources. Slob et al. (2022) find that the majority of their PS sample consists of HERGs, which they suggested indicates that they are a quickly evolving population.

## 5.2 Accretion mode classification

The most common method of classifying sources as either HERGs or LERGs is through optical emission line measurements. We used multiple criteria to classify our sources, following the method outlined by Best & Heckman (2012). A schematic view of the methods used can be found in Table 5.1, in which we also present the number of sources we were able to classify with each selection step.

There are multiple approaches of classifying AGN as either HERGs or LERGs, the majority of which rely on finding bimodal parameters. In literature, line flux ratios are

	LF			MPS		
	LERG	HERG	Unclassified	LERG	HERG	Unclassified
1. Best & Heckman	193	13	192	10	2	7
2. DR12 EI	105	53	622	4	4	7
3. DR12 [OIII] EW	N.A.	49	574	N.A.	3	4
4. $L_{[OIII]}$ criterion	255	N.A.	N.A.	2	N.A.	N.A.
Total	553	115	511	16	9	9
Total $z < 0.3$	269	16	89	10	2	6
Ratio of total $z < 0.3$	$0.72 \pm 0.05$	$0.04 \pm 0.01$	$0.24 \pm 0.02$	$0.6 \pm 0.2$	$0.11 \pm 0.06$	$0.3 \pm 0.2$
	HF			GPS		
	LERG	HERG	Unclassified	LERG	HERG	Unclassified
1. Best & Heckman	1,012	97	1,039	94	10	42
2. DR12 EI	1,000	410	4,376	92	29	189
3. DR12 [OIII] EW	N.A.	394	3,984	N.A.	19	170
4. $L_{[OIII]}$ criterion	1690	N.A.	N.A.	47	N.A.	N.A.
Total	3702	901	3316	233	58	163
Total $z < 0.3$	1789	194	696	145	21	31
Ratio of total $z < 0.3$	$0.67 \pm 0.02$	$0.072 \pm 0.005$	$0.260 \pm 0.008$	$0.73 \pm 0.07$	$0.11 \pm 0.02$	$0.16 \pm 0.02$

**Table 5.1:** The different classification methods of accretion modes, as well as the number of sources that were identified during each step. 'N.A.' is used to indicate that this step can only be used to classify a particular accretion mode. The ratios of each population as a fraction of the total population are also reported. We see that in the MPS and GPS samples, There is a higher HERG fraction than in the LF and HF samples. The row 'Total  $z < 0.3$ ' is used for the statistical test mentioned below.

often used to classify sources as HERGs or LERGs (e.g. Laing et al., 1994). A reliable method of classification is the Excitation Index (EI), as defined by Buttiglione & Capetti (2010);

$$EI = \log_{10}([OIII]/H\beta) - \frac{1}{3}[\log_{10}([NII]/H\alpha) + \log_{10}([SII]/H\alpha) + \log_{10}([OI]/H\alpha)]. \quad (5.1)$$

This parameter is bimodal, which allows us to divide the HERG and LERG populations at  $EI = 0.95$  (Buttiglione & Capetti, 2010). Another method that has been used as a classification for HERGs uses the equivalent width of the [OIII] emission line (e.g. Laing et al. 1994, Tadhunter et al. 1998).

### 5.2.1 Best & Heckman (2012) classification

We crossmatched our LF and HF master samples to the catalog produced by Best & Heckman (2012), using a crossmatching radius of  $2''$ . The Best sample consists of 18,286 radio-loud AGN, and was constructed by combining the seventh data release of the Sloan Digital Sky Survey (SDSS; Abazajian et al., 2009), NVSS, and FIRST.

In their catalog, Best & Heckman 2012 classify 481 sources as HERGs, 9,863 as LERGs and 7,942 sources could not definitively be identified. HERGs are generally brighter in the optical than LERGs and are therefore more easily classified than LERGs. This effect will be smaller at lower redshifts. Because HERGs are detectable out to further redshifts than LERGs, they perform any further analysis on a redshift cut of  $z < 0.3$ . We adopt this limit in further analysis.

When crossmatching the sources in our HF and LF samples to the Best catalog, we find for the LF sample we find 193 LERGs, 13 HERGs, and 192 sources remained unclassified. For the MPS sample, we find 10 LERGs, 2 HERGs, and 7 unclassified sources. In the HF sample, we find 1,012 LERGs, 97 HERGs, and 1,039 unclassified sources. For the GPS sample, we find 95 LERGs, 10 HERGs, and 42 sources remained unclassified.

### 5.2.2 Excitation index classification

We crossmatched all sources that were either unclassified in Best & Heckman (2012) or did not have a match in that catalog with a radius of 2'' to the Galaxy Properties catalog from the Portsmouth Group (hereafter called DR12; Thomas et al., 2013). This catalog provides optical emission lines for 492,450 galaxies from the 12th data release of SDSS. It makes use of the first release of Baryon Oscillation Spectroscopic Survey (BOSS) spectra. For all galaxies in this catalog, it contains emission line characteristics of 28 lines. We crossmatch our HF and LF master samples to this catalog and calculate for the sources that have not been classified yet the EI, using Equation 5.1.

We classified any source with EI larger than 0.95 as HERG, and any source with an EI lower than this value as LERG. This calculation demands that all six emission lines were detected. Missing emission lines could be caused by the lines being indeed faint, which indicates the source is a LERG. However, it could also be that such a source is undetected due to low signal-to-noise spectra. Therefore, we did not classify sources with any missing emission lines in the DR12 catalog. In the LF sample, we find 105 LERGs, 53 HERGs, and 622 unclassified sources. For the MPS sample, we find 4 LERGs, 4 HERGs, and 7 sources remained unclassified. In the HF sample, we find 1,000 LERGs, 410 HERGs, and 4,376 unclassified sources. For the GPS sample, we find 92 LERGs, 29 HERGs, and 189 sources remained unclassified.

### 5.2.3 Equivalent width of the [OIII] emission line classification

Any source with some undetected emission lines in DR12, that does have an [OIII] detection, can be identified as HERG when the equivalent width of the [OIII] emission line is larger than 5Å. If the equivalent width was lower than this value, we did not classify these sources as LERG, since the emission line could also be undetected due to low signal-to-noise spectra. In the LF sample, we classify 49 HERGs, while 574 sources remain unclassified. In the MPS sample, we find 3 HERGs, while 4 sources remain unclassified. In the HF sample, we classify 394 HERGs, while 3,984 sources remain unclassified. In the MPS sample, we find 19 HERGs, while 170 sources remain unclassified.

To test the reliability of the classifications, we check if there are any sources with conflicting classifications between the EI and [OIII] classification mechanisms. In the MPS sample, there are 2 sources that get classified as LERG using the EI, that would be classified as HERG using the [OIII] equivalent width. In the LF sample, there are 22 such sources, in the GPS 9 of such sources and in the HF sample 182. This is a small percentage of the entire sample. We follow the classification scheme of Best & Heckman (2012), which states that the EI is more reliable than the [OIII] equivalent width classification. Therefore, these sources with conflicting classifications remain classified as LERGs.

#### 5.2.4 [OIII] emission line luminosity

In Figure 5.1 we plot the [OIII] luminosities against the 1400 MHz luminosities separately for all MPS and GPS sources that are either classified as LERGs, HERGs or that could not be classified. We also plot the HERGs and LERGs from the LF respectively HF samples. The solid black line is the lower limit to the distribution of HERGs from Best & Heckman (2012). They state that unclassified sources below this line cannot be HERGs, while unclassified sources above this line can either be HERGs or LERGs. We indeed find that none of the HERGs in our GPS or MPS samples are below this line. Unclassified sources from the LF and HF samples were not plotted due to overcrowding of the plot.

We use this to further identify the sources that remained unclassified using the methods above. In the MPS sample, 8 unclassified sources are above the line, while 2 are below the line and must thus be LERGs. This compares to the LF sample where 332 sources are above the line while 255 are below the line and must thus be LERGs. This corresponds to there being roughly twice as many unclassified sources above the line in the MPS sample, while in the LF sample there are more sources below the line. These sources below the line get classified as LERG.

In the GPS sample, 143 sources are above the line, while 47 sources are below the line. In the HF sample, 2,348 sources are above the line and 1690 are below the line. The sources below the line again get classified as LERGs. This corresponds to there being roughly twice as many sources above the line than below in the GPS sample, while in the HF sample the ratio is roughly constant.

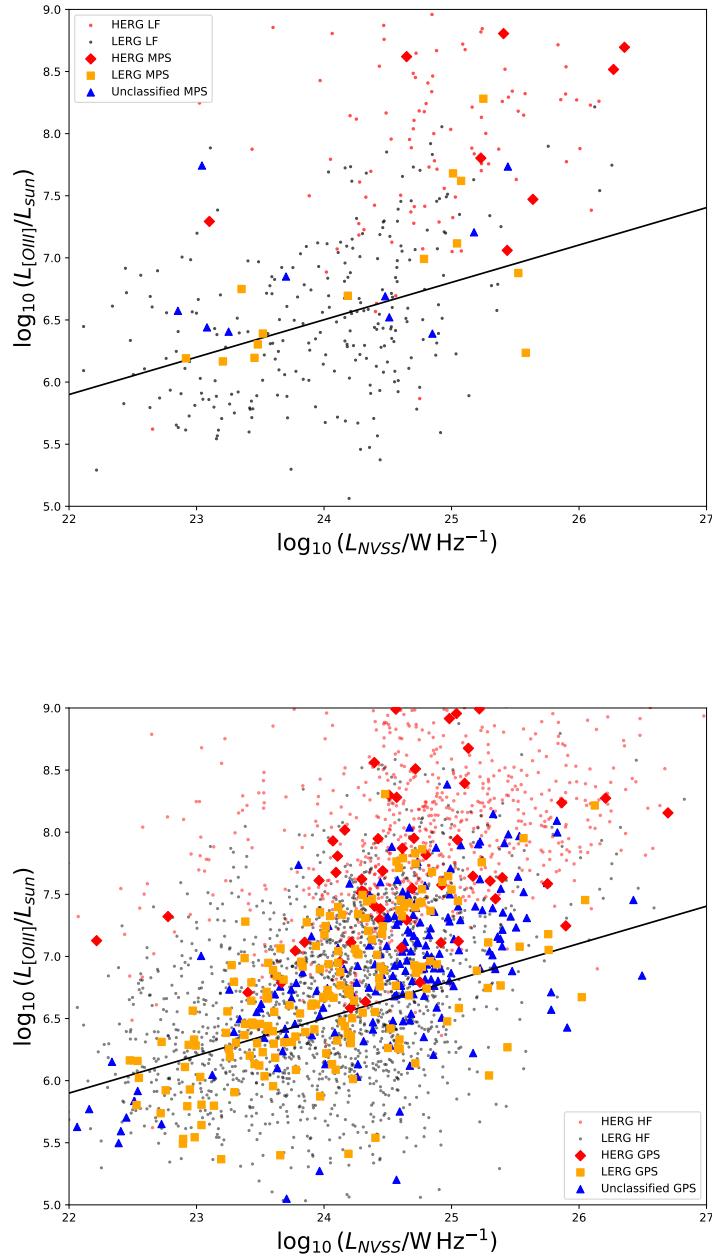
In our PS samples, there are relatively more sources above the line than in our LF and HF samples, from which we can also conclude that generally, PS sources have stronger [OIII] luminosities than AGN, and will thus likely contain a larger percentage of HERGs than the master sample.

## 5.3 The occurrence of HERGs and LERGs in our samples

For the LF, HF, MPS, and GPS samples we calculate the ratio of LERGs, HERGs, and unclassified sources in each of these samples, where  $z < 0.3$ . Uncertainties were propagated from Poisson statistics on the number of sources found. When less than 10 sources were found in a single category, we used the 84 % confidence limits as found by Poisson statistics (Gehrels, 1986). The ratios can be found in Table 5.1.

In the LF sample, the percentage of HERGs is  $\sim 4\%$ , while in the MPS this percentage is  $\sim 11\%$ . Similarly, we find that the percentage of HERGs in the GPS sample is  $\sim 11$ , compared to  $\sim 7\%$  in the HF sample. We thus find that HERGs are slightly more dominant in the MPS and GPS samples compared to the LF and HF samples, which would indicate that the PS sources in our sample are indeed more likely to be young radio galaxies. Below we will determine whether this difference is statistically significant.

In the GPS sample, there are fewer unclassified sources compared to the HF sample, which causes the fraction of LERGs in the sample to also increase, just as the fraction of HERGs. Relatively the HERG fraction between the HF and GPS samples increases with  $\sim 60\%$ , while the LERG fraction only increases with  $\sim 10\%$ . The increase in HERG fraction is thus much larger than the increase in LERG fraction.



**Figure 5.1:** The [OIII] luminosity against the 1400 MHz luminosities, for all sources in the LF and HF master samples that were classified as either HERG or LERG. Unclassified sources in the HF sources were not plotted due to overcrowding. Plotted were also all MPS respectively GPS sources that are HERGs, LERGs, or are not classifiable with our methods. The solid black line is the lower limit to the distribution of HERGs, as proposed by Best & Heckman (2012). Unclassified sources below this line cannot be HERGs, while unclassified sources above this line can either be HERGs or LERGs.

## 5.4 Statistics of the HERG/LERG populations

We performed the  $\chi^2$  test for independence between the distribution of HERGs and LERGs in our master samples and our PS samples. This test is designed to determine whether the two samples are drawn from the same distribution.

We defined our null hypothesis to be that the ratio of HERGs and LERGs in the LF sample and MPS sample, as well as the HF sample and GPS sample, are drawn from the same distribution. We show we can reject this null hypothesis, and thus that PS sources indeed are significantly more likely to be HERG than LERG. For simplicity, I will refer in the text below to the HF and GPS samples. The same analysis was performed between the LF and MPS samples.

We carefully need to deal with unclassified sources, as in a general sample of AGNs they are likely biased to being unidentified LERGs. For an explicit example, when we use the [OIII] equivalent width criteria, sources can either be classified as HERG or remain unclassified. Therefore we will also include the unclassified sources in our  $\chi^2$  test.

Furthermore, HERGs are more easily observed at higher redshifts than LERGs because they tend to be associated with more luminous sources, with stronger emission lines. LERGs tend to be associated with more diffuse and extended radio structures, which are more difficult to observe at higher redshifts. Therefore we limit the following calculation to sources with  $z < 0.3$ , as in Best & Heckman (2012). We have obtained these redshift measurements from the above-mentioned catalogs.

We calculate the test statistic  $\chi^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$ , where  $O_{ij}$  is the observed count, and  $E_{ij}$  is the expected count, found by calculating  $E_{ij} = \frac{R \cdot C}{G}$  for each combination of rows  $i$  and columns  $j$ . Here,  $R$  is the row total,  $C$  is the column total, and  $G$  is the grand total.

We found for the analysis of the GPS and HF samples that  $\chi^2 = 11.76$ . We compared this number to the critical value to determine whether we can reject the null hypothesis. There are 2 degrees of freedom in our sample. For a significance of 0.05, we found the critical value of  $\chi^2$  to be 5.99. Since  $11.76 > 5.99$ , we can reject the null hypothesis. We thus find that our GPS sample has a significantly higher percentage of HERGs compared to our HF sample.

We perform the same calculation for the LF and MPS samples and find a value of  $\chi^2 = 3.04$ . This value is not higher than the critical value, and therefore we cannot reject the null hypothesis, which means we cannot determine whether the MPS sample has significantly more HERGs than LERGs, compared to the LF sample. It should be noted however that this test should be used cautiously when dealing with low-number statistics. In the MPS sample we compute the test statistic for counts lower than five, which is required for the reliability of this test. A larger number of classifications in the LF sample will allow for a more reliable measurement. For the GPS and HF samples, the result is not skewed by low-number statistics.

# 6. Source counts

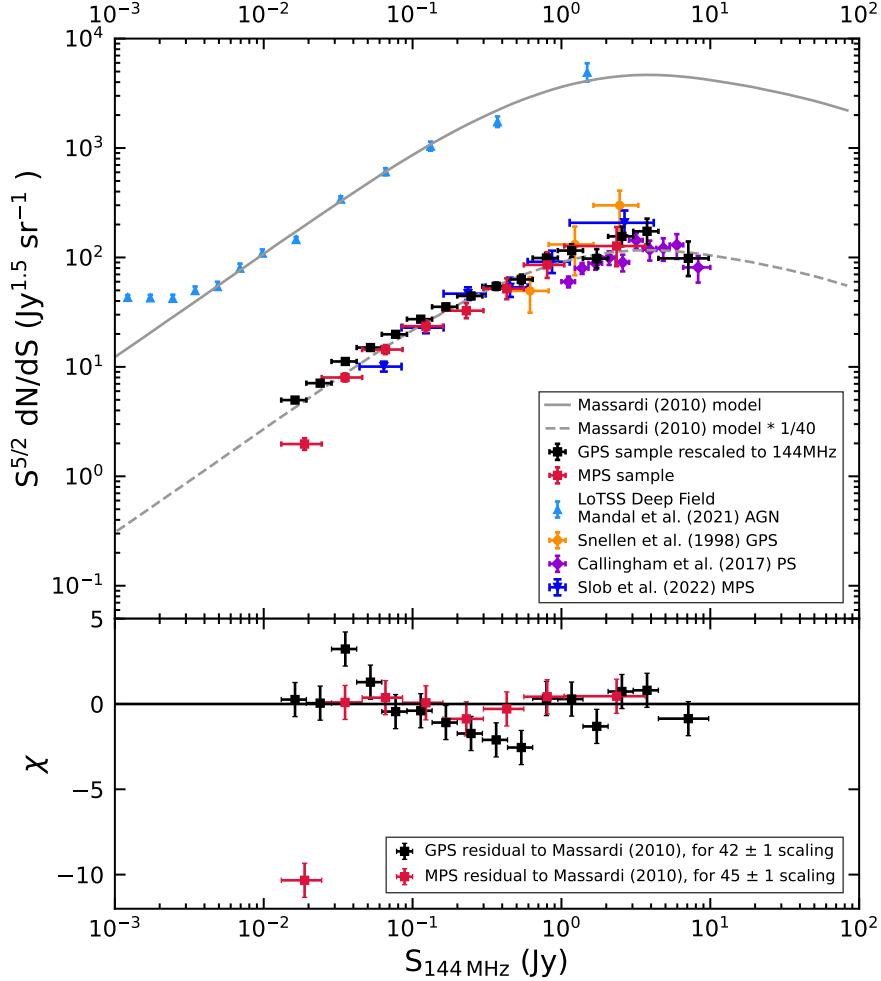
We further studied the evolution of the PS sources in our samples by computing the 144 MHz Euclidean normalized differential source counts for the MPS and GPS samples. We computed the source counts for the MPS and GPS samples near their respective peak frequencies of 144 MHz and 1400 MHz. To compare them, we then rescaled the GPS source counts to 144 MHz using a spectral index of  $-0.61$ , which corresponds to the median  $\alpha_{low}$  of our HF sample.

If PS are indeed young radio galaxies, we would expect an offset between the source counts of general radio galaxies and the PS sources, corresponding to the difference in lifetimes.

## 6.1 Source counts of the MPS sample

To compute the 144 MHz source counts for the MPS sample, we only included sources with LoTSS flux densities above  $S_{144\text{MHz}} > 13 \text{ mJy}$ , corresponding to the 95 % completeness limit of LoLSS of 11 mJy at 54 MHz (de Gasperin et al., 2021), rescaled to 144 MHz assuming a power law with a spectral index of 0.18. This spectral index corresponds to the median uncertainty in  $\alpha_{low}$  in the LF sample, which corresponds to the median of the selection limit of MPS sources.

There are 516 MPS sources with a LoTSS flux density brighter than 13 mJy. To find  $dN/dS$ , we divided these remaining sources over eight flux density bins, which we then divided by the detected area of the sky and the width of these flux density bins. The observed sky coverage for the MPS sample corresponds to the sky area of LoLSS ( $A=650$  square degrees; de Gasperin et al., 2023). We then normalized these source counts by  $S^{2.5}$ , where  $S$  is the central flux density of each bin, such that they are normalized to a uniformly distributed Euclidean space. The resulting normalized source counts can be found in Table 6.1 and in Figure 6.1. Note that significant incompleteness is likely impacting the lowest flux density bin.



**Figure 6.1:** The top figure shows the Euclidean normalized differential source counts, for the sources in our GPS and MPS sample. Plotted are also the source counts from PS samples as described in Snellen et al. (1998), Callingham et al. (2017), and Slob et al. (2022). To compare them to a general population of AGN, the 150 MHz source counts from the LoTSS Deep Field (Mandal et al., 2021) are also plotted. The model from Massardi et al. (2010) is indicated by the solid gray line, as well as this same model scaled down by a factor of 40 by the dashed gray line. In the bottom figure,  $\chi$  was plotted for the best fit for the Massardi et al. (2010) model to the MPS and GPS samples. The values for the scaling parameters are indicated as well.

## 6.2 Source counts of the GPS sample

The sources in our GPS sample have a turnover frequency near  $\sim 1400$  MHz. If we were to compute the 144 MHz source counts from the LoTSS flux densities, these source counts would be biased towards lower flux densities. Therefore, we first computed the source counts at 1400 MHz, which is near the turnover frequency for GPS sources. We then rescaled the 1400 MHz source counts to 144 MHz, such that they would be comparable to the MPS sample.

In the GPS sample, NVSS limits the flux density completeness and is 95% complete

at  $S_{1400\text{MHz}} = 3.24\text{ mJy}$  (Condon et al., 1998). We only included sources with an NVSS flux density above this completeness limit, although the lowest flux density bins might still be affected by incompleteness. There are 7,307 sources in the GPS sample above the completeness limit, which we used to make the source counts.

We then distributed the NVSS flux densities of these sources over sixteen bins, which gave us the value  $N_S$  as indicated in Table 6.1.  $N_S$  was then divided by the sky area of the sample and the bandwidth of each flux density bin to find  $dN/dS$ . For the GPS sample, LoTSS-DR2 is the survey with the limiting sky coverage (5634 square degrees; Shimwell et al., 2022). We then normalized our values for  $dN/dS$  by  $S^{2.5}$ , where  $S$  is the central flux density of each bin. This value gives us the Euclidean normalized source counts at 1400 MHz. They can be found in Table A.1 in the appendix.

To convert the source counts at 1400 MHz to 144 MHz, we rescaled the fluxes to 144 MHz assuming a power law model from Equation 4.1 with a spectral index of  $-0.61$ . This spectral index corresponds to the median  $\alpha_{low}$  of the HF master sample. The resulting 144 MHz Euclidean normalized source counts for the GPS sample can be found in Table 6.1 and in Figure 6.1.

### 6.3 Comparing to source counts of literature PS samples

We compare the source counts of our MPS and GPS samples to different samples from literature. The samples of PS sources have been introduced in Section 1.5. The source counts are shown in Figure 6.1.

Slob et al. (2022) employ a similar method we used to obtain our MPS sample, with the main differences being that we use a more sensitive and more reliable LoLSS catalog, and we are stricter in determining which sources are PS sources. We plot their 144 MHz source counts over six flux density bins as a reference and see there is a good agreement.

We also plot the source counts from the PS sample of Callingham et al. (2017). Only sources with flux density  $S_{143\text{MHz}} > 1\text{ Jy}$  were included, which corresponds to the reported 100 % flux density completeness of the GLEAM survey. Since the surveys we use to derive the GPS sample cover a larger area than those surveys used to form the MPS sample, we are more complete at the highest flux density end due to cosmic variance. As such, we observe that our GPS sample is consistent with that derived by Callingham et al. (2017).

We also plot the source counts for the sample of PS sources from Snellen et al. (1998). These source counts were evaluated at the individual peak frequencies for each PS source, corresponding to a median frequency of 2 GHz. We rescaled these source counts to 144 MHz assuming the power law from Equation 4.1 with a spectral index of  $-0.8$ .

In order to compare the populations of PS sources with a general population of radio sources, we also consider the source counts from the LoTSS Deep field (Mandal et al., 2021), which are the deepest 150 MHz source counts published to date. Among the observed deep fields are The Lockman Hole, the Boötes, and the Elais-N1 regions. The extremely sensitive source counts show the well-known flattening at  $\leq 1\text{ mJy}$ , associated with the emerging of the star-forming galaxy population(Mandal et al., 2021). These general source counts are well described by the Massardi et al. (2010) model. This model scaled down by approximately 1/40 corresponds to the offset we observe between the sample of AGNs and the sample of PS sources, as we will discuss below.

## 6.4 Lifetimes derived from source counts

In the upper plot from Figure 6.1 we see that all PS samples have a significant offset from the AGN sample, while the shape of the distribution remains similar. This offset, with no shift in its peak, indicates that PS sources are indeed the young progenitors of FR I and FR II sources, if the PS sources have shorter lifetimes. The PS source counts are roughly scaled down by a factor of 40 from the AGN source counts, which would then indicate that AGN live 40 times longer than PS sources.

For our MPS and GPS sources, we fit the model of Massardi et al. (2010) divided by a scaling parameter to the source counts of our MPS and GPS samples. We calculate the residuals between our source counts and the best fit for the scaled model, which is shown in the bottom plot of Figure 6.1. The lowest flux density bin was not taken into account for the fitting, since it is likely affected by incompleteness. We find that the MPS sample has a best-fitting scaling parameter of  $45 \pm 1$ , while the GPS sample has a scaling parameter of  $42 \pm 1$ . We could then conclude that MPS sources have a slightly shorter lifetime than GPS sources.

We see in the bottom plot of Figure 6.1 that at the lowest flux density bins, the GPS sample offset has a slight trend upwards, then an offset below the fitted model before approaching a roughly constant value. This trend in the GPS sample might be caused by blazars contaminating the GPS sample. Blazars are more likely to show a peak at higher frequencies and thus will be more likely to contaminate the GPS sample than the MPS sample (e.g. O'Dea et al. 1983, O'Dea et al. 1986, Kovalev et al. 2002, Kovalev 2005).

Snellen et al. (1998) report an offset of approximately 250 between typical large-scale radio sources and their GPS sample, for source counts at 2 GHz. This large offset would indicate that these PS sources have much shorter lifetimes than the MPS and GPS sources found in our samples, which were selected at 144 MHz respectively 1400 MHz. Slob et al. (2022) report that these shorter lifetimes could be caused by 2 GHz PS sources being smaller than the sources in our sample, due to their larger turnover frequencies. The detected radiation is thus much closer to the core, which causes the jet power at these places to be larger too. This power could cause a faster evolution. We do not observe significant evolution between our MPS and GPS samples, which we might expect according to this reasoning. We calculate the offset of the Snellen et al. (1998) GPS source counts rescaled to 144 MHz, to the Massardi et al. (2010) model. We find an offset of  $47 \pm 17$ , which is a much lower offset compared to what Snellen et al. (1998) report, and does not significantly differ from the offset of our MPS and GPS sample. We thus do not find any evidence that PS sources that peak at higher frequencies will evolve faster than PS sources that peak at lower frequencies.

We need to treat the offset cautiously since it does not take any redshift evolution into account. In Chapter 7 we compute luminosity functions for the sources in our sample, which is a more thorough method of dealing with evolution through redshift.

MPS sample			GPS sample		
$\langle S_{144\text{MHz}} \rangle [\text{Jy}]$	$N_s$	$S^{5/2} dN/ds_{-\sigma}^{+\sigma} [\text{Jy}^{1.5}/\text{sr}]$	$\langle S_{144\text{MHz}} \rangle [\text{Jy}]$	$N_s$	$S^{5/2} dN/ds_{-\sigma}^{+\sigma} [\text{Jy}^{1.5}/\text{sr}]$
0.019	75	$1.97^{+0.25}_{-0.23}$	0.016	1691	$4.97^{+0.12}_{-0.12}$
0.035	148	$8.0^{+0.71}_{-0.66}$	0.024	1326	$7.1^{+0.2}_{-0.19}$
0.066	121	$14.4^{+1.43}_{-1.31}$	0.035	1143	$11.21^{+0.34}_{-0.33}$
0.123	76	$23.53^{+3.02}_{-2.69}$	0.052	880	$15.02^{+0.52}_{-0.51}$
0.23	46	$32.63^{+5.55}_{-4.79}$	0.077	639	$19.84^{+0.82}_{-0.78}$
0.429	25	$51.87^{+12.58}_{-10.3}$	0.113	493	$27.29^{+1.29}_{-1.23}$
0.802	17	$85.5^{+26.15}_{-20.51}$	0.167	353	$35.4^{+1.99}_{-1.88}$
2.347	8	$127.14^{+62.6}_{-43.92}$	0.247	252	$44.46^{+2.98}_{-2.8}$
			0.364	176	$54.53^{+4.43}_{-4.11}$
			0.538	106	$62.96^{+6.73}_{-6.1}$
			0.793	100	$98.85^{+10.9}_{-9.87}$
			1.17	64	$115.57^{+16.32}_{-14.4}$
			1.727	29	$97.59^{+21.68}_{-18.01}$
			2.548	28	$155.79^{+35.33}_{-29.25}$
			3.759	17	$173.37^{+53.02}_{-41.59}$
			7.118	10	$98.08^{+41.79}_{-30.45}$

**Table 6.1:** The 144 MHz source counts for our GPS and MPS sample.  $\langle S \rangle$  is the central flux density of every flux bin in Jy,  $N_s$  is the number of sources in each bin, and  $N_{-\sigma}^{+\sigma}$  are the normalized differential source counts in  $[\text{Jy}^{1.5}/\text{sr}]$ . For the GPS sample, the source counts shown are the rescaled source counts from 1400 MHz to 144 MHz. The 1400 MHz source counts can be found in Table A.1.

# 7

# . Luminosity functions

## 7.1 Redshift information

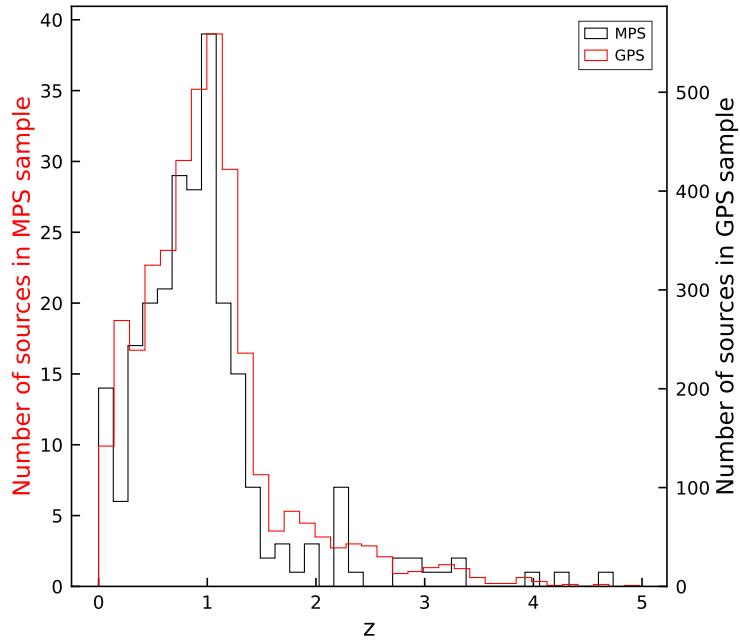
To further study the evolution of PS sources, we constructed the luminosity functions of our samples. We produced 144 MHz luminosity functions for the LF and MPS samples, and 1400 MHz luminosity functions for the HF and GPS samples. This allows us to study the MPS and GPS samples near their respective peak frequencies. We obtained redshift information from the LoTSS DR1 and DR2 optical catalogues (DR1: Williams et al. (2019), DR2: Hardcastle, 2023; in prep.). These works present value-added catalogs for radio source associations and optical/IR identifications from SDSS, Pan-STARRS (Chambers et al., 2016), and WISE (Wright et al., 2010). The DR1 value-added catalog is centered around the HETDEX region and contains optical information for 318,520 radio sources. The DR2 optical catalog does not overlap with DR1 and has optical identifications for 4,116,934 radio sources. The photometric redshifts for the sources in the value-added catalogs were obtained through the methods described in Duncan et al. (2019) for DR1 and Duncan et al. (2021) for DR2.

We combined these value-added catalogs and matched them to our sample. Spectroscopic redshifts were always preferred over photometric redshifts when available. The number of sources with redshifts available is shown in Table 7.1. We plot the distribution of the available redshifts in Figure 7.1, separately for the MPS and GPS samples. Both distributions peak near  $z = 1$ .

We need to account for both optical and radio selection effects when constructing our luminosity functions. We used *Astroquery*, to obtain  $g$  and  $i$  photometry measurements from SDSS DR16 (Ahumada et al., 2020), which is the same data release used in the DR2 LoTSS value-added catalog. This photometry will allow us to classify the optical selection effects. We crossmatched this photometry with a crossmatching radius of 2'' to our HF and LF samples.

Sample	Sources in sample	Redshift	Spectroscopic redshift
LF	13,891	6,938	2,197
MPS	503	244	67
HF	120,041	55,697	14,519
GPS	7,229	4100	1049

**Table 7.1:** The number of sources with reported redshifts in our master sample and PS samples. We also report how many of these redshifts are spectroscopic.



**Figure 7.1:** The distribution of redshifts for our MPS and GPS samples. Both distributions peak near  $z = 1$ .

## 7.2 The $1/V_{max}$ method

We used the standard  $1/V_{max}$  method (Schmidt, 1968) to construct the luminosity functions. The  $1/V_{max}$  method gives us the luminosity function of a bin centered at a luminosity  $L_j$  as

$$\frac{dN(L_j)}{d \log L} = \frac{1}{\Delta \log L_j} \sum_{i=1}^N \frac{1}{V_i}. \quad (7.1)$$

The summation in this equation is over all  $N$  sources  $i$  in luminosity bin  $j$ ,  $\Delta \log L_j$  is the width of the bin, and  $V_i$  is the total volume of space in which a source could be detected.  $V_i$  is calculated as  $V_i = V_{max} - V_{min}$ , where  $V_{max}$  is the volume of space that corresponds to the upper redshift limit  $z_{max}$  where the galaxy could still be detected, and depends on the true luminosity of the galaxy, the sky coverage of the survey, and the completeness limit of the sample.

To determine  $V_{max}$  we needed to take both optical and radio selection limits into account. The optical selection effects were considered to be identical for both the MPS and GPS samples. Although blazars might contaminate the GPS sample more than the MPS sample, we assumed that this contamination does not significantly affect the optical observations. The radio selection effects differ for each sample and are handled separately.

### 7.2.1 Luminosity functions of the MPS sample

To calculate  $V_{max}$  for the MPS sample, we first found the radio luminosity of the sources in our sample. The 144 MHz radio luminosity of a source was constructed by

$$L_{144\text{MHz}} = \frac{4\pi D_L^2 S_{144\text{MHz}}}{(1+z)^{1+\alpha_{high}}}, \quad (7.2)$$

where  $D_L$  is the luminosity distance,  $S_{144\text{MHz}}$  is the flux density at 144 MHz. The  $1/(1+z)^{1+\alpha_{high}}$  term is the k-correction, where  $\alpha_{high}$  is the upper spectral index.

In radio, the MPS and LF completeness limit is determined by the LoLSS completeness limit of 11 mJy at 44 MHz. For the PS sources in our MPS sample, we rescaled this limit to 144 MHz assuming a power law with a spectral index of 0.18, which corresponds to the mean of the error of  $\alpha_{low}$ . This completeness limit corresponds to the selection criteria of PS sources. At 144 MHz, the completeness limit for the MPS sample is then 13mJy.

For the other sources in the LF sample, we were not bound by the PS selection criteria, therefore we rescaled the LoLSS completeness limit with the lower 95 % limit of the distribution of  $\alpha_{low} = -1.05$ . For the LF sample, we find a completeness limit of 3.94mJy.

We found  $z_{max,radio}$  by computing the true radio luminosity of each source using Equation 7.2 and the measured redshift. We then considered a test source which we observed with a flux density exactly that of the flux density limit. We evaluated the luminosity such a source would have at a range of redshifts of  $\Delta z = 0.0001$  to a maximum redshift of  $z = 100$ . The redshift where the luminosity of this source equaled the true luminosity of our source is then  $z_{max,radio}$ .

We also need to consider the optical selection limit, which was more complicated because the redshifts of the sources in our sample were taken from different optical surveys. We therefore only considered sources of which both the SDSS  $g$  and  $i$  magnitudes have been detected. The 95 % completeness limit of SDSS is  $m_{i,lim} = 21.3\text{mag}$  was taken as the optical selection limit. The other optical survey from which we obtained redshifts is Pan-STARRS. Slob et al. (2022) have found that when calculating the optical selection effects with Pan-STARRS instead of SDSS, no significant differences in the luminosity functions were found. Therefore we neglected Pan-STARRS and only used SDSS to capture the optical incompleteness.

To find  $z_{max,opt}$ , we calculated the true absolute i-band magnitude  $M_i$  of each source using

$$M_i = m_i - DM - K_i(z). \quad (7.3)$$

In this equation,  $m_i$  is the apparent i-band magnitude of our sources, DM is the distance modulus, and  $K_i(z)$  is the  $k$ -correction. The  $k$ -correction was calculated using the *K-corrections calculator* (Chilingarian et al. 2010 , Chilingarian & Zolotukhin 2012) for the  $g - i$  color. We again considered a test source, that was observed with an apparent  $i$ -band magnitude at exactly that of the optical completeness limit. We evaluated this test source at the same redshift range as the radio calculation, and determine at which redshift the absolute magnitude of the test source equals the absolute magnitude of the true source. We consider this redshift as the maximum redshift our source could have to still be detectable in SDSS.

We then pick the lowest redshift out of  $z_{max,opt}$  and  $z_{max,radio}$ , and determine which volume of space corresponds with such a source. The fractional area of sky coverage for

our sample was also taken into account, which for the LF and MPS samples corresponds to  $650\text{deg}^2$ .

We then constructed the **144 MHz** luminosity functions for our LF and MPS samples. The **144 MHz** luminosity functions for the MPS sample are found in Figure 7.2 and Table B.1. We only included sources with  $L_{144\text{MHz}} > 1\text{e}23\text{W Hz}^{-1}$ , as below this limit, the luminosity function will be dominated by star-forming galaxies(SFGs; Sabater et al. 2019, Franzen et al. 2021). We are only interested in studying the AGN population. We plot the luminosity functions at redshift bins  $0.0 < z < 0.1$ ,  $0.1 < z < 0.5$ ,  $0.5 < z < 1.0$ ,  $1.0 < z < 1.5$  and  $1.5 < z < 3.0$ . We then calculated the luminosity functions for a range of luminosity bins in each redshift bin, using Equation 7.1. For  $V_{min}$ , we used the volume of space that corresponds to the lower redshift limit of each redshift bin.

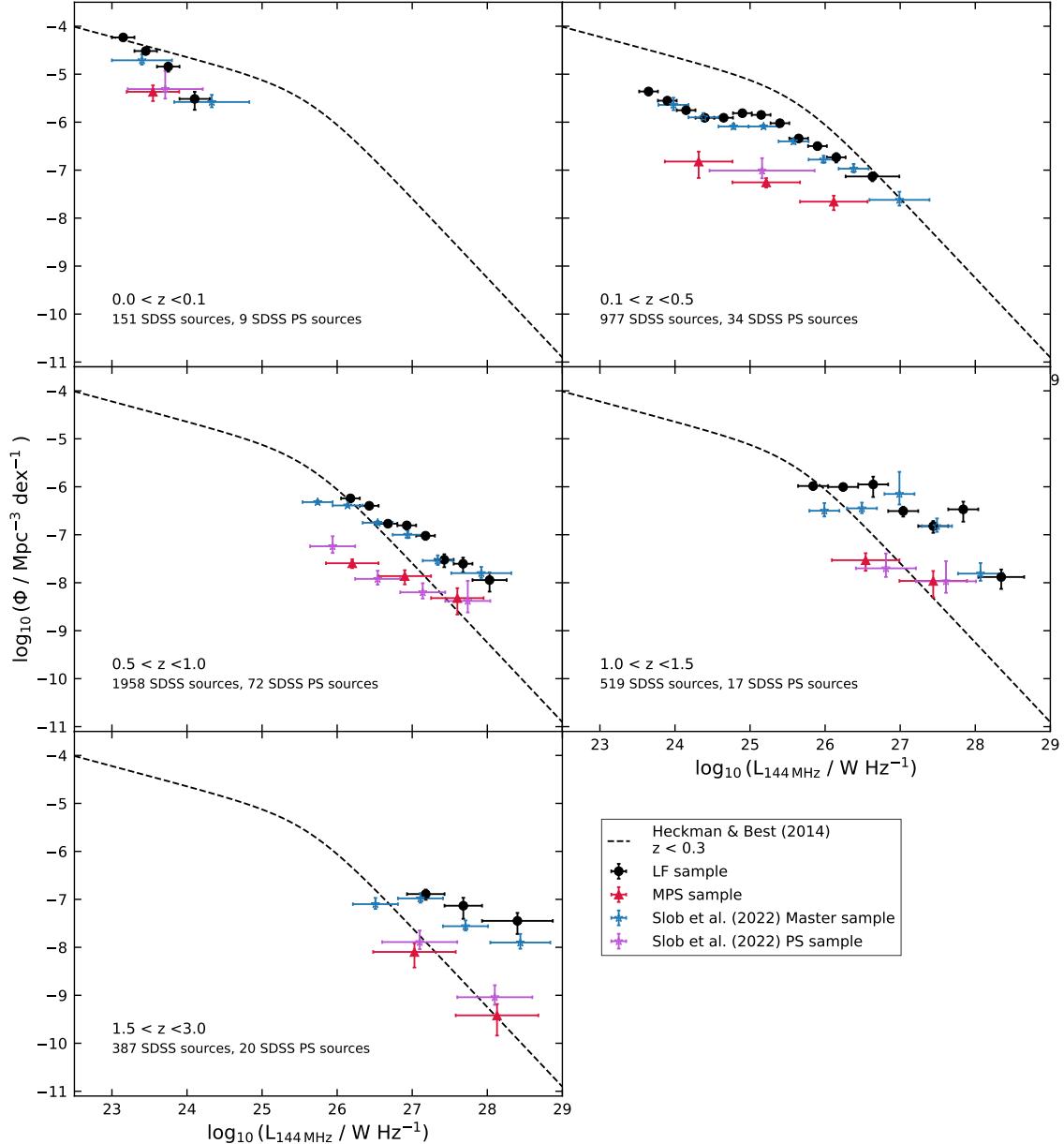
We calculated the Poissonian counting errors  $\sigma_j$  on  $dN(L_j)/d \log L$  for each luminosity bin  $j$  as

$$\sigma_j^2 = \frac{1}{(\Delta \log L_j)^2} \sum_{i=1}^N \frac{1}{V_i^2}. \quad (7.4)$$

We chose the luminosity bins such that each bin has at least 8 sources, though this was not always possible. For luminosity bins with 5 or fewer sources, we took the lower and upper **84%** confidence limits as found by Poisson statistics (Gehrels, 1986).

We plot the **1400 MHz** double power-law luminosity function model from Heckman & Best (2014), rescaled to **144 MHz** with a spectral index of  $-0.7$ . This model was derived for local ( $z \approx 0.1$ ) AGN, however we also plot it at higher redshift bins to guide the eye. We would expect a complete sample of AGN to coincide with this model at low redshift, which has been shown by Sabater et al. (2019), for a sample of AGN in LoTSS-DR1. We do not observe a complete overlap between our LF sample and the model, instead our LF sample is shifted downward in the two lowest redshift bins. This shift downward is likely caused by the selection criteria of our LF sample. We only select unresolved isolated sources in LoTSS. This selection criterion decreases the number of sources in our sample, causing an offset downwards compared to other samples of AGN. This effect is smaller at higher redshifts since most AGN will be unresolved there.

We also plot the **144 MHz** luminosity functions from Slob et al. (2022), for their master sample and PS sample. Due to the similar selection methods between Slob et al. (2022) and our work, we do not expect to find large differences between the two samples. We find our luminosity functions coincide reasonably well with their work. They indicate that their sample is affected by optical incompleteness for  $z > 1.0$ , which causes the shape of the luminosity function to not reflect that of a complete sample. However, these selection effects will be similar for the master and PS samples. Therefore they assume the offset between their PS and master samples reflects the true evolution between the samples. We applied a similar method of optical selection as Slob et al. (2022), so we expect that the observed offset between the LF and MPS samples is also reliable. We further investigate the offset to these luminosity functions in Section 7.3.



**Figure 7.2:** The 144 MHz luminosity functions for our LF and MPS samples are indicated in black circles respectively red triangles. The 144 MHz luminosity functions for the master sample and PS sample by Slob et al. (2022) were plotted in blue and purple. The Heckman & Best (2014) double power law model, which was derived at 1400 MHz, was rescaled to 144 MHz using a power law with a spectral index of  $-0.7$ . This model is valid for local AGN, and should not be physically compared to the luminosity functions at higher redshifts. In the lowest redshifts bin, the shape of the luminosity functions of our LF sample does not coincide with the Heckman & Best (2014) model, though it does coincide with that of the master sample by Slob et al. (2022). This is because both the Slob sample and ours are affected by selection effects. At higher redshift bins, optical incompleteness will affect our samples, which affects the LF and MPS samples in the same way. The offset between these samples remains reliable.

### 7.2.2 Luminosity functions of the GPS sample

We computed the 1400 MHz luminosity functions for our HF and GPS samples in a very similar way to the 144 MHz luminosity functions. We calculated  $V_{max,opt}$  in the same way, where we assume that the optical properties are the same between the MPS and GPS samples.

For  $V_{max,radio}$ , we also computed the radio luminosities of each source in our samples using Equation 7.2. The LoTSS flux density at 144 MHz was replaced with the NVSS flux density at 1400 MHz such that we are evaluating the luminosity functions near the peak frequency of the PS sources. NVSS is also the limiting flux density of these samples, with a 95% flux density limit of 3.24 mJy. We used this limit to determine  $V_{max,radio}$  for both the GPS and HF samples.

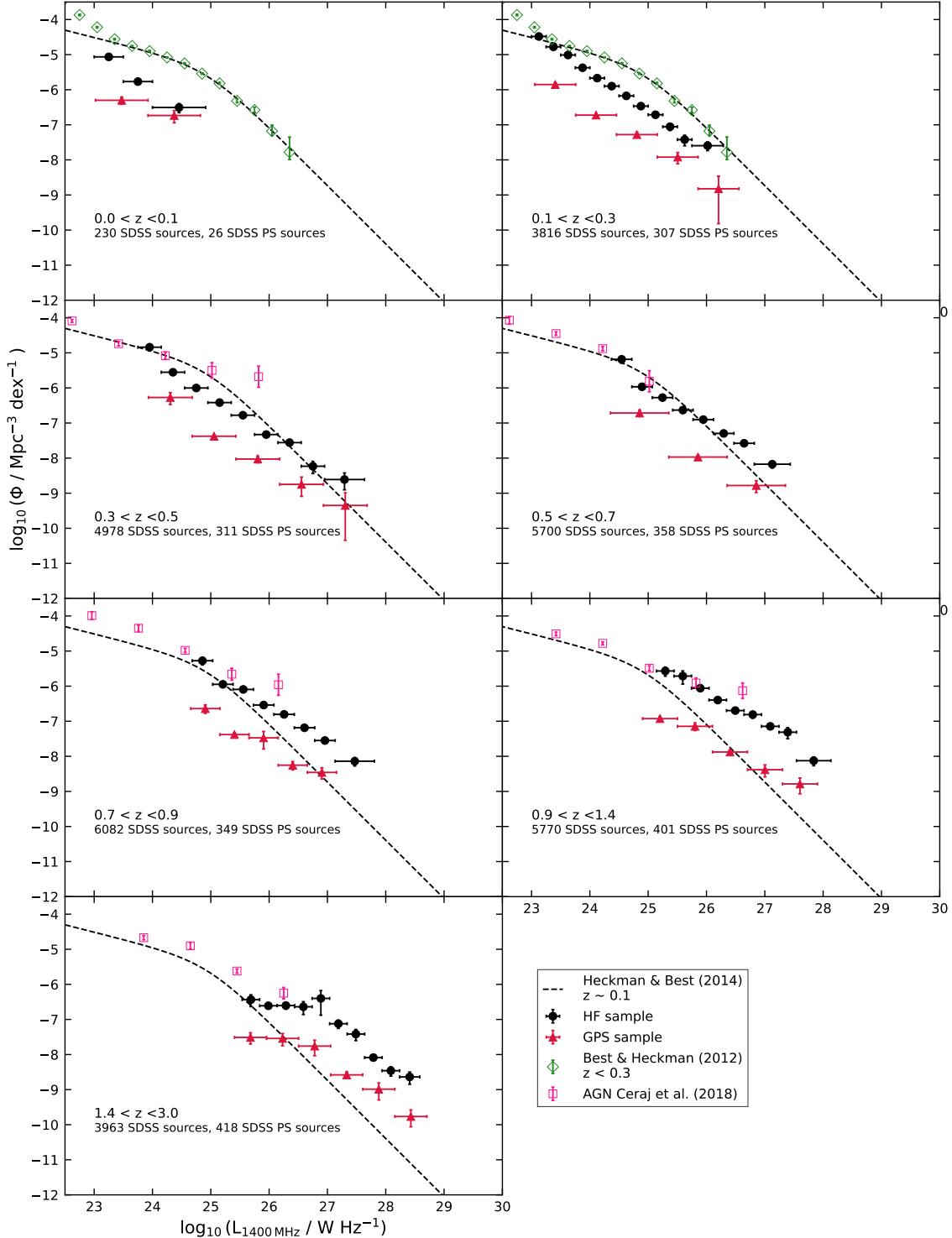
We again selected the lowest  $z_{max}$  out of  $z_{max,radio}$  and  $z_{max,opt}$  for each source, which we used to calculate  $V_{max}$ , where the fractional area of the sky of the HF sample was taken into account. The limiting sky area is of  $5634\text{deg}^2$  (LoTSS; Shimwell et al., 2022).

For the 1400 MHz HF and GPS luminosity functions, more sources were available than for the 144 MHz LF and MPS luminosity functions. Therefore we opted to use smaller redshift bins. The limits of these redshift bins were chosen as  $0.0 < z < 0.1$ ,  $0.1 < z < 0.3$ ,  $0.3 < z < 0.5$ ,  $0.5 < z < 0.7$ ,  $0.7 < z < 0.9$ ,  $0.9 < z < 1.4$ , and  $1.4 < z < 3.0$ . The lower redshift bin limit corresponds to  $V_{min}$ , which gives us  $V_i$  for each source.

We plot the 1400 MHz luminosity functions in Figure 7.3 and in Table B.2. We again did not include any source with  $L_{1400\text{MHz}} < 1e23\text{mJy}$ , due to the contamination by SFGs (Mauch & Sadler, 2007). Incompleteness could affect some lower luminosity bins, which were removed from the plot due to their unreliability.

The double power-law 1400 MHz luminosity function model from Heckman & Best (2014) was also plotted, to guide the reader's eye. It should be noted that this fit is only valid at  $z < 0.3$ , and should not be interpreted as a fit to our sample at higher bins. The  $0 < z < 0.3$  luminosity functions of Best & Heckman (2012) were plotted for their entire sample of radio sources. We also plot the Ceraj et al. (2018) 1400 MHz AGN luminosity functions. Their luminosity functions were reported in different redshift bins than ours. We have therefore plotted their  $0.1 < z < 0.5$ ,  $0.5 < z < 0.7$ ,  $0.7 < z < 0.9$ ,  $0.9 < z < 1.1$  and  $1.7 < z < 2.1$  bins, which provides the best overlap with the redshift bins chosen for our luminosity functions.

We find that the luminosity functions for both our HF sample and GPS sample do not follow the same shape as the Heckman & Best (2014) model in the lowest redshift bins, similar to our 144 MHz luminosity functions. This again is most likely caused by our selection of only unresolved sources. At higher redshifts, the shape between the Ceraj et al. (2018) luminosity functions and ours coincide relatively well.



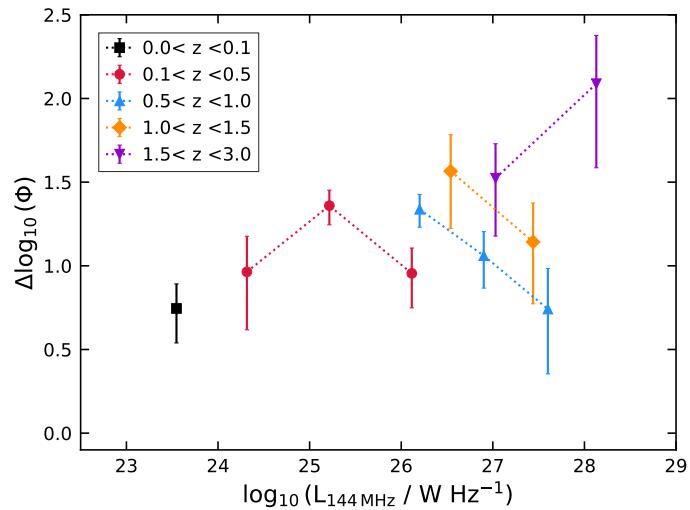
**Figure 7.3:** The 1400 MHz luminosity functions for our HF and GPS samples, at a range of redshifts are plotted in black circles and red triangles respectively. The curve indicates the Heckman & Best (2014) double power-law model, which has been plotted to guide the eye. It should be noted that this model is only valid near  $z \approx 0.1$ . We also plot the luminosity functions for a local sample of radio sources by Best & Heckman (2012) in green diamonds, and a sample of AGN by Ceraj et al. (2018) as pink squares. At low redshifts, the selection of only unresolved sources causes a difference between the luminosity functions of our sample and those by Best & Heckman (2012). At higher redshift, the shape is different between the Ceraj et al. (2018) luminosity functions and our HF sample because of optical selection effects. The offsets between our HF and GPS sample are affected by these effects similarly, and will therefore be considered reliable.

### 7.3 Luminosity function offsets

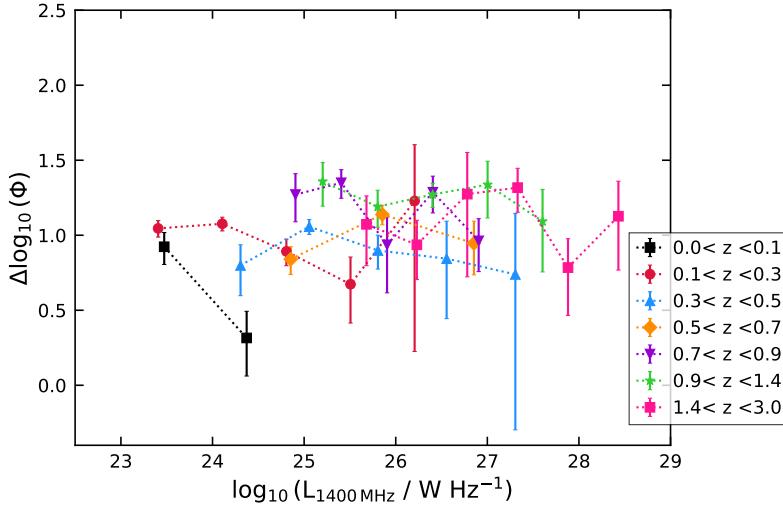
To further study the offset between the PS samples and master samples, we calculated the offsets between the LF and MPS samples, and between the HF and GPS sample, for each redshift bin. We calculated these by interpolating the LF and HF luminosity bins at the MPS respectively GPS luminosity bins. We plot these offsets in Figure 7.4 for the LF and MPS samples and in Figure 7.5 for the HF and GPS samples.

We find that there seems to be a trend slightly upward towards higher redshift bins for the offsets between the MPS and LF samples. This trend would indicate there is a larger percentage of MPS sources at lower redshift. It is generally assumed that PS sources are under-dense locally (e.g. Kunert-Bajraszewska et al. 2010, Labiano et al. 2006, Snellen et al. 1998). A smaller offset at lower redshift bins contradicts this statement. This apparent over-density could be caused by the selection effects of our master sample. PS sources are compact sources, and therefore approximately always unresolved, even at low redshifts. A general population of AGN at low redshift is more likely resolved at low redshift. Therefore, since we only select unresolved sources in our HF sample, we are more likely to be biased towards PS sources locally.

The offset between the GPS and HF samples in Figure 7.5 seems to be more constant for the different redshift bins. In Section 7.4 we further discuss the implications of the offset evolution.



**Figure 7.4:** The offset between the LF and MPS samples luminosity functions for different redshift ranges. A slight trend upward for higher redshift bins can be observed.



**Figure 7.5:** The offset between the HF and GPS samples luminosity functions for different redshift ranges. No significant redshift evolution can be observed.

## 7.4 Redshift evolution of the luminosity function offsets

We investigate the luminosity function offset evolution for the MPS and GPS samples. We produce the luminosity functions of the GPS and HF samples in the same redshift bins as chosen for the MPS sample. For every redshift bin in both samples, we calculate the weighted mean of the offsets. We assume that there is no evolution within the luminosity bins. Rescaling the 1400 MHz luminosity functions to 144 MHz was not necessary since this cancels out when taking the mean offsets. The average offset for each redshift bin for both samples is indicated in Table 7.2, and was plotted in Figure 7.6.

We study the evolution of these mean offsets as a function of redshift, by fitting both a horizontal and a sloped line to the mean offsets. If we find that the horizontal line provides a better fit, it implies there is no redshift evolution.

For the GPS sample, we find the best fit to be  $\Delta \log_{10}(\Phi) = 0.04 \pm 0.07 * z + 1.11 \pm 0.05$  for the sloped line, and  $\Delta \log_{10}(\Phi) = 1.14 \pm 0.03$  for the horizontal line. The slope we find is approximately zero, therefore we can conclude that the offsets between the GPS sample and the HF sample do not show a redshift evolution. To more reliably state this, we calculate the reduced  $\chi^2$  values for both fits, to find which model fits best. The sloped fit for the GPS sample is  $\chi^2/v = 2.3$ , while the horizontal fit has a value of  $\chi^2/v = 1.8$ . This indicates that the horizontal line provides a better fit, implying at each redshift the ratio between the number of GPS sources and the number of AGN in the HF sample is approximately the same.

For the MPS sample, we find the best fits to be  $\Delta \log_{10}(\Phi) = 0.3 \pm 0.1 * z + 1.0 \pm 0.1$  for the sloped line, and for the horizontal line  $\Delta \log_{10}(\Phi) = 1.22 \pm 0.09$ . The sloped fit has a value of  $\chi^2/v = 0.8$ , while the horizontal fit has a value of  $\chi^2/v = 0.6$ . This suggests a slight preference for the sloped fit. This slope is significantly non-zero, so we find that for the MPS sample, there could indeed be a redshift evolution. At higher redshift, the luminosity function offset between the MPS and LF samples is larger, indicating there are fewer MPS sources compared to our general sample of AGN.

The slope in the MPS offset could be explained by the youth hypothesis. At high redshift, fewer GPS sources have been able to grow into MPS sources. Eventually, the

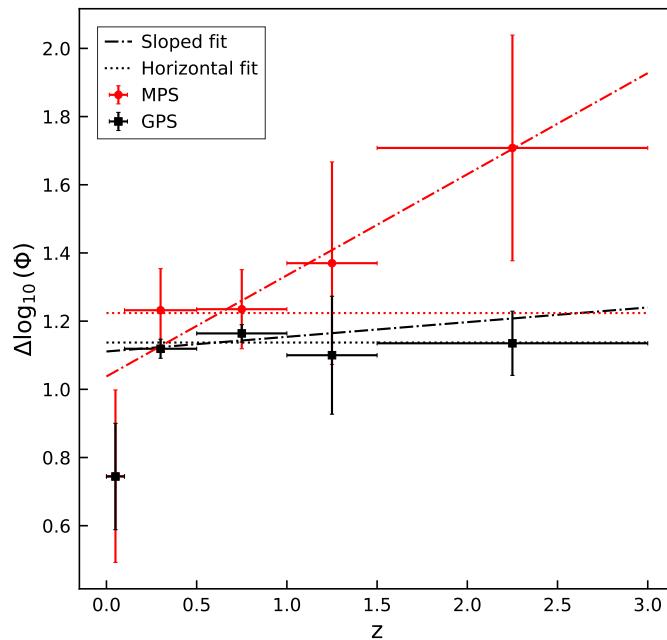
	$0.0 < z < 0.1$	$0.1 < z < 0.5$	$0.5 < z < 1.0$	$1.0 < z < 1.5$	$1.5 < z < 3.0$
$\Delta \log_{10}(\Phi)$ MPS	$0.7 \pm 0.3$	$1.2 \pm 0.1$	$1.2 \pm 0.1$	$1.4 \pm 0.3$	$1.7 \pm 0.3$
$\Delta \log_{10}(\Phi)$ GPS	$0.7 \pm 0.2$	$1.12 \pm 0.03$	$1.16 \pm 0.03$	$1.1 \pm 0.2$	$1.14 \pm 0.09$
$\Delta \log_{10}(\Phi)$ ratio GPS/MPS	$1.0 \pm 0.4$	$0.91 \pm 0.09$	$0.94 \pm 0.09$	$0.8 \pm 0.2$	$0.7 \pm 0.1$

**Table 7.2:** The mean offset in each redshift bin between the luminosity functions of our MPS and LF, respectively our GPS and HF samples, as a function of redshift. The ratio between the offset of the GPS and MPS samples has been indicated as well.

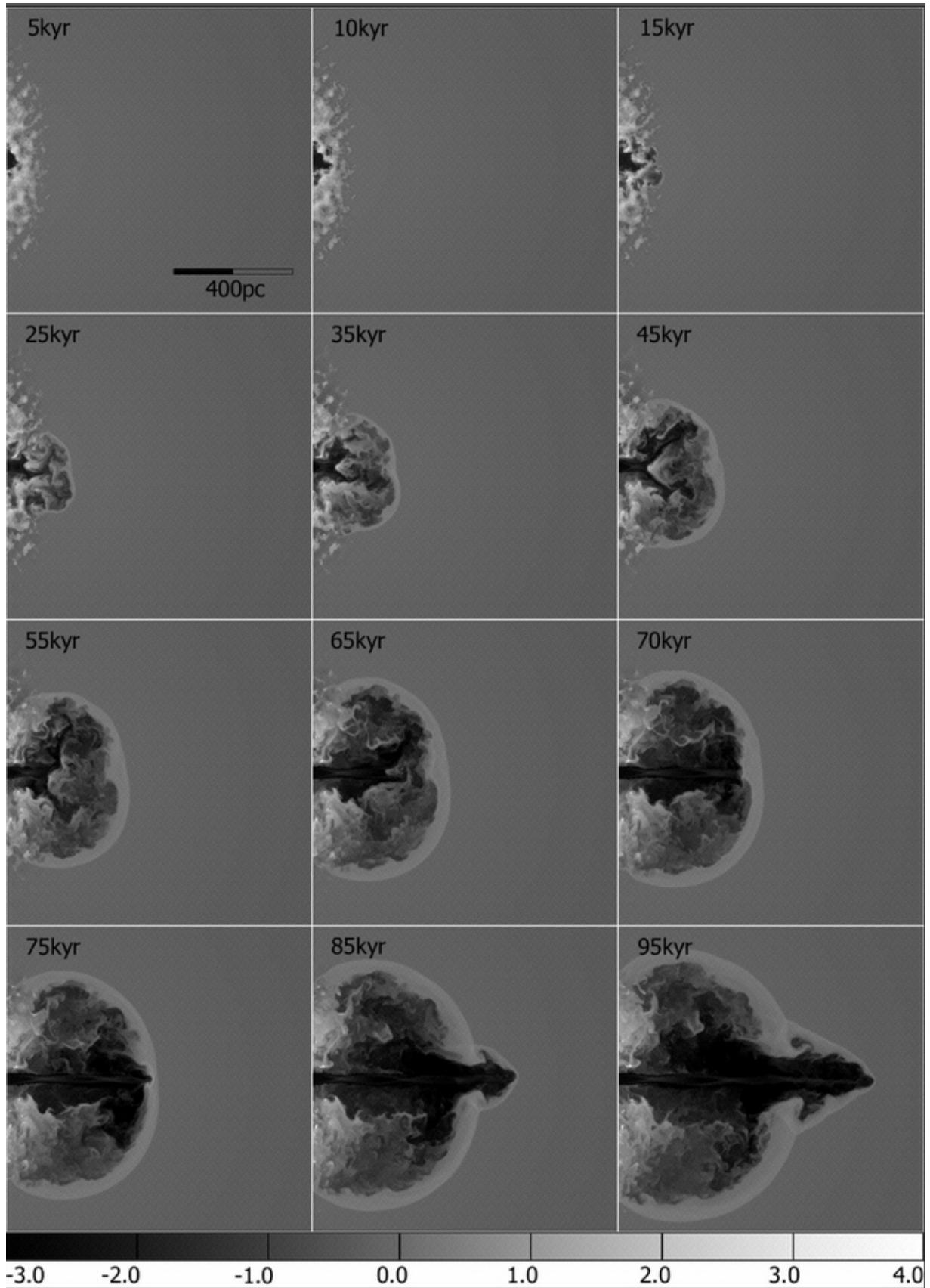
offset should flatten out at low  $z$ , when enough GPS sources have been able to form. We indeed observe this in Figure 7.6. The dip for the lowest redshift bin is likely due to incompleteness.

To explain this evolution with the youth hypothesis, we need to understand the physical circumstances and formation of young AGN. Three-dimensional simulations by Sutherland & Bicknell (2007) demonstrated the interaction of a light hypersonic jet with an inhomogeneous thermal and turbulently supported disk in an elliptical galaxy. The mid-plane densities of such a simulation can be found in Figure 7.7. They identify four phases in the evolution of a jet. The first phase is the flood and channel phase, which occurs at 5, 10, 15, and 25 kyr. In this phase, the jet is making its way through the porous, fractal disk. The pressure in this phase is very high. After this, the energy-driven bubble phase occurs, which can be seen at 35, 45, and 55 kyr. In this phase, a high-pressure, pseudospherical bubble starts to form. This bubble becomes larger than the size of the disk. The jet is still disrupted by the disk, and the bubble grows almost adiabatically. In the 55, 65, and 70 kyr phase, the jet breakout phase starts. There are no more clouds in the path of the jet and it starts to propagate towards the edge of the bubble. Around this phase, the jets can travel much quicker because no more material is in their way. After 75 kyr, the classical phase starts. The jets start to form the classical radio lobe with hot-spot, cocoon, bow shock, and back flow.

If we compare the sizes of the jets in the different phases with the sizes of different categories of PS sources, we find that the GPS phase ends roughly when the linear size is 1 kpc (O'Dea, 1998). A single jet would then be  $\sim 500$  pc, which occurs around the 75 kyr phase in Figure 7.7. This coincides well with the start of the classical phase. We could then conclude that the GPS phase occurs when the jets are traveling quite slowly and constantly through the bubble, but for the MPS phase the jets have already escaped the bubble and are growing roughly twice as fast as the bubble in previous phases. From this, we would predict that the MPS phase is shorter than the GPS phase. We indeed see in Figure 7.6 that the horizontal fit of the MPS sample is higher than that of the GPS sample, corresponding to a shorter lifetime. In the source counts from Chapter 6 we also find that the MPS sample has a slightly larger offset than the GPS sample. However, we warrant caution with this interpretation. Additional MPS sources are required to be identified to increase the significance of this result.



**Figure 7.6:** The mean offset in each redshift bin between the luminosity functions of our MPS and LF, respectively our GPS and HF samples, as a function of redshift. A higher offset corresponds to fewer MPS or GPS sources relative to our LF or HF samples. We find that the offset of the GPS sample does not show a redshift evolution, while the MPS sample seems to have a larger offset at a higher redshift. The lowest redshift bin is likely affected by incompleteness for both samples.



**Figure 7.7:** This image, taken from Sutherland & Bicknell (2007), shows the logarithm of the density at various phases of evolution in their simulation. In the images from 5 to 25 kyr, we find the flood and channel phases. The panels from 35 to 55 kyr are representative of the energy-driven bubble phase. In panels 55 to 70 kyr we find the jet breakout phase. In the images from 75 to 95 kyr and beyond the classical phase was depicted.

# 8. Discussion

Using the samples selected in this work, we are able to study the evolution between MPS and GPS sources for the first time. By constructing the source counts and luminosity functions, we find a difference in offset between the two samples, which allows us to draw conclusions about the lifetimes of the populations. Below, we discuss the implications of these findings, as well as any caveats to our work.

## 8.1 Luminosity function offset evolution

In Figure 7.6, we found evidence for a redshift evolution in the luminosity functions of our MPS sample. We found that at higher redshifts, our MPS sample appears to have a larger offset, which could be explained by the youth hypothesis. MPS sources would be lower in number density at higher redshift, as all MPS sources must have been GPS sources first. At higher redshift, fewer sources have been able to make the transition from a GPS source to an MPS source.

We do not observe a redshift evolution in the offsets of the GPS luminosity functions. This would indicate that an equilibrium is already occurring, with the same number of sources going from HFP sources to GPS sources, as the number of sources going from GPS to MPS sources. More research into HFP and high turnover-frequency GPS sources could perhaps tell us why we only observe the redshift evolution in the MPS sample, and not in the GPS sample.

Further, we find that the median offset of the MPS sample is slightly higher than offsets of the GPS sample, in both the computed source counts and the luminosity functions. This might be evidence that MPS sources are evolving faster than GPS sources. This hypothesis is supported by simulations of a light hypersonic jet with a nonuniform interstellar medium, as done by Sutherland & Bicknell (2007). They find that the jet break-out phase in their simulations occurs when the linear size of a source is roughly 1kpc, which corresponds to the transition between GPS and MPS sources. After the jet break-out phase, so when a source is an MPS source, the jet evolves faster, which could correspond to shorter lifetimes. This result warrants further research.

## 8.2 Accretion mode classifications

In Chapter 5 we computed the accretion mode classifications, which told us that PS sources are more likely to be high-excitation galaxies than large-scale AGN. Our conclusions for the GPS sample were reliable as they were based on large enough counts,

but for the MPS sample we did not have enough sources classified to properly perform the statistical test. More optical spectra would be required to classify a larger amount of MPS sources, to increase the reliability of these arguments.

### 8.3 Frequency selection window

We should carefully consider the validity of comparing our samples of PS sources with those from literature. Different definitions and frequency selection windows can significantly affect the number of sources identified.

In this work we can only identify GPS sources that peak between approximately 250 MHz and 2 GHz. However, GPS sources in literature are defined as those that can peak between  $\sim 500$  MHz and  $\sim 5$  GHz. We are thus unable to identify GPS sources that peak above  $\sim 2$  GHz. This will affect our source counts and luminosity functions, as we are likely missing a large fraction of GPS sources.

In order to capture these higher-frequency GPS sources, we would require a sensitive wide-field survey of the northern sky around  $\sim 5$  GHz. Such a survey does not exist yet.

Even though we are unable to observe GPS sources that peak above  $\sim 2$  GHz, we still find  $\sim 6$  times more GPS sources than the largest sample of GPS sources so far, which is the sample by Callingham et al. (2017). If a suitable high-frequency radio survey were to exist, we could increase this number even further.

We can identify MPS sources that peak between 50 MHz and 1 GHz. In literature, MPS sources are defined as any source that peaks below 1 GHz. Below 30 MHz we cannot observe these due to atmospheric cutoff (O'Dea, 1998). Our selection window thus matches the definition of MPS sources very well, and we do not expect to miss a lot of MPS sources as a result of our selection window.

The difference in frequency selection window for the GPS sources might become problematic in comparing the source counts and luminosity functions from our samples and those from literature. We have suggested that the offset between the GPS sample and the HF master sample is caused by a difference in lifetimes. If we change the frequency window in which we are able to select GPS sources, these offsets might change as well, influencing our conclusions about the lifetimes.

It should further be noted that the definitions of the different classes of PS sources are not based on any physical parameters, but rather on observational limits. The frequency windows that define when a source is an MPS source or a GPS source have arbitrary boundaries, and a different definition might be just as valid. This will influence the source counts and luminosity functions further.

### 8.4 PS selection criteria

Not only does the frequency window influence our sample sizes, but the definitions of when a source is considered a PS source also differ among different samples from literature. We define a source to be a PS source if  $\alpha_{low} > \sigma_{\alpha\_low}$  and  $\alpha_{high} < 0$ .

Slob et al. (2022) determine that a source is a PS source when  $\alpha_{low} > 0.1$  and  $\alpha_{high} < 0$ . Callingham et al. (2017) define multiple sub-samples of PS sources, where their high-frequency soft sample is defined as  $\alpha_{low} > 0.1$  and  $\alpha_{high} < -0.5$  and contains sources with a turnover between  $\sim 200$  MHz and  $\sim 1.4$  GHz. They also define a high-frequency hard sample, where  $\alpha_{low} > 0.1$  and  $-0.5 < \alpha_{high} < 0$ , which consists of

sources with higher frequency turnovers than their high-frequency soft sample.

These works, which employ similar methods to ours, thus have different classifications for when a source can be considered a PS source. This effect will propagate to the sample size, especially when there is a dense population of sources near these boundaries. We have chosen our boundaries such that the limit on  $\alpha_{high}$  is the same as in Callingham et al. (2017) and Slob et al. (2022), but have defined a different selection criterion for  $\alpha_{low}$ . This should be considered when comparing samples to each other. Our method is more statistically justified, as we take into account the individual uncertainties of each source, while these other works use arbitrary values for the spectral index limit.

# 9

# . Conclusions & future outlook

## 9.1 Conclusions

The recent revolution in wide-field radio surveys, both in sensitivity, sky area, resolution, and frequency range, has made the study of statistical samples of PS sources possible for the first time. In this work, we present a sample of 8,975 GPS sources with spectral turnovers near 1400 MHz, and a sample of 522 MPS sources with turnovers near 144 MHz. These samples have been selected using LoTSS, LoLSS, NVSS, and VLASS. Our MPS sample is 1.4 times larger than any previously known sample of PS sources in the same frequency range. Our GPS sample is over six times larger than the previously known largest sample of GPS sources.

We first selected all unresolved, isolated sources in LoTSS with a crossmatch in NVSS. This sample was crossmatched to LoLSS to create our LF master sample of 13,891 sources, and it was crossmatched to all isolated sources in VLASS to create our HF master sample of 120,041 sources.

In our LF master sample, sources were classified as PS sources if the power law spectral index between LoLSS and LoTSS is positive, and the spectral index between LoTSS and NVSS is negative. Similarly for the HF sample, a source is considered a PS source if the spectral index between LoTSS and NVSS is positive, and the spectral index between NVSS and VLASS is negative. This selects sources with a concave SED in our selection window. 3.8% of the sources in our LF sample were identified as MPS sources, and 7.5% of the sources in our HF sample were identified as GPS sources.

We identified four sources with spectral indices  $> 2.5$ , which is the ultimate limit SSA could achieve. These spectral turnovers must thus be associated with FFA, indicating that these sources are the only sources in our sample from which we can determine that they are likely frustrated.

We classified the accretion modes of the sources in our sample using emission line criteria, following the methods described by Best & Heckman (2012). We found, when accounting for a redshift bias, that the fraction of high-excitation radio galaxies is higher in our PS samples than in our HF and LF samples. We determined that this difference was statistically significant for our GPS sample, however we did not have enough classified sources in our MPS sample to draw reliable conclusions. The higher fraction of HERGs in our GPS sample suggests that these represent a population undergoing rapid evolution, agreeing with the youth hypothesis.

We further calculated the 144 MHz Euclidean normalized source counts for our MPS and GPS samples and found that their shapes match those of the source counts of large-scale AGN, scaled down by a factor  $\sim 45$  for the MPS sample and a factor  $\sim 42$  for the GPS

sample. These offsets imply that both MPS and GPS sources have shorter lifetimes than large-scale AGN, with MPS sources tending to have slightly shorter lifetimes compared to GPS sources. We suspect that the GPS sample is more likely to be contaminated by blazars, which affects the shape of the GPS source counts slightly.

Source counts do not account for redshift evolution, therefore we computed the 144 MHz respectively 1400 MHz luminosity functions for our MPS and GPS samples. Using the LoTSS value-added optical catalog, we were able to obtain redshifts for the sources in our sample. The shape of the PS luminosity functions roughly matches that of their respective master samples, but with significant offsets. We find that our samples are likely affected by optical completeness at high redshifts, but conclude that this will affect our master samples and PS samples in the same way. Therefore the offsets between these samples remain reliable.

The luminosity function of the GPS sample has an offset of approximately  $\sim 14$  to the luminosity function of the HF sample, which remains relatively constant over redshift. The MPS sample has an average offset of about  $\sim 17$ , which appears to evolve with redshift. The youth model suggests that this evolution is caused by sources first evolving out of their GPS phase before becoming MPS sources. The offsets also imply that MPS sources generally have slightly shorter lifetimes compared to GPS sources. We find from simulations by Sutherland & Bicknell (2007) that these shorter lifetimes are likely caused by the MPS phase coinciding with the jet break-out phase of PS sources, which increases the speed at which MPS sources evolve compared to GPS sources.

## 9.2 Future outlook

### 9.2.1 Completion of wide-field radio survey used in this research

The survey area of the samples we defined is limited by the survey with the smallest sky area considered, which is LoTSS for our HF and GPS sample. It currently has covered 27% of the northern sky (Shimwell et al., 2022). Upon completion of LoTSS, we thus expect to find roughly three times as many sources, which could give us 33,000 GPS sources.

The MPS and LF samples are limited by the LoLSS sky area, which will also eventually cover the entire northern sky. Currently, 650 degree<sup>2</sup> has been observed (de Gasperin et al., 2023), which is only  $\sim 3\%$  of the northern sky. We would then expect to find roughly 17,000 MPS sources once both LoLSS and LoTSS have been completed.

For the selection of our GPS sample, we make use of the VLASS QL epoch 2 catalog. This is still a preliminary data release. Single Epoch (SE) component catalogs and, eventually, cumulative VLASS component catalogs will be released. The SE catalogs will have a better dynamic range close to bright sources by about a factor of two (Lacy, 2022). Furthermore, these catalogs will have better peak flux densities, since no self-calibration was applied to the QL images, which will be applied to the SE images. Otherwise, the SE images will be similar to the QL images.

VLASS was only used to constrain  $\alpha_{high}$  for the HF sample, which is less important in classifying PS sources than  $\alpha_{low}$ .  $\alpha_{high}$  does not constrain us in whether there is a spectral turnover, rather it tells us roughly at which frequency it occurs. Therefore the preliminary QL catalogs were deemed reliable for this work.

### 9.2.2 Future wide-field radio surveys

LOFAR is a pathfinder for the Square Kilometer Array (SKA), which is a radio telescope currently being built in Australia and South Africa. SKA's low-frequency array will be eight times more sensitive and will have a 135 times faster survey speed than LOFAR, which is currently the most sensitive radio telescope in the low-frequency range. SKA's mid-frequency array will be almost five times more sensitive and will have a 60 times faster survey speed than the VLA (SKAO, 2022). SKA will provide opportunities for studying larger and more sensitive samples of PS sources than ever before.

In the nearer future, we could expand our work to the southern sky using the Australian SKA Pathfinder (ASKAP; Johnston et al., 2007) in Western Australia. ASKAP is a synthesis array consisting of 36 dish antennas of 12m in diameter each. The dishes are spread out in two dimensions with baselines up to 6km. This telescope, which is a pathfinder for SKA just like LOFAR, is in the process of producing the Rapid ASKAP Continuum Survey (McConnell et al., 2020), which will survey the entire southern sky visible from the ASKAP site in Western Australia, at 887.5 MHz(RACS-low), 1367.5 MHz(RACS-mid), and 1667.5 MHz(RACS-high). RACS-low has been completed (Hale et al., 2021), resulting in the deepest radio survey of the full Southern sky to date at this frequency. It has a 95% completeness at 5 mJy for point sources, which is less sensitive than LoTSS and VLASS, but more sensitive than NVSS. When all three of the RACS surveys have been completed, this will provide us with a treasure chamber of new potential PS sources. The first data release of RACS-mid has been published (Duchesne et al., 2023), and RACS-high will follow in the coming years.

### 9.2.3 SED modeling

We classify PS samples using spectral index measurements between two pairs of flux density measurements. This is often reliable, though it makes the SED of individual sources vulnerable to variability. On a large statistical sample variability will not affect the results we find, as it is a symmetric effect. We do not consider the SEDs of individual sources in the conclusions of this work. The method we use limits us in determining the exact spectral turnover frequency and flux density. We can only provide a rough frequency window in which we know our sources peak. Spectral index fitting on more than three flux density measurements will allow us to determine the turnover frequency, which we can use to classify the absorption mechanism dominant in individual sources. We already have a subsample of 10,281 sources that are present in both the LF and HF samples. These sources have at least four flux density measurements. We could then determine the spectral turnover frequency of roughly between  $\sim 400$  and  $\sim 800$  PS sources, based on the percentage of MPS and GPS present in the LF respectively HF samples. The four flux density measurements from the LF and HF samples will be supplemented by detections in the surveys mentioned in Section 2.3.

Determining the turnover frequency with spectral modeling could help us determine the linear sizes of these sources according to Equation 1.1 by O'Dea (1998). We could then plot these sources in a diagram of linear size against radio power. Such a diagram provides a snapshot of radio source evolution, where individual sources will trace out trajectories on the plane. Comparing our MPS and GPS samples on this plane will be valuable to further study the evolution between these sources.

### 9.2.4 Luminosity function modeling

Snellen et al. (2000) have studied the luminosity evolution of GPS sources, in which they argue that GPS sources increase in luminosity, while large-scale radio sources decrease in luminosity with time. This evolution follows for ram-pressure confined radio sources in a surrounding medium with a King profile density. They could not derive local luminosity functions, which would be required to confirm such a model, due to a low number of GPS sources known at the time. With our sample, we could model this luminosity evolution. Understanding and modeling the evolution of GPS and MPS sources could help understand the slope of the luminosity functions, and help us understand the evolution between our MPS and GPS samples.

### 9.2.5 Potential high-frequency peaked spectrum sources

In this work, we only classified the sources in the bottom right quadrant of the color-color plots in Figure 4.1 and Figure 4.3. These are the sources with a positive  $\alpha_{low}$  and a negative  $\alpha_{high}$ . We know that any source with a positive spectral index must turn over at some frequency, therefore any source in the upper right quadrant of our color-color plots must have a turnover at higher frequencies than 1400 MHz for the LF sample and higher than 3 GHz for the HF sample. We have already shown that the majority of such sources in the LF sample indeed get classified as GPS in the HF sample. There are then 3,150 sources in the HF sample that might be GPS or HFP sources. Of these sources, we only know that the spectral turnover occurs  $> 3 \text{ GHz}$ , but we could still use this sample to study the earliest stages of evolution in the youth hypothesis.

# Bibliography

- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS , 182, 543, doi: [10.1088/0067-0049/182/2/543](https://doi.org/10.1088/0067-0049/182/2/543)
- Ahumada, R., Allende Prieto, C., Almeida, A., et al. 2020, ApJS , 249, 3, doi: [10.3847/1538-4365/ab929e](https://doi.org/10.3847/1538-4365/ab929e)
- An, T., & Baan, W. A. 2012, The Astrophysical Journal, 760, 77, doi: [10.1088/0004-637x/760/1/77](https://doi.org/10.1088/0004-637x/760/1/77)
- Baum, S. A., O'Dea, C. P., Murphy, D. W., & de Bruyn, A. G. 1990, A&A , 232, 19
- Becker, R. H., White, R. L., & Edwards, A. L. 1991, ApJS , 75, 1, doi: [10.1086/191529](https://doi.org/10.1086/191529)
- Becker, R. H., White, R. L., & Helfand, D. J. 1995, The Astrophysical Journal, 450, 559, doi: [10.1086/176166](https://doi.org/10.1086/176166)
- Best, P. N., & Heckman, T. M. 2012, Monthly Notices of the Royal Astronomical Society, 421, 1569, doi: [10.1111/j.1365-2966.2012.20414.x](https://doi.org/10.1111/j.1365-2966.2012.20414.x)
- Bicknell, G. V., Dopita, M. A., & O'Dea, C. P. O. 1997, ApJ , 485, 112, doi: [10.1086/304400](https://doi.org/10.1086/304400)
- Buttiglione, S., & Capetti, A. 2010, in Astronomical Society of the Pacific Conference Series, Vol. 427, Accretion and Ejection in AGN: a Global View, ed. L. Maraschi, G. Ghisellini, R. Della Ceca, & F. Tavecchio, 147
- Callingham, J. R., Gaensler, B. M., Ekers, R. D., et al. 2015, The Astrophysical Journal, 809, 168, doi: [10.1088/0004-637x/809/2/168](https://doi.org/10.1088/0004-637x/809/2/168)
- Callingham, J. R., Ekers, R. D., Gaensler, B. M., et al. 2017, The Astrophysical Journal, doi: [10.48550/ARXIV.1701.02771](https://doi.org/10.48550/ARXIV.1701.02771)
- Carvalho, J. C. 1985, MNRAS , 215, 463, doi: [10.1093/mnras/215.3.463](https://doi.org/10.1093/mnras/215.3.463)
- Ceraj, L., Smolčić, V., Delvecchio, I., et al. 2018, Astronomy & Astrophysics, 620, A192, doi: [10.1051/0004-6361/201833935](https://doi.org/10.1051/0004-6361/201833935)
- Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, The Pan-STARRS1 Surveys, arXiv, doi: [10.48550/ARXIV.1612.05560](https://doi.org/10.48550/ARXIV.1612.05560)
- Chilingarian, I. V., Melchior, A.-L., & Zolotukhin, I. Y. 2010, MNRAS , 405, 1409, doi: [10.1111/j.1365-2966.2010.16506.x](https://doi.org/10.1111/j.1365-2966.2010.16506.x)

- Chilingarian, I. V., & Zolotukhin, I. Y. 2012, MNRAS , 419, 1727, doi: [10.1111/j.1365-2966.2011.19837.x](https://doi.org/10.1111/j.1365-2966.2011.19837.x)
- Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, The Astronomical Journal, 115, 1693, doi: [10.1086/300337](https://doi.org/10.1086/300337)
- Dallacasa, D., Stanghellini, C., Centonza, M., & Fanti, R. 2001, Astronomy and Astrophysics, 363
- de Gasperin, F., Williams, W. L., Best, P., et al. 2021, Astronomy & Astrophysics, 648, A104, doi: [10.1051/0004-6361/202140316](https://doi.org/10.1051/0004-6361/202140316)
- de Gasperin, F., Edler, H. W., Williams, W. L., et al. 2023, The LOFAR LBA Sky Survey II. First data release, arXiv, doi: [10.48550/ARXIV.2301.12724](https://doi.org/10.48550/ARXIV.2301.12724)
- de Vries, N., Snellen, I. A. G., Schilizzi, R. T., & Mack, K. H. 2009, Astronomische Nachrichten, 330, 214, doi: [10.1002/asna.200811159](https://doi.org/10.1002/asna.200811159)
- Dewdney, P., Hall, P., Schilizzi, R., & Lazio, T. 2009, Proceedings of the IEEE, 97, 1482, doi: [10.1109/jproc.2009.2021005](https://doi.org/10.1109/jproc.2009.2021005)
- Douglas, J. N., Bash, F. N., Bozian, F. A., Torrence, G. W., & Wolfe, C. 1996, VizieR Online Data Catalog, VIII/42
- Duchesne, S. W., Thomson, A. J. M., Pritchard, J., et al. 2023, arXiv e-prints, arXiv:2306.07194, doi: [10.48550/arXiv.2306.07194](https://doi.org/10.48550/arXiv.2306.07194)
- Duncan, K. J., Sabater, J., Röttgering, H. J. A., et al. 2019, Astronomy & Astrophysics, 622, A3, doi: [10.1051/0004-6361/201833562](https://doi.org/10.1051/0004-6361/201833562)
- Duncan, K. J., Kondapally, R., Brown, M. J. I., et al. 2021, Astronomy & Astrophysics, 648, A4, doi: [10.1051/0004-6361/202038809](https://doi.org/10.1051/0004-6361/202038809)
- Fanti, R., Fanti, C., Schilizzi, R. T., et al. 1990, A&A , 231, 333
- Franzen, T., Seymour, N., Sadler, E., et al. 2021, Publications of the Astronomical Society of Australia
- Gehrels, N. 1986, ApJ , 303, 336, doi: [10.1086/164079](https://doi.org/10.1086/164079)
- Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., et al. 2019, AJ , 157, 98, doi: [10.3847/1538-3881/aafc33](https://doi.org/10.3847/1538-3881/aafc33)
- Gopal-Krishna, Patnaik, A. R., & Steppe, H. 1983, A&A , 123, 107
- Gordon, Y. 2023, CIRADA: VLASS Catalog User Guide. <https://cirada.ca/catalogues>
- Gordon, Y. A., Rudnick, L., Andernach, H., et al. 2023, arXiv e-prints, arXiv:2303.12830, doi: [10.48550/arXiv.2303.12830](https://doi.org/10.48550/arXiv.2303.12830)
- Gregory, P. C., & Condon, J. J. 1991, The Astrophysical Journal Supplement Series, 75, 1011, doi: [10.1086/191559](https://doi.org/10.1086/191559)

- Hale, C. L., McConnell, D., Thomson, A. J. M., et al. 2021, Publications of the Astronomical Society of Australia, 38, doi: [10.1017/pasa.2021.47](https://doi.org/10.1017/pasa.2021.47)
- Hales, S. E. G., Riley, J. M., Waldram, E. M., Warner, P. J., & Baldwin, J. E. 2007, Monthly Notices of the Royal Astronomical Society, 382, 1639, doi: [10.1111/j.1365-2966.2007.12392.x](https://doi.org/10.1111/j.1365-2966.2007.12392.x)
- Hancock, P. J., Murphy, T., Gaensler, B. M., Hopkins, A., & Curran, J. R. 2012, Monthly Notices of the Royal Astronomical Society, 422, 1812, doi: [10.1111/j.1365-2966.2012.20768.x](https://doi.org/10.1111/j.1365-2966.2012.20768.x)
- Hancock, P. J., Trott, C. M., & Hurley-Walker, N. 2018, Publications of the Astronomical Society of Australia, 35, doi: [10.1017/pasa.2018.3](https://doi.org/10.1017/pasa.2018.3)
- Healey, S. E., Romani, R. W., Taylor, G. B., et al. 2007, ApJS , 171, 61, doi: [10.1086/513742](https://doi.org/10.1086/513742)
- Heckman, T., & Best, P. 2014, Annual Review of Astronomy and Astrophysics, doi: [10.48550/ARXIV.1403.4620](https://doi.org/10.48550/ARXIV.1403.4620)
- Hill, G. J., Gebhardt, K., Komatsu, E., et al. 2008, Panor. Views Galaxy Form. Evol., 399, 115. <https://arxiv.org/abs/0806.0183>
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, The Astrophysical Journal Supplement Series, 208, 19, doi: [10.1088/0067-0049/208/2/19](https://doi.org/10.1088/0067-0049/208/2/19)
- Hunter, J. D. 2007, Computing In Science & Engineering, 9, 90
- Intema, H. T., Jagannathan, P., Mooley, K. P., & Frail, D. A. 2017, Astronomy & Astrophysics, 598, A78, doi: [10.1051/0004-6361/201628536](https://doi.org/10.1051/0004-6361/201628536)
- Johnston, S., Bailes, M., Bartel, N., et al. 2007, PASA , 24, 174, doi: [10.1071/AS07033](https://doi.org/10.1071/AS07033)
- Kaiser, C. R., & Best, P. N. 2007, Monthly Notices of the Royal Astronomical Society, 381, 1548, doi: [10.1111/j.1365-2966.2007.12350.x](https://doi.org/10.1111/j.1365-2966.2007.12350.x)
- Kapahi, V. K. 1981, A&AS , 43, 381
- Kovalev, Y. Y. 2005, Baltic Astronomy, 14, 413
- Kovalev, Y. Y., Kovalev, Y. A., Nizhelsky, N. A., & Bogdantsov, A. B. 2002, PASA , 19, 83, doi: [10.1071/AS01109](https://doi.org/10.1071/AS01109)
- Kunert-Bajraszewska, M., Gawroński, M. P., Labiano, A., & Siemiginowska, A. 2010, Monthly Notices of the Royal Astronomical Society, 408, 2261, doi: [10.1111/j.1365-2966.2010.17271.x](https://doi.org/10.1111/j.1365-2966.2010.17271.x)
- Labiano, A., Barthel, P. D., O'Dea, C. P., et al. 2006, Astronomy & Astrophysics, 463, 97, doi: [10.1051/0004-6361:20066183](https://doi.org/10.1051/0004-6361:20066183)
- Lacy, M. 2022, Vlass project memo: 17 characterization of vlass single epoch continuum ... [https://library.nrao.edu/public/memos/vla/vlass/VCLASS\\_017.pdf](https://library.nrao.edu/public/memos/vla/vlass/VCLASS_017.pdf)
- Lacy, M., Baum, S. A., Chandler, C. J., et al. 2020, Publications of the Astronomical Society of the Pacific, 132, 035001, doi: [10.1088/1538-3873/ab63eb](https://doi.org/10.1088/1538-3873/ab63eb)

- Laing, R. A., Jenkins, C. R., Wall, J. V., & Unger, S. W. 1994, in Astronomical Society of the Pacific Conference Series, Vol. 54, The Physics of Active Galaxies, ed. G. V. Bicknell, M. A. Dopita, & P. J. Quinn, 201
- Lane, W. M., Cotton, W. D., van Velzen, S., et al. 2014, Monthly Notices of the Royal Astronomical Society, 440, 327, doi: [10.1093/mnras/stu256](https://doi.org/10.1093/mnras/stu256)
- Large, M. I., Cram, L. E., & Burgess, A. M. 1991, The Observatory, 111, 72
- Longair, M. S. 2011, High Energy Astrophysics (Cambridge University Press)
- Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285, doi: [10.1086/300353](https://doi.org/10.1086/300353)
- Mandal, S., Prandoni, I., Hardcastle, M. J., et al. 2021, Astronomy & Astrophysics, 648, A5, doi: [10.1051/0004-6361/202039998](https://doi.org/10.1051/0004-6361/202039998)
- Marecki, A., Falcke, H., Niezgoda, J., Garrington, S. T., & Patnaik, A. R. 1999, A&AS, 135, 273, doi: [10.1051/aas:1999450](https://doi.org/10.1051/aas:1999450)
- Massardi, M., Bonaldi, A., Negrello, M., et al. 2010, Monthly Notices of the Royal Astronomical Society, doi: [10.1111/j.1365-2966.2010.16305.x](https://doi.org/10.1111/j.1365-2966.2010.16305.x)
- Mauch, T., Murphy, T., Buttery, H. J., et al. 2003, Monthly Notices of the Royal Astronomical Society, 342, 1117, doi: [10.1046/j.1365-8711.2003.06605.x](https://doi.org/10.1046/j.1365-8711.2003.06605.x)
- Mauch, T., & Sadler, E. M. 2007, MNRAS, 375, 931, doi: [10.1111/j.1365-2966.2006.11353.x](https://doi.org/10.1111/j.1365-2966.2006.11353.x)
- McConnell, D., Hale, C. L., Lenc, E., et al. 2020, PASA, 37, e048, doi: [10.1017/pasa.2020.41](https://doi.org/10.1017/pasa.2020.41)
- Mohan, N., & Rafferty, D. 2015, PyBDSF: Python Blob Detection and Source Finder, Astrophysics Source Code Library, record ascl:1502.007. <http://ascl.net/1502.007>
- Murphy, T., Sadler, E. M., Ekers, R. D., et al. 2010, Monthly Notices of the Royal Astronomical Society, 402, 2403, doi: [10.1111/j.1365-2966.2009.15961.x](https://doi.org/10.1111/j.1365-2966.2009.15961.x)
- Narayan, R., & Yi, I. 1995, ApJ, 452, 710, doi: [10.1086/176343](https://doi.org/10.1086/176343)
- O'Dea, C., Dent, W., Balonek, T., & Kapitzky, J. 1983, Astronomical Journal (ISSN 0004-6256), vol. 88, Nov. 1983, p. 1616-1625. NSF-supported research., 88, 1616
- O'Dea, C. P. 1998, Publications of the Astronomical Society of the Pacific, 110, 493, doi: [10.1086/316162](https://doi.org/10.1086/316162)
- O'Dea, C. P., Dent, W. A., Kinzel, W. M., & Balonek, T. J. 1986, AJ, 92, 1262, doi: [10.1086/114260](https://doi.org/10.1086/114260)
- O'Dea, C. P., & Saikia, D. J. 2021, The Astronomy and Astrophysics Review, 29, doi: [10.1007/s00159-021-00131-w](https://doi.org/10.1007/s00159-021-00131-w)
- Owsianik, I., & Conway, J. E. 1998, A&A, 337, 69, doi: [10.48550/arXiv.astro-ph/9712062](https://doi.org/10.48550/arXiv.astro-ph/9712062)
- Peacock, J. A., & Wall, J. V. 1982, MNRAS, 198, 843, doi: [10.1093/mnras/198.3.843](https://doi.org/10.1093/mnras/198.3.843)

- Peck, L. 1999, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 3713, Unattended Ground Sensor Technologies and Applications, ed. E. M. Carapezza, D. B. Law, & K. T. Stalker, 188–196, doi: [10.1117/12.357133](https://doi.org/10.1117/12.357133)
- Pracy, M. B., Ching, J. H. Y., Sadler, E. M., et al. 2016, MNRAS , 460, 2, doi: [10.1093/mnras/stw910](https://doi.org/10.1093/mnras/stw910)
- Rengelink, R. B., Tang, Y., de Bruyn, A. G., et al. 1997, Astronomy and Astrophysics Supplement Series, 124, 259, doi: [10.1051/aas:1997358](https://doi.org/10.1051/aas:1997358)
- Sabater, J., Best, P. N., Hardcastle, M. J., et al. 2019, Astronomy & Astrophysics, 622, A17, doi: [10.1051/0004-6361/201833883](https://doi.org/10.1051/0004-6361/201833883)
- Schmidt, M. 1968, ApJ , 151, 393, doi: [10.1086/149446](https://doi.org/10.1086/149446)
- Shimwell, T. W., Hardcastle, M. J., Tasse, C., et al. 2022, Astronomy & Astrophysics, 659, A1, doi: [10.1051/0004-6361/202142484](https://doi.org/10.1051/0004-6361/202142484)
- SKAO. 2022. <https://www.skao.int/en/resources/291/frequently-asked-questions>
- Slob, M. M., Callingham, J. R., Röttgering, H. J. A., et al. 2022, Astronomy & Astrophysics, 668, A186, doi: [10.1051/0004-6361/202244651](https://doi.org/10.1051/0004-6361/202244651)
- Snellen, I. A., Schilizzi, R. T., de Bruyn, A. G., et al. 1998, Astronomy and Astrophysics Supplement Series, 131, 435, doi: [10.1051/aas:1998281](https://doi.org/10.1051/aas:1998281)
- Snellen, I. A. G., Schilizzi, R. T., Miley, G. K., et al. 2000, Monthly Notices of the Royal Astronomical Society, 319, 445, doi: [10.1111/j.1365-8711.2000.03935.x](https://doi.org/10.1111/j.1365-8711.2000.03935.x)
- Sobolewska, M., Siemiginowska, A., Guainazzi, M., et al. 2019, ApJ , 871, 71, doi: [10.3847/1538-4357/aaee78](https://doi.org/10.3847/1538-4357/aaee78)
- Sohn, J., Geller, M. J., Zahid, H. J., & Fabricant, D. G. 2019, ApJ , 872, 192, doi: [10.3847/1538-4357/ab0213](https://doi.org/10.3847/1538-4357/ab0213)
- Sohn, J., Geller, M. J., Zahid, H. J., et al. 2017, ApJS , 229, 20, doi: [10.3847/1538-4365/aa653e](https://doi.org/10.3847/1538-4365/aa653e)
- Souchay, J., Andrei, A. H., Barache, C., et al. 2009, A&A , 494, 799, doi: [10.1051/0004-6361:200809602](https://doi.org/10.1051/0004-6361:200809602)
- Stanghellini, C., Baum, S. A., O'Dea, C. P., & Morris, G. B. 1990, A&A , 233, 379
- Stanghellini, C., O'Dea, C. P., Dallacasa, D., et al. 1998, A&AS , 131, 303, doi: [10.1051/aas:1998270](https://doi.org/10.1051/aas:1998270)
- Sutherland, R. S., & Bicknell, G. V. 2007, ApJS , 173, 37, doi: [10.1086/520640](https://doi.org/10.1086/520640)
- Tadhunter, C. N., Morganti, R., Robinson, A., et al. 1998, MNRAS , 298, 1035, doi: [10.1046/j.1365-8711.1998.01706.x](https://doi.org/10.1046/j.1365-8711.1998.01706.x)
- Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29

- Thomas, D., Steele, O., Maraston, C., et al. 2013, MNRAS , 431, 1383, doi: [10.1093/mnras/stt261](https://doi.org/10.1093/mnras/stt261)
- Tingay, S. J., & de Kool, M. 2003, The Astronomical Journal, 126, 723, doi: [10.1086/376600](https://doi.org/10.1086/376600)
- Urry, C. M., & Padovani, P. 1995, Publications of the Astronomical Society of the Pacific, 107, 803, doi: [10.1086/133630](https://doi.org/10.1086/133630)
- van Breugel, W., Miley, G., & Heckman, T. 1984, AJ , 89, 5, doi: [10.1086/113480](https://doi.org/10.1086/113480)
- van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, Astronomy & Astrophysics, 556, A2, doi: [10.1051/0004-6361/201220873](https://doi.org/10.1051/0004-6361/201220873)
- Wayth, R. B., Lenc, E., Bell, M. E., et al. 2015, PASA , 32, e025, doi: [10.1017/pasa.2015.26](https://doi.org/10.1017/pasa.2015.26)
- Wilkinson, P. N., Booth, R. S., Cornwell, T. J., & Clark, R. R. 1984, Nature , 308, 619, doi: [10.1038/308619a0](https://doi.org/10.1038/308619a0)
- Williams, W. L., Hardcastle, M. J., Best, P. N., et al. 2019, Astronomy & Astrophysics, 622, A2, doi: [10.1051/0004-6361/201833564](https://doi.org/10.1051/0004-6361/201833564)
- Wright, A., & Otrupcek, R. 1990, PKS Catalog (1990, 0
- Wright, A. E., Griffith, M. R., Burke, B. F., & Ekers, R. D. 1994, The Astrophysical Journal Supplement Series, 91, 111, doi: [10.1086/191939](https://doi.org/10.1086/191939)
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ , 140, 1868, doi: [10.1088/0004-6256/140/6/1868](https://doi.org/10.1088/0004-6256/140/6/1868)

# A. 1400 MHz source counts

$\langle S_{144 \text{ MHz}} \rangle [\text{Jy}]$	$N_s$	$S^{5/2} dN/ds_{-\sigma}^{+\sigma} [\text{Jy}^{1.5}/\text{sr}]$
0.004	1691	$0.61^{+0.02}_{-0.01}$
0.006	1326	$0.87^{+0.02}_{-0.02}$
0.009	1143	$1.38^{+0.04}_{-0.04}$
0.013	880	$1.84^{+0.06}_{-0.06}$
0.019	639	$2.44^{+0.1}_{-0.1}$
0.028	493	$3.35^{+0.16}_{-0.15}$
0.041	353	$4.35^{+0.24}_{-0.23}$
0.061	252	$5.46^{+0.37}_{-0.34}$
0.09	176	$6.7^{+0.54}_{-0.5}$
0.133	106	$7.73^{+0.83}_{-0.75}$
0.196	100	$12.14^{+1.34}_{-1.21}$
0.289	64	$14.2^{+2.01}_{-1.77}$
0.427	29	$11.99^{+2.66}_{-2.21}$
0.63	28	$19.14^{+4.34}_{-3.59}$
0.929	17	$21.3^{+6.51}_{-5.11}$
1.759	10	$12.05^{+5.13}_{-3.74}$

**Table A.1:** The 1400 MHz source counts for our GPS sample.  $\langle S \rangle$  is the central flux density of every flux bin in Jy,  $N_s$  is the number of sources in each bin, and  $N_{-\sigma}^{+\sigma}$  are the normalized differential source counts in  $[\text{Jy}^{1.5}/\text{sr}]$ .

# B. Luminosity functions

$\log L [\text{W Hz}^{-1}]$	N	$\log \phi [\text{Mpc}^{-3} \text{ dex}^{-1}]$	$\log L [\text{W Hz}^{-1}]$	N	$\log \phi [\text{Mpc}^{-3} \text{ dex}^{-1}]$
0.0 < $z$ < 0.1; LF Sample					
23.15 ± 0.2	80	$-4.24^{+0.05}_{-0.05}$	23.55 ± 0.3	9	$-5.37^{+0.13}_{-0.19}$
23.45 ± 0.2	44	$-4.52^{+0.06}_{-0.07}$			
23.75 ± 0.2	21	$-4.84^{+0.09}_{-0.11}$			
24.1 ± 0.2	6	$-5.51^{+0.15}_{-0.23}$			
0.1 < $z$ < 0.5; LF Sample					
23.65 ± 0.1	37	$-5.36^{+0.07}_{-0.08}$	24.32 ± 0.4	4	$-6.82^{+0.21}_{-0.34}$
23.9 ± 0.1	48	$-5.55^{+0.07}_{-0.09}$	25.22 ± 0.4	21	$-7.25^{+0.09}_{-0.11}$
24.15 ± 0.1	69	$-5.75^{+0.05}_{-0.06}$	26.12 ± 0.4	9	$-7.66^{+0.12}_{-0.18}$
24.4 ± 0.1	100	$-5.91^{+0.05}_{-0.05}$			
24.65 ± 0.1	140	$-5.9^{+0.04}_{-0.04}$			
24.9 ± 0.1	173	$-5.81^{+0.03}_{-0.03}$			
25.15 ± 0.1	162	$-5.85^{+0.03}_{-0.04}$			
25.4 ± 0.1	108	$-6.02^{+0.04}_{-0.04}$			
25.65 ± 0.1	52	$-6.34^{+0.06}_{-0.06}$			
25.9 ± 0.1	36	$-6.5^{+0.07}_{-0.08}$			
26.15 ± 0.1	21	$-6.73^{+0.09}_{-0.11}$			
26.63 ± 0.4	24	$-7.13^{+0.08}_{-0.1}$			
0.5 < $z$ < 1.0; LF Sample					
26.18 ± 0.1	215	$-6.24^{+0.03}_{-0.04}$	26.2 ± 0.3	26	$-7.59^{+0.08}_{-0.1}$
26.43 ± 0.1	150	$-6.4^{+0.04}_{-0.04}$	26.9 ± 0.3	13	$-7.86^{+0.12}_{-0.17}$
26.68 ± 0.1	65	$-6.77^{+0.06}_{-0.06}$	27.6 ± 0.3	4	$-8.32^{+0.21}_{-0.34}$
26.93 ± 0.1	46	$-6.81^{+0.07}_{-0.09}$			
27.18 ± 0.1	35	$-7.02^{+0.07}_{-0.09}$			
27.43 ± 0.1	12	$-7.52^{+0.11}_{-0.16}$			
27.68 ± 0.1	9	$-7.6^{+0.13}_{-0.18}$			
28.03 ± 0.2	6	$-7.94^{+0.15}_{-0.24}$			
1.0 < $z$ < 1.5; LF Sample					
25.84 ± 0.2	112	$-5.98^{+0.07}_{-0.08}$	26.54 ± 0.4	13	$-7.53^{+0.15}_{-0.22}$
26.24 ± 0.2	145	$-6.0^{+0.07}_{-0.08}$	27.44 ± 0.4	4	$-7.96^{+0.21}_{-0.34}$
1.0 < $z$ < 1.5; MPS Sample					

$26.64 \pm 0.2$	116	$-5.95^{+0.16}_{-0.26}$				
$27.04 \pm 0.2$	54	$-6.51^{+0.09}_{-0.11}$				
$27.44 \pm 0.2$	32	$-6.82^{+0.11}_{-0.14}$				
$27.84 \pm 0.2$	26	$-6.47^{+0.16}_{-0.26}$				
$28.35 \pm 0.3$	8	$-7.88^{+0.16}_{-0.25}$				
$1.5 < z < 3.0$ ; LF Sample			$1.5 < z < 3.0$ ; MPS Sample			
$27.18 \pm 0.2$	95	$-6.89^{+0.09}_{-0.12}$	$27.03 \pm 0.6$	17	$-8.1^{+0.18}_{-0.32}$	
$27.68 \pm 0.2$	55	$-7.13^{+0.17}_{-0.27}$	$28.13 \pm 0.6$	3	$-9.42^{+0.24}_{-0.42}$	
$28.4 \pm 0.5$	42	$-7.45^{+0.17}_{-0.27}$				

**Table B.1:** 144 MHz luminosity functions for the LF and MPS sample, as plotted in Chapter 7. Incomplete bins that were removed are not indicated here.

$\log L[\text{W Hz}^{-1}]$	N	$\log \phi [\text{Mpc}^{-3} \text{dex}^{-1}]$	$\log L[\text{W Hz}^{-1}]$	N	$\log \phi [\text{Mpc}^{-3} \text{dex}^{-1}]$
0.0 < $z$ < 0.1; HF Sample			0.0 < $z$ < 0.1; GPS Sample		
23.25 ± 0.2	182	$-5.06^{+0.03}_{-0.03}$	23.47 ± 0.4	19	$-6.3^{+0.09}_{-0.11}$
23.75 ± 0.2	36	$-5.77^{+0.07}_{-0.08}$	24.37 ± 0.4	7	$-6.73^{+0.14}_{-0.21}$
24.46 ± 0.5	12	$-6.51^{+0.11}_{-0.15}$			
0.1 < $z$ < 0.3; HF Sample			0.1 < $z$ < 0.3; GPS Sample		
23.13 ± 0.1	359	$-4.48^{+0.03}_{-0.03}$	23.41 ± 0.3	149	$-5.85^{+0.05}_{-0.05}$
23.38 ± 0.1	620	$-4.78^{+0.02}_{-0.02}$	24.11 ± 0.3	114	$-6.72^{+0.04}_{-0.04}$
23.63 ± 0.1	895	$-5.01^{+0.01}_{-0.02}$	24.81 ± 0.3	35	$-7.28^{+0.07}_{-0.08}$
23.88 ± 0.1	803	$-5.37^{+0.02}_{-0.02}$	25.51 ± 0.3	8	$-7.92^{+0.13}_{-0.19}$
24.13 ± 0.1	508	$-5.67^{+0.02}_{-0.02}$	26.21 ± 0.3	1	$-8.82^{+0.36}_{-0.99}$
24.38 ± 0.1	301	$-5.9^{+0.02}_{-0.03}$			
24.63 ± 0.1	160	$-6.17^{+0.03}_{-0.04}$			
24.88 ± 0.1	81	$-6.47^{+0.05}_{-0.05}$			
25.13 ± 0.1	46	$-6.71^{+0.06}_{-0.07}$			
25.38 ± 0.1	21	$-7.05^{+0.09}_{-0.11}$			
25.63 ± 0.1	9	$-7.42^{+0.12}_{-0.18}$			
26.02 ± 0.3	13	$-7.6^{+0.11}_{-0.14}$			
0.3 < $z$ < 0.5; HF Sample			0.3 < $z$ < 0.5; GPS Sample		
23.95 ± 0.2	823	$-4.84^{+0.09}_{-0.11}$	24.31 ± 0.4	191	$-6.27^{+0.14}_{-0.2}$
24.35 ± 0.2	2076	$-5.55^{+0.01}_{-0.01}$	25.06 ± 0.4	94	$-7.38^{+0.04}_{-0.05}$
24.75 ± 0.2	1185	$-6.0^{+0.01}_{-0.01}$	25.81 ± 0.4	21	$-8.03^{+0.09}_{-0.11}$
25.15 ± 0.2	456	$-6.42^{+0.02}_{-0.02}$	26.56 ± 0.4	4	$-8.75^{+0.21}_{-0.34}$
25.55 ± 0.2	199	$-6.78^{+0.03}_{-0.03}$	27.31 ± 0.4	1	$-9.35^{+0.36}_{-0.99}$
25.95 ± 0.2	56	$-7.33^{+0.05}_{-0.06}$			
26.35 ± 0.2	33	$-7.56^{+0.07}_{-0.08}$			
26.75 ± 0.2	7	$-8.23^{+0.14}_{-0.21}$			
27.29 ± 0.3	5	$-8.61^{+0.19}_{-0.29}$			
0.5 < $z$ < 0.7; HF Sample			0.5 < $z$ < 0.7; GPS Sample		
24.55 ± 0.2	1687	$-5.18^{+0.09}_{-0.11}$	24.85 ± 0.5	293	$-6.71^{+0.08}_{-0.1}$
24.9 ± 0.2	1834	$-5.97^{+0.01}_{-0.01}$	25.85 ± 0.5	57	$-7.97^{+0.05}_{-0.06}$
25.25 ± 0.2	978	$-6.28^{+0.01}_{-0.01}$	26.85 ± 0.5	8	$-8.78^{+0.14}_{-0.2}$
25.6 ± 0.2	437	$-6.63^{+0.02}_{-0.02}$			
25.95 ± 0.2	227	$-6.9^{+0.03}_{-0.03}$			
26.3 ± 0.2	91	$-7.3^{+0.04}_{-0.05}$			
26.65 ± 0.2	49	$-7.58^{+0.06}_{-0.07}$			
27.13 ± 0.3	21	$-8.17^{+0.09}_{-0.11}$			
0.7 < $z$ < 0.9; HF Sample			0.7 < $z$ < 0.9; GPS Sample		
24.86 ± 0.2	1662	$-5.28^{+0.09}_{-0.12}$	24.91 ± 0.2	169	$-6.64^{+0.1}_{-0.14}$
25.21 ± 0.2	1845	$-5.94^{+0.02}_{-0.03}$	25.41 ± 0.2	107	$-7.38^{+0.05}_{-0.06}$
25.56 ± 0.2	1011	$-6.06^{+0.07}_{-0.09}$	25.91 ± 0.2	48	$-7.47^{+0.18}_{-0.32}$
25.91 ± 0.2	502	$-6.54^{+0.04}_{-0.04}$	26.41 ± 0.2	16	$-8.25^{+0.1}_{-0.13}$

$26.26 \pm 0.2$	248	$-6.8^{+0.03}_{-0.04}$	$26.91 \pm 0.2$	9	$-8.46^{+0.13}_{-0.19}$
$26.61 \pm 0.2$	112	$-7.19^{+0.05}_{-0.06}$			
$26.96 \pm 0.2$	45	$-7.55^{+0.07}_{-0.08}$			
$27.47 \pm 0.3$	20	$-8.14^{+0.1}_{-0.13}$			
$0.9 < z < 1.4$ ; HF Sample			$0.9 < z < 1.4$ ; GPS Sample		
$25.29 \pm 0.2$	1429	$-5.57^{+0.11}_{-0.15}$	$25.2 \pm 0.3$	152	$-6.93^{+0.06}_{-0.07}$
$25.59 \pm 0.2$	1126	$-5.71^{+0.15}_{-0.23}$	$25.8 \pm 0.3$	151	$-7.14^{+0.09}_{-0.12}$
$25.89 \pm 0.2$	834	$-6.06^{+0.06}_{-0.06}$	$26.4 \pm 0.3$	72	$-7.88^{+0.06}_{-0.07}$
$26.19 \pm 0.2$	533	$-6.4^{+0.04}_{-0.04}$	$27.0 \pm 0.3$	17	$-8.38^{+0.14}_{-0.2}$
$26.49 \pm 0.2$	336	$-6.7^{+0.04}_{-0.05}$	$27.6 \pm 0.3$	9	$-8.79^{+0.17}_{-0.28}$
$26.79 \pm 0.2$	179	$-6.81^{+0.08}_{-0.1}$			
$27.09 \pm 0.2$	102	$-7.15^{+0.07}_{-0.09}$			
$27.39 \pm 0.2$	55	$-7.31^{+0.13}_{-0.19}$			
$27.84 \pm 0.3$	28	$-8.12^{+0.11}_{-0.14}$			
$1.4 < z < 3.0$ ; HF Sample			$1.4 < z < 3.0$ ; GPS Sample		
$25.69 \pm 0.2$	563	$-6.43^{+0.14}_{-0.2}$	$25.68 \pm 0.3$	87	$-7.51^{+0.13}_{-0.19}$
$25.99 \pm 0.2$	656	$-6.61^{+0.07}_{-0.08}$	$26.23 \pm 0.3$	135	$-7.54^{+0.14}_{-0.21}$
$26.29 \pm 0.2$	677	$-6.6^{+0.08}_{-0.1}$	$26.78 \pm 0.3$	119	$-7.76^{+0.17}_{-0.27}$
$26.59 \pm 0.2$	612	$-6.64^{+0.15}_{-0.22}$	$27.33 \pm 0.3$	52	$-8.58^{+0.08}_{-0.1}$
$26.89 \pm 0.2$	394	$-6.4^{+0.22}_{-0.48}$	$27.88 \pm 0.3$	20	$-8.99^{+0.18}_{-0.31}$
$27.19 \pm 0.2$	285	$-7.12^{+0.1}_{-0.13}$	$28.43 \pm 0.3$	5	$-9.77^{+0.19}_{-0.29}$
$27.49 \pm 0.2$	154	$-7.42^{+0.13}_{-0.19}$			
$27.79 \pm 0.2$	84	$-8.09^{+0.07}_{-0.09}$			
$28.09 \pm 0.2$	37	$-8.46^{+0.11}_{-0.15}$			
$28.41 \pm 0.2$	22	$-8.64^{+0.14}_{-0.21}$			

**Table B.2:** 1400 MHz luminosity functions for the HF and GPS sample, as plotted in Chapter 7. Incomplete bins that were removed are not indicated here.