A study of quasars: an investigation into the molecular gas of a high-redshift quasar and the radio loudness of radio-quiet quasars.

Hana Josephine Schumacher
Institute of Cosmology and Gravitation
University of Portsmouth

Thesis submitted in fulfilment of the requirements for the award of the degree of Doctor of Philosophy of the University of Portsmouth

April 2013
Abstract

This thesis is composed of two parts; the first part deals with observations of the molecular gas towards an unlensed, obscured quasar AMS12, and the second part investigates radio undetected, optically selected quasi-stellar objects (QSOs) to determine the nature of the radio flux density distributions of these objects. AMS12 is an unlensed, obscured, $z = 2.767$ quasar which we observed with the Plateau de Bure Interferometer to detect carbon monoxide rotational transitions and atomic carbon fine structure lines in the molecular gas. We present new detections of the CO(5-4), CO(7-6), [CI](3P_1 - 3P_0) and [CI](3P_2 - 3P_1) molecular and atomic line transitions in this thesis. AMS12 is the first unlensed, high redshift source to have both atomic carbon ([CI]) transitions detected. The highly-excited molecular gas probed by CO(3-2), (5-4) and (7-6), is modelled with large velocity gradient models. The gas kinetic temperature $T_G$, density $n(H_2)$, and the characteristic size $r_0$, are determined using the dust temperature from the far-infrared spectral energy distribution which had the following best-fitting parameters $\log_{10}[L_{\text{FIR}}/L_\odot] = 13.5$, dust temperature $T_D = 88$ K and emissivity index $\beta = 0.6$, as a prior for the gas temperature. The best fitting parameters are $T_G = 89.6$ K, $n(H_2) = 10^{3.9}$ cm$^{-3}$ and $r_0 = 0.8$ kpc. The ratio of the [CI] lines gives a [CI] excitation temperature of $43 \pm 10$ K, indicating the [CI] and the high-excitation CO are not in thermal equilibrium. The [CI] excitation temperature is below that of the dust temperature and the gas kinetic temperature of the high-excitation CO, perhaps because [CI] lies at a larger radius where there may also be a large reservoir of CO at a cooler temperature, which may be detectable through the CO(1-0). Using the [CI](3P_1 - 3P_0) line we can estimate the strength of the CO(1-0) line and hence the gas mass. This suggests that a significant fraction ($\sim 30\%$) of the molecular gas is missed from the high-excitation line analysis, giving a gas mass higher than that inferred from the assumption that the high-excitation gas is a good tracer of the low-excitation gas. The stellar mass was estimated from the mid-/near-infrared spectral energy distribution to be $M_\star \sim 3 \times 10^{11} M_\odot$. The Eddington limited black hole mass is found from the bolometric luminosity to be $M_\bullet \gtrsim 1.5 \times 10^9 M_\odot$. These give a black hole - bulge mass ratio of $M_\bullet/M_\star \gtrsim 0.005$. This is in agreement with studies on the evolution of the $M_\bullet/M_\star$ relationship at high redshifts, which find a departure from the local value $\sim 0.002$.

In the second half of the thesis we investigate the possible existence of a lower envelope in the radio luminosity versus optical luminosity plane. We select a population of QSOs from the Sloan Digital Sky Survey photometric quasar catalogue from Richards et al. The QSOs are within a narrow redshift band $0.3 < z_{\text{phot}} < 0.5$ and cross-matched with the 1.4 GHz National Radio Astronomy Observatory Very Large Array Sky Survey. The radio images extracted from the positions of the optical QSOs are retained if the flux integrated over the beam size of the radio survey is less than $3\sigma_{\text{rms}} \leq 1.35$ mJy. The radio-undetected QSO population is split into eight samples depending on their optical magnitudes and stacked to determine the mean
flux in each sample. The stacked mean flux is detected in all but the faintest optical magnitude sample. The radio versus optical luminosity relation from the stacked samples hint at a lower envelope in the radio luminosity which may be interpreted as there being a minimum radio jet power for a given accretion rate. Stacking assumes the underlying distribution of the property being measured is fairly represented by the stacked result. We investigate the underlying distribution of the radio flux density from the QSOs taking the noise of the sample into account. We find the distribution of the QSO flux density is modelled by a power-law with a negative index in all eight optical magnitude samples. This implies the mean stacked result is not a good representation of the distribution of the flux density of the QSOs and that there is no lower envelope. This highlights the danger of interpreting results from stacking without verifying the distribution is characterised by the mean stacked value. We use the distribution of the radio flux density of the QSOs to model the radio loudness. We appear to recover the quasar optical luminosity function when we model the distribution of radio loudness parameters suggesting that, since we are essentially holding the radio flux density fixed, the radio loudness is a function of the optical luminosity. This suggest that the radio loudness is not a fundamental property of the QSO but rather the ratio of two independent properties, the radio and optical luminosities. We convert the radio loudness parameter to jet efficiencies and find a minimum jet efficiency of $\eta_{\text{min}} = 4 \times 10^{-4}$. We find there is no sign of a minimum jet efficiency as far as our data’s sensitivity limit allows, so we expect $\eta < \eta_{\text{min}}$. Hence we provide an observational constraint for theoretical models of jet production in the minimum jet efficiency.
Whaia te iti kahurangi
ki te tuohu koe
me he maunga teitei

Dedicated to my family
in particular Hilda, Bill, Val and Peter
Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

The work in this thesis is my own unless otherwise stated. The reduction of CO(3-2) line used in Chapters 2 and 3 was reduced by Alejo Matrínez-Sansigre. The fitting of the far-infrared spectral energy distribution that is presented in Appendix A was done by Alejo Martínez-Sansigre and the Herschel data was provided by Mark Lacy.

Part of this thesis has been published in:

Acknowledgements

This thesis has been funded by SEPnet Astro and the Institute of Cosmology and Gravitation, University of Portsmouth.

Foremost, I would like to acknowledge the help, support and guidance of my primary supervisor Alejo Martínez-Sansigre. Your support throughout this PhD has made it possible, and in particular your ongoing support even after you left the field. I am extremely grateful that you continue to be my advisor from afar.

To Anna Kapinska thank you so much for your guidance and dedication over the past year and a bit. I hope I’ve not scared you off supervising more people. Your support in the second half of my thesis was excellent and invaluable.

To Bob Nichol, thank you for taking up the reigns almost a year ago even though you were already overloaded with work. Your experience and guidance was invaluable and helped this thesis along.

I would like to acknowledge Steve Rawlings. Although I only had his guidance for the first two years it was really inspirational, as was his encouragement at the beginning stages of the PhD process.

My office mates and friends I have made throughout the past three and a half years in Pompey have really enriched the whole experience. Particular thanks go to Heather Campbell, Dominic Galliano, Niko Meures, Guido Pettinari for making our year such a supportive and awesome group of people. Janine Pforr and Emma Beynon for being the best office mates and for their many words of advice and help.

Finally to my family, thank you for supporting me throughout this process. And most importantly Sam, who I married along the way, thank you so much for everything. I literally couldn’t have done it without you.
## Contents

1 Introduction 1

1.1 Active Galactic Nuclei ................................. 1

1.1.1 Eddington limit .................................. 3

1.1.2 Observed Emission from the AGN .................. 4

1.2 Radio emission from extragalactic sources ............... 9

1.2.1 Accretion paradigm of AGN ......................... 12

1.2.2 Spin paradigm of AGN .............................. 13

1.3 AGN and the connection with their host galaxies ....... 14

1.3.1 Co-evolution of the black hole and host galaxy ... 15

1.3.2 Dust in high redshift galaxies ....................... 17

1.3.3 Molecular gas in high redshift galaxies .......... 18

1.4 Outline ............................................. 20

I Molecular gas in AMS12 22

2 The molecular gas in a high redshift, obscured quasar 23

2.1 Introduction ......................................... 23

2.2 Molecular line observations ................................ 26

2.2.1 Radio observations ................................ 26

2.2.2 IRAM PdBI observations .......................... 28

2.3 Spectral line fitting .................................. 35

2.4 PdBI continuum measurements .......................... 39

2.5 The far-infrared dust .................................. 40

2.5.1 Dust mass ....................................... 41

2.6 Summary ............................................. 42
3 Modelling the molecular gas in AMS12

3.1 Motivation ........................................ 44
3.2 LVG modelling ..................................... 44
3.3 CO LVG model results .............................. 48
   3.3.1 No assumption on prior knowledge of the gas kinetic temperature 50
   3.3.2 Using the dust temperature PDF as a prior for $T_G$ ............... 55
3.4 Atomic carbon ................................... 58
   3.4.1 Excitation temperature and mass of atomic carbon ............... 59
   3.4.2 The [CI] temperature and mass towards AMS12 ................. 60
3.5 Discussion ......................................... 62
   3.5.1 Gas mass ..................................... 62
   3.5.2 Discriminating between LVG models ............................... 65
   3.5.3 Dynamical mass ................................ 70
   3.5.4 [CI] abundance and cooling contribution in AMS12 .............. 72
   3.5.5 AGN bolometric luminosity and the scale of AGN heating ....... 73
   3.5.6 Black hole, stellar, gas and dust masses ......................... 74
   3.5.7 Comparison to other galaxies ................................. 77
3.6 Conclusion and future work ........................... 78

II Radio-quiet quasars

4 Searching for a radio-quiet lower envelope

4.1 Introduction ....................................... 81
   4.1.1 Radio loudness parameter .......................... 82
   4.1.2 Stacking ...................................... 87
4.2 The SDSS and NVSS surveys ......................... 89
4.3 Radio emission from our sources ................... 89
   4.3.1 Multiple sources ................................ 90
   4.3.2 Missing lobes ................................ 90
   4.3.3 Radio emission from star formation ..................... 91
4.4 Stacking procedure ................................ 92
   4.4.1 SDSS QSO selection criteria ....................... 93
4.5 Matching with the NVSS catalogue ................. 94
4.6 Stacking ........................................... 96
4.6.1 Bootstrap resampling ........................................ 100
4.6.2 Splitting the optically brightest bin ....................... 103
4.7 Testing the behaviour of the noise ......................... 103
4.8 Searching for a radio-quiet quasar envelope ............... 109
4.9 Discussion .................................................. 112
4.10 Conclusion .................................................. 114

5 Beyond stacking: modelling the distributions of QSO flux densities 115

5.1 Motivation .................................................. 115
5.2 Modelling .................................................. 117
  5.2.1 Quasar flux density distributions ....................... 117
  5.2.2 Building a noise profile ................................ 122
5.3 The models of the underlying QSO flux density distributions .............................................. 124
  5.3.1 Power-law with noise .................................... 126
  5.3.2 Single Gaussian with noise ............................... 126
5.4 Parameter estimation and model comparison .................. 127
  5.4.1 Model Comparison ....................................... 128
5.5 Results .................................................... 129
  5.5.1 Parameter estimation results ............................. 129
  5.5.2 Noise as a model ....................................... 145
5.6 Results from the model comparison ......................... 148
  5.6.1 Goodness of fit tests .................................... 153
5.7 Investigating the effect of the model resolution ............. 154
5.8 Conclusions from the modelling ............................. 156

6 Radio loudness parameter: distribution and modelling 160

6.1 Motivation .................................................. 160
  6.1.1 Radio loudness parameter ................................. 163
6.2 Investigating the distribution of the radio-loudness parameter .............................................. 164
  6.2.1 The R parameter of the mean stacked radio flux densities .......................... 165
  6.2.2 The distribution of the R parameter given the underlying QSO flux density models .......... 166
  6.2.3 Modelling the distribution of $R$ ........................ 169
  6.2.4 Parameter estimation .................................... 171
  6.2.5 Model selection ....................................... 178
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.3 Discussion</td>
<td>181</td>
</tr>
<tr>
<td>6.3.1 Jet efficiencies</td>
<td>181</td>
</tr>
<tr>
<td>6.4 Conclusion</td>
<td>186</td>
</tr>
<tr>
<td>6.4.1 Future work</td>
<td>189</td>
</tr>
<tr>
<td>7 Conclusion</td>
<td>191</td>
</tr>
<tr>
<td>7.1 Molecular gas in an obscured quasar</td>
<td>191</td>
</tr>
<tr>
<td>7.2 Stacking of radio undetected QSOs</td>
<td>194</td>
</tr>
<tr>
<td>7.3 Future work</td>
<td>197</td>
</tr>
<tr>
<td>A The far infrared spectral energy distribution</td>
<td>200</td>
</tr>
<tr>
<td>A.1 Continuum measurements</td>
<td></td>
</tr>
<tr>
<td>A.1.1 Existing data: MAMBO and Spitzer</td>
<td>200</td>
</tr>
<tr>
<td>A.1.2 New data: Herschel and PdBI</td>
<td>201</td>
</tr>
<tr>
<td>A.2 Modelling the FIR emission</td>
<td>202</td>
</tr>
<tr>
<td>A.2.1 Far-infrared luminosity</td>
<td>203</td>
</tr>
<tr>
<td>References</td>
<td>208</td>
</tr>
</tbody>
</table>
# List of Figures

1.1 Schematic diagram of an AGN ........................................ 6
1.2 The composite optical quasar spectrum ................................ 7

2.1 Carbon monoxide and atomic carbon molecular lines towards AMS12 33
2.2 The average CO profile ............................................. 37

3.1 Cartoon showing the concept of LVG modelling ...................... 47
3.2 The PDFs of the LVG model fit to the CO ladder assuming flat priors 51
3.3 The best-fitting CO SED assuming flat priors ........................ 54
3.4 The PDFs of the LVG models fit to the CO data assuming dust temperature as prior ............................................. 56
3.5 The best-fitting CO SED assuming dust temperature as prior .... 58
3.6 Two component LVG models .......................................... 69

4.1 Mean stacked image of each of the six NVSS-undetected, optical magnitude samples ............................................. 99
4.2 Bootstrap resampling of the six samples’ integrated fluxes .... 101
4.3 The three sub-sample stacked images from splitting the optically brightest sample ............................................. 104
4.4 The bootstrapping results of the mean flux from the three sub-samples 105
4.5 The measured rms errors on the NVSS samples ...................... 107
4.6 Radio luminosity density vs. optical luminosity density ........ 111

5.1 The radio flux density distributions in each \( m_I \) sample .......... 119
5.2 The NVSS noise distribution ........................................... 123
5.3 Illustration of the construction of the models used in comparison with data ............................................. 125
5.4 The best-fitting power-law plus noise models ...................... 130
5.5 The power-law index PDFs ........................................... 133
5.6 The best-fitting Gaussian plus noise models in each \( \langle m_I \rangle \) sample ... 136
5.7 The best-fitting Gaussian model of the underlying QSO flux density in the \( \langle m_I \rangle = 17.88 \) sample ........................................... 138
5.8 Two dimensional PDFs of the Gaussian plus noise fits .................. 139
5.9 Marginalised PDF of the mean of the Gaussian model ................... 141
5.10 Marginalised PDF of the width of the Gaussian models ................ 143
5.11 The QSO and noise radio flux density distributions for each \( \langle m_I \rangle \) sample........... 146

6.1 The \( R \) values for the mean stacked radio luminosity densities versus B-band optical luminosity density ........................................... 167
6.2 The power-law model distributions of the underlying QSO radio flux density .................. 168
6.3 The power-law model indices versus the \( R \) parameter ................. 172
6.4 The PDF of model A .................................................. 173
6.5 The PDF of model B .................................................. 174
6.6 The marginalised PDFs of model C .................................. 175
6.7 The three models of radio loudness parameter distribution ............. 177
6.8 Model B: the distribution of \( R \) ..................................... 184
6.9 Model C: the distribution of \( R \) ..................................... 185

A.1 The PDFs of the dust properties for AMS12 ................................. 205
A.2 The fit to the FIR dust observations for AMS12 ............................ 207
## List of Tables

2.1 Observed line parameters towards AMS12 ........................................ 38

3.1 The calculated CO(1-0) line strength .............................................. 65

4.1 Properties of the six samples of the NVSS-undetected QSOs sorted in optical magnitude .......................................................... 97

4.2 The flux measurements of the six stacked images .................................. 98

4.3 The flux measurements and the errors from the bootstrap resampling ...... 102

4.4 The brightest optical luminosity bin split into three bins ....................... 106

4.5 The stacked mean radio and optical luminosity densities ....................... 110

5.1 Best-fitting power-law indices ...................................................... 132

5.2 The best-fitting Gaussian parameters ............................................... 135

5.3 Results from the Kolmogorov-Smirnov test ..................................... 145

5.4 The odds ratios of the Gaussian plus noise and noise alone models ........ 150

5.5 The odds ratios of the power-law plus noise and the noise alone models .... 151

5.6 The odds ratio of the power-law plus noise and the Gaussian plus noise models ................................................................. 152

5.7 The $\chi^2$ goodness-of-fit test ...................................................... 154

5.8 Varying the resolution of the power-law model .................................. 155

6.1 The stacked mean radio and optical luminosity densities ....................... 166

6.2 The odds ratio comparing the models for the distribution of the radio loudness ................................................................. 178

6.3 The bolometric luminosities for the eight optical magnitude samples ...... 183

A.1 The FIR measurements ............................................................. 202
Glossary of abbreviations

2QZ 2 degree field quasar redshift survey
AGN Active Galactic Nuclei
ASKAP Australian SKA Pathfinder
BLR Broad Line Region
[CI ] Atomic Carbon
CASA Common Astronomical Software Applications
CMB Cosmic Microwave Background
CO Carbon monoxide
DR# Data Release #
EMU Evolutionary Map of the Universe
FIR Far-Infrared
FIRST Faint Images of the Radio Sky at Twenty centimetres survey
FWHM Full-Width Half-Maximum
H$_2$ Molecular Hydrogen
ICM Intercluster Medium
IDL Interactive Data Language
IGM Intergalactic Medium
IMF Initial Mass Function
IR Infrared
IRAM Institut de Radioastronomie Millimétrique
ΛCDM A Cold Dark Matter
LF Luminosity Function
LOFAR Low Frequency Array
LTE Local Thermodynamic Equilibrium
LVG Large Velocity Gradient
MAMBO Max-Planck Millimetre Bolometer Array
MHD Magnetohydrodynamic
NLR Narrow Line Region
NRAO National Radio Astronomy Observatory
NVSS NRAO VLA Sky Survey
OLF Optical Luminosity Function
PdBI Plateau de Bure Interferometer
PDF Probability Density/Distribution Function
QSG Quasi-Stellar Galaxy
QSO Quasi-Stellar Object
QSRS Quasi-Stellar Radio Source
RLF Radio Luminosity Function
RLQ Radio-Loud Quasar
RQQ Radio-Quiet Quasar
SDSS Sloan Digital Sky Survey
SED Spectral Energy Distribution
SF Star Formation
SFR Star Formation Rate
SKA Square Kilometre Array
SMBH Supermassive Black Hole
SMG Submillimetre Galaxy
SNR Signal to Noise Ratio
ULIRG  Ultra-Luminous Infrared Galaxy
UV   Ultraviolet
VLA  Very Large Array
Chapter 1

Introduction

Quasars are the most powerful objects in the observed Universe. They are an important ingredient in galaxies and their role in galaxy evolution is a widely investigated topic with correlations observed between the quasars and their hosts. The aims of this thesis follow two tracks: we investigate the effect a powerful quasar may have on the molecular gas from which stars are formed via observations towards a high-redshift, obscured quasar; then we probe a volume-limited sample of optically selected, radio-undetected quasars in order to investigate the radio flux density distribution of this sample. The aim of the latter is to attempt to constrain whether radio and optical emission are related and how this sample can be used to constrain models of relativistic jet production by supermassive black holes.

1.1 Active Galactic Nuclei

Quasars are the brightest examples of a class of objects called Active Galactic Nuclei (AGN). AGN emit radiation from the central regions of galaxies that cannot be attributed to stars. They are often split into two classes, those that outshine their host galaxies (quasars) and those that do not (Seyferts). AGN were first discovered
Active Galactic Nuclei

by the optical spectrum of NGC 1068 taken by E. A. Fath at Lick Observatory in 1908. The central region of NGC 1068 showed strong emission lines, which with higher resolution observations by V. M. Slipher at Lowell Observatory, were shown to have widths of several hundred km s$^{-1}$ (see e.g. [Peterson 1997]). Further spectra taken of nearby galaxies found similar characteristics between them. These galaxies all shared the property of high surface brightness at their centres, i.e. their cores appeared star-like, while not outshining their host galaxies; they were classified as Seyfert galaxies ([Seyfert 1943]).

In the second half of the twentieth century the emergence of radio observations and surveys revealed strong radio sources. These often had corresponding optical counterparts, that were very bright point sources. The star-like quality of the appearance of their optical emission gave rise to their name quasi-stellar radio sources (QSRS) or quasars. On analysis of the optical spectrum of 3C273 it was discovered the strong emission lines were from the hydrogen-Balmer series and this source was at the highest redshift yet observed ($z = 0.158$, [Schmidt 1963]). A collection of these objects was compiled by [Schmidt 1968] which shared the same properties: star-like sources with radio counterparts, broad emission lines, large redshifts, variable continuum emission and large UV fluxes. Following the discovery of quasars, it was noticed that there were some point-like objects that did not have strong radio counterparts. These were named quasi-stellar galaxies (QSGs) since their optical emission was similar to the quasars but they were not detected in the radio ([Sandage 1965]). The term quasi-stellar object (QSO) was coined to encompass both quasars and QSGs ([Hoyle & Burbidge 1966]). In this thesis we will use the term quasars in general unless we are specifically addressing radio-quiet, optically selected quasars which we refer to as QSOs.

The emission from quasars is seen across the spectrum with some of this emission exhibiting variability on very short timescales (e.g. the UV or X-rays) suggesting a
Active Galactic Nuclei

1.1

Small region. The variability in optical is on longer timescales, and in the infrared, variability is seen on even longer timescales, suggesting the production of this emission is at a greater distance from the centre than the short wavelength emission (Krolik, 1999). The emission from these objects is observed over a wide range of frequencies and coupled with the fact they show different variabilities, suggest a range of physical mechanisms are acting to produce this emission, and they occur at vastly different scales (e.g. Krolik, 1999). The strong emission lines from the optical spectroscopy of these objects are very broad, high-excitation lines, suggesting rapid velocities. Assuming the clouds which these emission lines come from are gravitationally bound to the centre of the galaxy, the mass of the nucleus must be very large, since from virial arguments \( M \approx \frac{v^2 r}{G} \). Therefore a central, compact source of a large mass was required to explain these observations, e.g. a black hole (e.g. Burbidge & Perry, 1976; Sargent et al., 1978; Soltan, 1982; Rees, 1984).

1.1.1 Eddington limit

The mass of the black hole can also be determined from Eddington limiting luminosity arguments. This relates the mass and the luminosity of the AGN. In the case of the gas around the black hole being spherically symmetric, and assuming it consists of fully ionized hydrogen, the flux \( S \) at distance \( r \) is given by,

\[
S = \frac{L}{4\pi r^2},
\]

(1.1)

where \( L \) is the luminosity. The radiation pressure is given by \( P_{\text{rad}} \propto \frac{L}{c} \) and since the flux is the flow of energy \( E \) we can express \( P_{\text{rad}} \) as

\[
P_{\text{rad}} = \frac{L}{4\pi r^2 c}.
\]

(1.2)
1.1 Active Galactic Nuclei

This is the radiation pressure at a distance $r$. The outward radiation force on an individual electron is given by,

$$F_{\text{rad}} = \frac{L\sigma_T}{4\pi r^2 c},$$

(1.3)

where $\sigma_T$ is the Thomson cross-section. The gravitational force is given by,

$$F_{\text{grav}} = \frac{GMm_p}{r^2},$$

(1.4)

where $G$ is the gravitational constant, $M$ is the mass of the central object, and $m_p$ is the mass of a proton. The Eddington limit is where the radiation force balances the gravitational force ($F_{\text{rad}} = F_{\text{grav}}$), and thus we can express the Eddington luminosity as,

$$L_{\text{Edd}} = \frac{4\pi GcMm_p}{\sigma_T}.$$  

(1.5)

This is the maximum limit on the luminosity of a system with mass $M$ powered by accretion. Inverting this argument, we can estimate the mass of a source with an observed luminosity. This would be the minimum mass of an accreting system that is radiating at the Eddington limit. For example a source with a typical observed luminosity for a quasar of $L = 10^{39}$ W, would have a mass of at least $M = 10^8$ M$_\odot$. This is assuming it is not radiating or it cannot radiate with a luminosity significantly above the Eddington luminosity.

1.1.2 Observed Emission from the AGN

For a detailed account of the physical processes responsible for the emission observed from an AGN system I refer the reader to [Krolik (1999)]. The following
paragraphs briefly outline the emission observed and where the emission originates using Figure 1.1 as a visual reference.

The structure of the quasar-like AGN system represented in Figure 1.1 (taken from Urry & Padovani, 1995) illustrates the unification scheme of AGN. At the centre is the black hole itself. The black hole is surrounded by an accretion disk which is formed by diffuse material that is orbiting the accreting black hole and distributed into a disk from the mixing of the angular momentum within the gas. It is heated to temperatures of $\sim 10^3 - 10^6$ K emitting thermal continuum radiation in optical and ultraviolet (UV) wavelengths. X-rays and $\gamma$-rays are produced near the centre by inverse-Compton and synchrotron processes. Nearby gas clouds, photoionised by absorption of the UV/optical continuum radiation, emit line radiation in the X-ray, UV, and optical. The gas clouds that are close to the black hole show very broad velocity dispersion, with lines widths of 1000 - 2000 km s$^{-1}$. Collectively, the gas clouds which emit these lines are from the broad line region (BLR) shown in Figure 1.1. Ionised gas which is further away from the supermassive black hole (SMBH), and hence has a narrower velocity dispersion, emits narrow lines including forbidden lines (narrow line region, NLR). The composite optical quasar spectrum from Vanden Berk et al. (2001) is shown in Figure 1.2 to give a detailed impression of the typical optical emission from quasars. The spectrum peaks at blue wavelengths and shows broad and narrow emission lines throughout the spectral range. Note this is the optical spectrum and as such it is accumulated from unobscured quasar emission.

The obscuring torus is shown further out from the nucleus as the orange bulge in Figure 1.1. It is believed to be made up of clumps of dust and absorbs the UV/optical continuum from the accretion disk. It is heated to $\sim 2000$ K and is thought to be responsible for the observed infrared (IR) continuum (e.g. Krolik & Begelman, 1988; Nenkova, Ivezić & Elitzur, 2002; Tristram et al., 2007; Nenkova
Figure 1.1: Schematic diagram of a “unified” radio loud AGN from [Urry & Padovani, 1995]. In this figure (which is not to scale) the black hole and accretion disk are shown at the centre. The green arrows indicate approximately from which direction the emission is observed in the case of different classifications of AGN. The Seyfert galaxies are split into type-1 and type-2 based on their optical spectroscopy; type-1 showing permitted, broad lines superposed with forbidden, narrow lines, while a type-2’s optical spectrum shows only the forbidden, narrow emission lines. The narrow line radio galaxies and the broad line radio galaxies similarly are classified by the appearance of emission lines in their optical spectra as well as the presence of strong radio emission. Blazars are highly variable AGN; they are oriented so they are observed face-on, looking ‘down’ the jet. Radio loud quasars are observed at an angle close to the angle of the jets (not as close as a blazar), the radio emission is dominated by the core emission which is beamed, and the optical spectroscopy reveals broad lines also. Radio quiet quasars appear as radio-loud quasars in the optical, but are not detected in the radio. Image credit: Urry & Padovani (1995), acquired from http://www.auger.org/news/PRagn/images/agn4_prouza.png
Figure 1.2: The composite optical quasar spectrum from \(\sim 2200\) quasars in the Sloan Digital Sky Survey (SDSS) \cite{Vanden Berk et al. 2001}. This is the geometric mean composite spectrum (Figure 5 in \cite{Vanden Berk et al. 2001}). Note that the spectral indices in this plot follow a different convention than we use in this thesis which is given by \(s_\lambda \propto \lambda^{-(\alpha_\nu+2)}\) with \(s_\lambda\) as the flux density at wavelength \(\lambda\).
Figure 1.1 shows an outflow from the vicinity of the black hole to represent the radio jets seen in some quasars. The mechanism that produces these jets is still not fully established. It is thought that magnetic fields tied to the accretion disk or the surface of the black hole are able to transport plasma to great distances (e.g. see Frank, King & Raine 2002 for a detailed description of accretion disks). One theory is that the jets are a mechanism for shedding angular momentum from the accretion disk (i.e. through poloidal magnetic fields anchored on the surface of the accretion disk; see e.g. Blandford & Payne 1982; Lovelace 1976). In other theories, the power of the jets are related to the spin of the black hole (e.g. via the Blandford-Znajek mechanism Blandford & Znajek 1977; Wilson & Colbert 1995, among others).

The unified scheme of AGN (e.g. Rees 1984; Antonucci 1993; Urry & Padovani 1995), states that the orientation of the quasar with the observer’s viewing angle is responsible for the obscuration. The obscured and unobscured quasars are intrinsically identical. The dusty torus surrounding the central black hole is impeding the observer’s view of the broad lines and the optical and UV continuum. However, obscuration via dust distributed in the host galaxy is also possible.

Host obscuration of quasars was suggested by Sanders et al. (1988) and Fabian (1999) proposed a model where the phase of the growth of massive black hole happens during an epoch where the host galaxy itself is undergoing a phase of growth, and hence contains more dust and gas. The model that Fabian (1999) suggests is that both the obscured quasar and its host galaxy are at different evolutionary phases, and hence different to the unobscured quasars.

Recent studies of obscured quasars have revealed quasars in which the narrow line region is also obscured resulting in observed featureless optical spectra. Moreover, a few sources also exhibit flat radio spectra indicating the jets emanating from the obscured source are face-on to the observer; this discovery discounts the premise
that the dusty torus is solely responsible for the obscuration, \cite{Martinez-Sansigre2005, Martinez-Sansigre2006b, Rigby2006, Kloeckner2009}. These cases suggest that dust on kpc scales is the source of at least some of the obscuration.

\section*{1.2 Radio emission from extragalactic sources}

Radio emission seen from extragalactic sources is often dominated by the emission from non-thermal processes rather than from thermal processes. The radio emission from thermal processes such as blackbody and thermal Bremsstrahlung emission is dwarfed by the emission from non-thermal processes at low frequencies (i.e. below 30 GHz, see Figure 1 of \cite{Condon1992}). The dominant non-thermal process is synchrotron radiation from free electrons in a plasma travelling at close to the speed of light, interacting with a magnetic field. This synchrotron emission is responsible for the bulk of the radio luminosity from quasars as well as from normal galaxies, where in the latter it is produced by relativistic electrons, accelerated in magnetic fields, in supernova remnants (e.g. \cite{Condon1992}). The spectrum from synchrotron emission can be approximated by the power law,

\[ S(\nu) \propto \nu^{-\alpha}, \]

where \( S(\nu) \) is the flux density at a frequency \( \nu \) and \( \alpha \) is the spectral index.

Radio surveys uncovered many sources with radio emission showing a variety of morphologies and could be further classified by their optical counterpart appearance. The features of the radio emission from extragalactic sources (e.g. AGN) are commonly identified in four categories: the core, jet, lobe and hotspot emission (e.g. \cite{Begelman1984, Bridle1984}). The core emission is compact and often has a flat spectrum (\( \alpha \sim 0 \)) which is believed to be the su-
perposition of optically thick synchrotron emission regions. The jets and lobes are extended, optically thin, radio emission. The jets are collimated outflows extending from the nucleus or the radio source. Bridle & Perley (1984) defined the jet to be any region of radio emission that is four times as long as it is wide, is a separable feature of the radio emission and is pointed towards the nucleus of the host. Jets that are observed at a large viewing angle generally have steep spectra (i.e. $\alpha \sim 0.7$). The radio lobes can be very large, sometimes spanning megaparsecs. The hotspots are sometimes seen at the ends of extended lobes and are assumed to be the termination of the powerful collimated flows which meet the intergalactic medium (IGM) and is drastically decelerated.

Radio-loud AGN can be further subdivided into two classes by their radio emission. They are sorted into either Fanaroff-Riley class Is (FRIs) which are dominated by cores and jets, or Fanaroff-Riley class IIs (FRIIs) where the brightest regions are further away from the central source (Fanaroff & Riley, 1974). The classification by Fanaroff & Riley (1974) of these objects took the ratio of the distance between the regions of brightest on either side of the galaxy (or quasar) to the total distance extended by the radio emission. The cut-off between the two classes was a ratio of 0.5, with FRIs classified as those objects with ratios $< 0.5$ and FRIIs $> 0.5$. This corresponded also to a division in the radio luminosities at 178 MHz with nearly all the FRIs having luminosities below $L_{178\, \text{MHz}} \sim 10^{25} \, \text{W Hz}^{-1} \, \text{sr}^{-1}$, and mostly all FRIIs having luminosities greater than this. Below we briefly outline several extragalactic radio sources that are classified by their appearance in optical and radio wavelengths.

**Radio galaxies**

Radio galaxies (RGs) are strong radio sources showing extended radio emission in the form of jets and/or lobes. RGs are selected at radio wavelengths. The optical
photometry of radio galaxies revealed that they are typically elliptical galaxies and, from their optical spectroscopy show forbidden, narrow emission lines similar to the type-2 Seyferts (narrow line radio galaxies, NLRG), and sometimes also permitted, broad emission lines (i.e. type-1, broad line radio galaxies, BLRG). Figure 1.1 shows RGs as interpreted by the unified model of AGN. It is thought that the central engine in RGs is obscured and they are probably radio-loud obscured quasars (e.g. Urry & Padovani, 1995; Krolik, 1999).

Quasars

We introduced quasars in Section 1.1 and referred to the part the radio emission played in the discovery of these objects. Quasars are split into radio loud and radio quiet objects (RLQ and RQQ respectively) based on their radio emission. The radio loudness can be parameterised by the radio loudness parameter $\mathcal{R}$ which is the ratio of the luminosity density in radio and optical wavelengths/frequencies. Strittmatter et al. (1980) noticed the distribution of the radio-to-optical flux densities for optically selected QSOs appeared bimodal. Kellermann et al. (1989) defined the radio loudness by the 5 GHz radio luminosity density and B-band optical luminosity as,

$$\mathcal{R} = \frac{L_{5 \text{ GHz}}}{L_B}.$$  \hspace{1cm} (1.7)

Kellermann et al. (1989) considered a quasar to be radio loud when $\mathcal{R} > 10$. The distribution of the radio loudness of quasars is a matter under investigation with the prospect of determining whether RLQs and RQQs form two distinct populations (manifesting in a bimodal distribution) or rather their distributions being more continuous (e.g. Kellermann et al. 1989; Ivezić et al. 2002; Cirasuolo et al. 2003b; Ivezić et al. 2004; Stawarz, Sikora & Lasota 2008; Baloković et al. 2012). Constraining the shape of the distribution of the radio loudness is complicated by
selection effects such as incompleteness in surveys of optically selected QSOs and that \(\sim 90\%\) of QSOs are radio-quiet (e.g. Kellermann et al., 1989). In Chapter 6 we will investigate the radio-quiet end of the radio loudness using optically selected QSOs that are not detected in a 1.4 GHz radio survey.

The mechanism for producing jets of collimated plasma must account for observations of varying radio luminosities from quasars with the same optical luminosities (the latter are used as proxies for the bolometric luminosities that are in turn proxies for the accretion rate). The radio jets are thought to provide an efficient angular momentum transport mechanism in the accretion disk (e.g. Blandford & Payne, 1982), via magnetic field lines that are anchored onto the surface of either the disk or the event horizon. These magnetic field lines are further twisted with the angular momentum of the disc or the black hole where extra energy may be extracted via, for example, the Blandford-Znajek mechanism (Blandford & Znajek, 1977).

### 1.2.1 Accretion paradigm of AGN

The accretion paradigm of AGN states that most, if not all AGN, are driven by accretion onto a supermassive black hole. Here the accretion rate \(\dot{m}\) plays the most defining role in determining the emission properties, and hence the appearance of the central source. In sources where the accretion rate is high \((\dot{m}/\dot{m}_{\text{Edd}} \gtrsim 0.1\), where \(\dot{m}_{\text{Edd}} = 4\pi G M_\ast / \epsilon \kappa_{\text{es}} c\) is the Eddington accretion rate with radiative efficiency \(\epsilon\) and electron scattering opacity \(\kappa_{\text{es}}\), the AGN appears as an optical quasar (also bright in X-ray). Whereas in cases where the accretion is of a sub-Eddington rate, a weak radio core and less optical emission are observed.

Not all AGN follow the conventional picture portrayed in Figure 1.1 where there is strong broadband continuum radiation and high-excitation emission lines. In objects with significant radio emission - radio galaxies, the conventional AGN are
known as high-excitation radio galaxies (HERGs). Objects that do not show strong emission lines are known as low-excitation radio galaxies (LERGs) (e.g. Hine & Longair 1979; Laing et al. 1994). Their nuclear emission is consistent with coming from radio jets, i.e. their energy is mostly from kinetic energy from the jets (Hardcastle, Evans & Croston 2006; Best & Heckman 2012). They are believed to be fuelled by inefficient accretion or hot accretion often termed ‘radio-mode’, while HERGS are thought to be fuelled by efficient or cold accretion - ‘quasar-mode’ (see for example: Best & Heckman 2012). These LERGs have been shown to be hosted by different host galaxies than the quasar-mode AGN (Best et al. 2005).

While the accretion rate determines the appearance of the quasar and accounts for several basic optical properties of the AGN, the variation in jet power presents a riddle. For example, two identical supermassive black holes (SMBHs, identical in mass, accretion rate and X-ray to infrared SEDs), can show radio luminosities varying by several orders of magnitude. Hence, there must exist some intrinsic property of the system which accounts for the jet power. A strong candidate is the angular momentum of the SMBH, the spin, that is most likely the variable responsible for the differing jet powers (e.g. Wilson & Colbert 1995; Meier, Koide & Uchida 2001).

1.2.2 Spin paradigm of AGN

The spin paradigm, for example as outlined in Meier (2002), states that to the first order, it is the normalised black hole angular momentum, \( \hat{a} \), that determines whether or not a strong radio jet is produced. Blandford & Znajek (1977) and Punsly & Coroniti (1990) suggest that jet power increases as the square of black hole angular momentum. However, recent work has indicated that at very high spins (\( \hat{a} > 0.9 \)) the jet power scales at higher orders of the angular momentum (for a detailed discussion we refer the reader to the Appendix in Tchekhovskoy, Narayan
& McKinney, 2010, and references within). The magnetic field lines are twisted as the black hole spins, and from the effects of differential rotation the material (free electrons from the plasma) is expelled along the field lines.

The synchrotron radiation is produced via relativistic electrons that spiral around these magnetic field lines, changing direction constantly and emitting radio light which is beamed in the direction of travel. The plasma experiences pinching from the frame dragging of the magnetic field lines introducing hoop stresses which, since the plasma cannot cross the magnetic lines and the frame dragging of the lines create a tighter coil, accelerate the plasma along the axial direction outwards; the plasma becomes collimated due to the pinch effect. For detailed discussions on the jet production mechanisms (which are beyond the scope of this thesis) see e.g. Meier, Koide & Uchida (2001), Meier (2002), Hawley & Krolik (2006), Beckwith, Hawley & Krolik (2008) and Tchekhovskoy, Narayan & McKinney (2010).

### 1.3 AGN and the connection with their host galaxies

AGN feedback is the interaction between the products of AGN activity such as the radiation, wind, and jets, and the host galaxy. Feedback is thought to quench star formation (and affect the growth of the black hole) by heating and ejecting interstellar dust and gas.

Fabian (2012) has produced a recent review on the observational evidence for AGN feedback and identified two modes of interaction, radiative or wind and kinetic. The radiative mode of feedback is due to the radiation pressure exerted on nearby dust particles embedded in the gas of the interstellar medium (ISM) (e.g. Laor & Draine, 1993; Scoville & Norman, 1995; Murray, Quataert & Thompson, 2005).
Quasar winds is another mechanism which expels the ISM quenching star formation.

The kinetic mode of feedback refers to the radio jets emitted by radio galaxies and their interaction with the intercluster medium (or intergalactic medium, IGM). In the IGM, the initial gravitational collapse which forms galaxies heats the IGM to hot temperatures. The hot gas cools via bremsstrahlung emission in the X-ray, the cooling is strongest in the densest regions. The hot gas surrounding these dense, cool regions flows towards the centre due to its pressure, which is known as the cooling flow (see [Fabian](#) [1994] for a review). However, it has been observed that the mass of this cooled gas is much less than expected and so there appears to be an external source heating the intergalactic medium which is termed the ‘cooling flow problem’ (e.g. [Heckman et al.](#) [1989] [O'Dea et al.](#) [1994] [Peterson et al.](#) [2003] [Peterson & Fabian](#) [2006]). AGN feedback is believed to be the primary heating mechanism though others such as supernova and cosmic ray heating and thermal conduction are postulated also.

Further observational evidence for the radio jet interaction with the ambient medium can come in the form of bubbles or cavities (e.g. [McNamara & Nulsen](#) [2007] for a review and references). It has been observed that the cavities in X-ray coincide with radio lobes, and the power of the radio jets can be inferred from the work done by the lobes to expand into the cavities.

For a recent comprehensive review on the topic see [Fabian](#) [2012] and references within. The following sections focus on the molecular gas and interstellar dust which will be analysed in detail in an obscured, high redshift quasar in Chapters [2](#) and [3](#).

### 1.3.1 Co-evolution of the black hole and host galaxy

The connection between an AGN and its host galaxy is an area of research with implications in galaxy evolution and the AGN and galaxies coeval evolution.
Magorrian et al. (1998) found a correlation between $M_\bullet$ and the bulge mass $M_{\text{bulge}}$ indicating there is a connection in how the galaxy and the black hole grow. Silk & Rees (1998) proposed that the formation of massive black holes may have peaked at an earlier redshift than the peak of galaxy formation ($z \sim 2 - 3$ as opposed to $z \sim 1 - 2$ for galaxies), implying that protogalactic star formation (SF) may have been influenced by the nucleus’ activity. Furthermore, Ferrarese & Merritt (2000) found an even tighter correlation between the black hole mass $M_\bullet$ and the stellar velocity dispersion in the bulge $\sigma_{\text{bulge}}$. The $M_\bullet-M_{\text{bulge}}$ and $M_\bullet-\sigma_{\text{bulge}}$ correlations are expected to be due to AGN feedback. AGN feedback is believed to be a major suppressant of star formation, and so is able to explain this correlation (e.g. Di Matteo, Springel & Hernquist 2005; Bower et al., 2006; Croton et al., 2006).

Forms of AGN feedback in close proximity to the SMBH include: quasar outflows of ionizing radiation, which would suppress star formation by UV flux destroying H$_2$ molecules, and from powerful winds generated when a SMBH is accreting material at a sufficient rate. Such winds radiate outwards and, if powerful enough, will push out the surrounding gas. The wind luminosity is a fraction of the Eddington luminosity (see Silk & Rees, 1998; Fabian, 1999 for details), hence a more massive AGN ($M_\bullet \gtrsim 10^7 M_\odot$) could in theory eject all the material from their host galaxies (since $L_{\text{Edd}} = 4\pi GcM_\bullet/\kappa_{\text{es}}$), thereby shutting down the star formation and the black hole feeding mechanism. The winds generated by the SMBH suppress star formation in its host galaxy, and this feedback effect on star formation is an important consideration in star formation models (e.g. Springel, Di Matteo & Hernquist 2005; Bower et al., 2006; Croton et al., 2006; Lagos, Cora & Padilla, 2008; Booth & Schaye, 2009; McCarthy et al., 2010).

On larger scales AGN feedback from the radio jets can influence the wider surroundings. Although highly collimated and directional, the powerful jets would heat surrounding material within the host galaxy out to a large distance from the centre,
thus suppressing star formation at a distance from the SMBH. Where jets extend out into the IGM, the shocks created by the jets heat the surrounding IGM and ICM, thereby preventing cooling and collapse of the matter responsible for further star formation into the galaxy (combined with other heating processes e.g. Fabian, 1994; Binney & Tabor, 1995; McNamara & Nulsen, 2007; Fabian, 2012 and references within).

### 1.3.2 Dust in high redshift galaxies

A source of extinction in the optical and UV wavelengths is interstellar dust. Dust absorbs the emission of photons at these wavelengths and re-emits thermal radiation in the infrared (re-processing). Dust is important for star formation; it promotes the production of molecular hydrogen which forms on the surfaces of the dust grains, and the absorption of the UV photons which causes the dissociations of molecules. The dust grain surfaces help with the production of other molecules as well. The source of heating the dust in galaxies is nearby star formation; the latter can be estimated from the far-infrared (FIR) luminosity of the dust (e.g. Kennicutt, 1998). However, there is some evidence that the dust may be heated by a host’s AGN as well as SF (e.g. Barvainis, 1987; Sanders et al., 1989; Siebenmorgen et al., 2004). The strong association of the dust with the molecular gas means that characterising the dust properties can help to further our understanding on the gas properties and increase our knowledge on the star formation in the galaxies. See Desert, Boulanger & Puget (1990) and Osterbrock & Ferland (2006) for good reviews on the interstellar dust.
1.3.3 Molecular gas in high redshift galaxies

Molecular gas is an important constituent in galaxies. It is, in general, quite cold (10-50 K), which is a requirement for gravitational collapse, though it can be heated up to $\approx 100$ K through violent starbursts or AGN heating. It is an important probe for understanding dynamics and structures of galaxies, as well as providing information on the properties of star formation.

Molecular gas clouds consist almost entirely of molecular hydrogen, $\text{H}_2$, which requires cool temperatures in order to form on the surface of dust grains, and photon absorption causes dissociation of the molecule from the grain. $\text{H}_2$ is symmetric, and lacks a permanent dipole, leading to no rotational transitions. The vibrational states of $\text{H}_2$ may be excited by high temperatures. However, the emission from these occurs in the mid-IR, meaning direct observations of $\text{H}_2$ are very difficult (i.e. from ground-based telescopes). Fortunately the second most abundant molecule (and the brightest), carbon monoxide (CO), has a permanent electric dipole and emission from its rotational transitions can be observed in the millimeter wavelengths. Therefore it is used as a tracer for $\text{H}_2$ mass (e.g. Osterbrock & Ferland 2006).

The conditions within the molecular cloud, which leads to the excitation of molecules, are shaped by a balance between the ambient temperature and the density of the gas. A thorough representation of the gas requires solving the coupled equations of radiative transfer and statistical equilibrium at each point. The high optical depth of the CO transitions make the determination of the gas properties difficult to infer, therefore detailed models are often required (e.g. Goldreich & Kwan 1974, Scoville & Solomon 1974).

$^{12}\text{CO}$ (hereafter CO), is the best tracer for $\text{H}_2$ mass due to its stability and abundance, and its ability to be excited and thermalised by collisions with $\text{H}_2$ at low densities (e.g. Solomon & Vanden Bout 2005). In order to be detected, the
intensity from the transitions needs to be in excess of the background radiation. The non-zero permanent electric dipole of CO means radiation is emitted from rotational transitions within the molecule. The rotational angular momentum quantum level, \( J = 0, 1, 2, 3, \ldots \), has energy \( E_J = \hbar \nu_{\text{CO}} J(J + 1)/2 \), where \( \nu_{\text{CO}} = 115 \text{ GHz} \), is the rest frequency of the transition CO(1-0). Since the rotational levels are discrete, each transition emits a line of frequency

\[
\nu_{\text{line}} = \frac{\hbar J}{2\pi I}
\]  

(1.8)

where \( I \) is the moment of inertia of a molecule. Thus the rotational spectrum of CO appears as a ladder with steps that are harmonics of the fundamental frequency (e.g. Goldreich & Kwan 1974; Obreschkow et al. 2009, whose Appendix gives a clear presentation).

The minimum excitation temperature \( T_{\text{ex}} \), needed for significant collisional excitation for each rotational level, is given by the rotational energy via

\[
T_{\text{ex,min}} \approx \frac{E_{\text{rot}}}{k} \approx \frac{\nu_{\text{line}} \hbar (J + 1)}{2k},
\]  

(1.9)

where \( k \) is the Boltzmann constant.

The critical density of the cloud \( n(\text{H}_2) \), required for substantial excitation depends upon the temperature of the cloud. The critical density occurs where the spontaneous and the collisional radiative rates are equal. It is therefore given by the ratio of the Einstein coefficient of spontaneous emission from an upper level to a lower level, and the collisional coefficient of the transition \( A_{ul}/(n(\text{H}_2)\gamma_{ul}) \). The typical critical density of the CO(1-0) transitions is \( \sim 10^3 \text{ cm}^{-3} \) (e.g. Genzel 1992).

Mechanisms responsible for heating the cloud include radiation from nearby star formation and in some cases AGN producing ionizing radiation, and shocks from
the quasar-driven winds described above (Silk & Rees, 1998). The excitation of CO to higher rotational transitions occurs in an optically thick environment. Therefore, the physical interpretation of these lines is not straightforward and requires detailed modelling. To circumvent this need, we can use observations of optically thin lines. Atomic carbon ([CI]) lines are one such probe due to the \( ^3P \) fine-structure arrangement of carbon which forms a simple three-level system. The lower and upper fine structure transitions of atomic carbon [CI](\( ^3P_2 - ^3P_1 \)) and [CI](\( ^3P_1 - ^3P_0 \)), (at rest-frame frequencies 492 GHz and 809GHz respectively), can be used to directly and independently determine the excitation temperature, neutral carbon column density and mass (e.g. Weiß et al., 2003). These quantities in turn, can be used to determine the properties of the molecular gas using conversion relations with CO and \( \text{H}_2 \). CO and [CI] appear to be intrinsically linked from spatial distribution studies in Orion by Ikeda et al. (2002), who found [CI] to trace the regions of CO emission. This connection to CO is further supported by the critical density for both the [CI](\( ^3P_1 - ^3P_0 \)) and the CO(1-0) lines being similar \( n_{\text{crit}} \approx 10^3 \text{ cm}^{-3} \) suggesting the transitions arise from the same volume and the excitation temperatures are thought to be equal (e.g. Ikeda et al., 2002).

1.4 Outline

This thesis is composed of two parts; the first part is an investigation into the molecular gas towards an obscured, \( z = 2.8 \) quasar, while the second part deals with studying a population of radio-undetected QSOs at a specific redshift (\( z \sim 0.4 \)). The two individual parts have more detailed introductions associated with them; however I will briefly outline their aims here.

The study into the molecular gas towards AMS12, an obscured quasar, is done with the detection of several carbon monoxide (CO) rotational transition lines and
both [CI] fine structure lines. Chapter 2 introduces the project and provides details on the observations of AMS12 from the millimetre interferometer Plateau de Bure Interferometer (PdBI). The CO SED is modelled with detailed large velocity gradient (LVG) models to infer the temperature, density and size of the emitting region in Chapter 3. The properties of the gas are compared with the dust properties in this object which were modelled by A. Martínez-Sansigre (details are given in Appendix A). This work appears in Schumacher et al. (2012).

In part two, we investigate the radio-undetected, volume-limited sample of optically selected QSOs. Chapter 4 motivates the study into these objects and outlines the stacking technique, which is used in many wavelengths to obtain detections of faint objects. In Chapter 5 we look beyond the stacked results into the distribution of the radio flux densities from each individual QSO. We fit three models to the distribution to try and characterise the underlying QSO emitted flux density. Then we investigate the implications of our results in the context of jet efficiencies and compare them to recent simulations of the jet efficiencies in Chapter 6.

Throughout this thesis we assume a \( \Lambda \) Cold Dark Matter (\( \Lambda \)CDM) cosmology with \( H_0 = 70 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1} \), \( \Omega_\Lambda = 0.70 \) and \( \Omega_m = 0.30 \).
Part I

Molecular gas in AMS12
Chapter 2

The molecular gas in a high redshift, obscured quasar

In this chapter we present new detections of the CO(5-4), CO(7-6), [CI](^3P_1 - ^3P_0) and [CI](^3P_2 - ^3P_1) molecular and atomic line transitions towards the unlensed, obscured quasar AMS12 (z = 2.7672), observed with IRAM’s Plateau de Bure Interferometer (PdBI). This is the first unlensed, high redshift source to have both atomic carbon ([CI]) transitions detected. Continuum measurements between 70 µm and 3 mm are used to constrain the far infrared (FIR) spectral energy distribution (SED), and we find a best fit FIR luminosity of $\log_{10}[L_{\text{FIR}}/L_\odot] = 13.5$, dust temperature $T_D = 88$ K and emissivity index $\beta = 0.6$. We combine the new CO measurements with an existing CO(3-2) observation to construct the CO SED (also known as the CO ladder). By fitting large velocity gradient models to this ladder we can infer physical properties of the gas in AMS12.

2.1 Introduction

Obscured quasars offer the opportunity to investigate their hosts as the ultraviolet (UV) and optical emission from their central engines is obscured along the line of sight by intervening gas and dust. Some obscured quasars have been proposed to be at an evolutionary phase where the host galaxies contain more gas and dust
### 2.1 Introduction

(Sanders et al. 1988; Fabian 1999). Therefore obtaining stellar, gas and dust mass estimates, inferring the star formation rate (SFR) and the properties of the quasar, can reveal the connections between the quasar and its host, and test the evolution of the ratio between the black hole mass ($M_\bullet$) and the host galaxy’s bulge properties, i.e. stellar velocity dispersion ($\sigma_\star$), luminosity and mass ($M_{\text{bulge}}$).

Studies into the gas and dust in high redshift ($z \sim 2$) quasar hosts allow us to observe an important epoch in galaxy formation. Determining the physical properties of the gas and dust allow for the characterisation of these galaxies, providing comparisons to the nearby Universe. The gas and dust are thought to be heated by nearby star formation (SF) (i.e. heating by young OB stars), and possibly by the AGN itself (e.g. Blain et al. 2002, and references within). Investigations into possible AGN heating of the gas and dust are important to distinguish the SF and AGN contributions to the far-infrared luminosity ($L_{\text{FIR}}$), and the CO luminosity (e.g. Barvainis 1987; Siebenmorgen et al. 2004; Solomon & Vanden Bout 2005).

An important constituent in the host galaxy for star formation is the molecular gas from which stars are made. Molecular gas consists almost entirely of molecular hydrogen, H$_2$, which requires cool temperatures in order to form on the surface of dust grains. Photon absorption causes dissociation of the molecule from the grain. H$_2$ is symmetric and lacks a permanent dipole, leading to no rotational transitions. The vibrational states of H$_2$ can be excited by high temperatures. However, the emission from these states falls in the mid-infrared regime, making direct observations of H$_2$ very difficult from ground-based telescopes. Therefore the second most abundant molecule carbon monoxide (CO), is used as a tracer for the molecular gas. CO is excited by collisions with H$_2$ which highlights the importance of the gas density. An important property of molecules in the gas is the density at which the transition lines’ collisional excitation and the spontaneous emission rates are equal. This is called the critical density, ($n_{\text{crit}}$) and the CO molecule’s rotational
transitions have low critical densities. Therefore they are able to be excited and thermalised by collisions with $\text{H}_2$ at low gas densities.

Observing multiple CO rotational transition lines provides the CO SED, or “CO ladder”. This can be used to infer the physical properties, e.g. the kinetic temperature and density, of the molecular gas via the fitting of detailed models such as large velocity gradient (LVG) models (Scoville & Solomon [1974]; Goldreich & Kwan [1974]).

Other emission lines such as HCN, [CI] and [CII], have been observed in galaxies at high redshifts, often with the help of gravitational lensing (e.g. Solomon & Vanden Bout [2005]; Walter et al. [2011]). The atomic carbon molecule [CI] is closely related to CO emission; the critical density of [CI] is close to that of CO(1-0) ($\sim 10^2 \text{ cm}^{-3}$), and is thought to map the CO emission; studies of the [CI] and CO in Orion by Ikeda et al. (2002) show the emission comes from the same regions. [CI] can be described by the two optically-thin lines of a 3-level system and we can use the line ratio to directly determine the properties of the gas.

Detailed studies of the CO SED of galaxies at low and high redshifts have revealed a range of gas properties, from conditions similar to nearby normal galaxies, to extreme conditions where the gas temperature reaches in excess of 200 K (e.g. APM 08279+5255, Weiß et al. [2007]). These appear, in some cases, to have a dependence on the possible contributors to the heating mechanism of the gas with known AGN hosts and starburst galaxies appearing to have higher kinetic gas temperatures than regular galaxies (e.g. Solomon & Vanden Bout [2005], and references within).

The first high redshift CO source detected was IRAS F10214+4724 with $z = 2.286$ (Brown & Vanden Bout [1991]; Solomon, Downes & Radford [1992]). Since then an increasing library of high redshift CO sources have been detected and studied (see Solomon & Vanden Bout [2005] for a review). Some of the high redshift galaxies detected in CO and other tracers of the molecular gas have been magnified by
gravitational lensing (e.g. [Solomon & Vanden Bout, 2005]). Therefore the studies of unlensed high redshift galaxies are useful to derive the physical conditions of the gas without the possibility of differential magnification complicating the interpretation of the observations, i.e. [Deane et al., 2013] found the gravitational lensing of IRAS F10214+4274 showed different magnification factors depending on the size and location of the emitting region.

The unlensed, obscured quasar, AMS12, at redshift $z = 2.767$, has been observed at multiple wavelengths ([Martínez-Sansigre et al., 2005, 2006a,b, 2009, Klöckner et al., 2009]). The redshift was determined from the Ly$\alpha$, CIV and HeII lines in the optical spectrum ([Martínez-Sansigre et al., 2006a]). Its mid-IR SED suggests the galaxy corresponds to a progenitor of the present-day $\sim 2L^*$ galaxies ([Martínez-Sansigre et al., 2005, 2006a,b]). A strong detection of the CO(3-2) line was first presented in [Martínez-Sansigre et al., 2009], prompting further investigation into the CO ladder, which we conduct here.

2.2 Molecular line observations

2.2.1 Radio observations

Fundamentals

Here we briefly outline a few concepts used in radio interferometry. We are ignoring the $w$ or $\sqrt{1 - l^2 - m^2}$ terms since this is a very basic overview. For a comprehensive introduction to radio interferometry see [Taylor, Carilli & Perley, 1999].

Radio telescopes measure the fluxes of objects in brightness temperature units which are converted into observer units such as the Jansky (where 1 Jy = $10^{-26}$ Wm$^{-2}$Hz$^{-1}$) via a conversion factor that is specific to each telescope and frequency.
dependent. Radio interferometers are an array of radio telescopes/receivers that are connected to one another; the space between each telescope is called the baseline and it is often expressed in spatial frequency units \((u, v)\) which are the distance in wavelength units. The signals from each individual antenna are correlated to produce visibilities \(V(u, v)\) which can be expressed via,

\[
V(u, v) = \int \int A(l, m)I_\nu(l, m)e^{-2\pi(ul+vm)}dl\,dm,
\]

where \((l, m)\) are angular distances on the sky, \(A(l, m)\) is the response from the interferometer at each baseline, \(I_\nu(l, m)\) is the source function, i.e. the intensity of the source observed. The visibilities are the Fourier transform of the interferometer response at each spatial frequency \((u, v)\). The interferometer can only measure discrete points in \(uv\) space, i.e. the \((u, v)\) of each baseline, so the actual data we get from an array is the sampled visibility function (the sampling function multiplied by the visibility function). We can perform the inverse Fourier transform on the sampled visibility function to produce a dirty map (which is the true sky convolved with the dirty beam). The dirty beam is the Fourier transform of the sampling function and contains positive and negative sidelobes due to the radiation pattern of the antenna (positive) and the missing spaces in the discrete sampling pattern (negative).

Deconvolution

To recover a higher quality image, deconvolution is required. This process effectively is an attempt to recover the true sky image from the dirty image given the incomplete (discrete) sampling of the visibility function. One common approach is the CLEAN algorithm from Högbom ([Högbom](#)) [1974]. This algorithm assumes the sky is made up of point sources. It takes the brightest point (pixel) in the image and measures its strength and position. It then takes the dirty beam multiplied by
the strength and some damping factor away from the dirty image at the position of this pixel, and repeats until there are no more bright peaks (i.e. all are below some specified level). This resulting image, with no bright sources remaining, is called the residual image. The accumulated bright source map is generated and then convolved with a synthesised beam (usually an elliptical Gaussian fitted to the primary lobe of the dirty beam). The residuals are then added to the “CLEANed” image. CLEANing is effective at removing the negative sidelobes present in the dirty image and performs well for point sources, but not so well for resolved, extended emission. We used the CLEAN algorithm since our observations of the molecular gas are detection campaigns, and are not observed with high resolution telescope configurations. As such we are not expecting to resolve the emission and they will effectively be point sources (albeit smeared by the beam).

2.2.2 IRAM PdBI observations

Interferometer

The CO observations were done with the Institut de Radioastronomie Millimétrique (IRAM) Plateau de Bure Interferometer (PdBI)\(^1\) which is a millimeter interferometer in the French Alps on the Plateau de Bure at an elevation of 2550 meters. The PdBI consists of six 15 meter diameter antennas each equipped with four receiver bands with dual polarisation in the 3, 2, 1.3 and 0.8 mm atmospheric windows. The telescopes can reach a maximum separation of 368 m in the north-south direction, and 760 m in the east-west direction (“A” configuration). The space between each telescope and a reference telescope is what is commonly referred to as a baseline. All our observations were carried out in the most compact “D” configuration which has the lowest angular resolution (\(~5^\prime\) at 100 GHz\), but is suitable

\(^1\)http://www.iram-institute.org/EN/content-page-56-7-56-0-0-0-0.html
for detection observations. In this most compact configuration the short spacings between the telescopes become an issue when observing a source at low elevation angles where one telescope’s line of sight is partially blocked by another; this effect is called shadowing and affects observations.

2.2 Molecular line observations

CO and [CI] line observations and data reduction

The CO(3-2) molecular line towards AMS12 was observed in 2009 using the 3 mm band of PdBI centred on 91.796 GHz. These observations were part of a search for CO in the host galaxies of two of the obscured quasars studied by Martínez-Sansigre et al. (2009). The observations of AMS12 were made on 30 April and 13 May 2009, using the narrow-band correlator with 1 GHz bandwidth. The calibrators 3C 345, MWC 349, 2145+067 and 0923+392 were used as flux or bandpass calibrators and the phase calibrator for both days was 1637+574. On 30 April, only 5 antennas were available and also 50% of the data were flagged on this day. The final data cube achieved an rms noise of 0.70 mJy beam$^{-1}$ per 30 kms$^{-1}$ channel. The resulting spectrum can be seen in Figure 2.1a (as published in Martínez-Sansigre et al., 2009).

The CO(5-4) and CO(7-6) rotational transitions were observed with the PdBI for 4 nights during April and May 2010. These observations used 6 antennas in the compact D configuration, with dual polarization using the new wide-band correlator, WideX, simultaneously with the narrow-band correlator. The WideX correlator covers a bandwidth of 3.6 GHz with a fixed spectral resolution of 2 MHz. Using the WideX correlator which quadruples the bandwidth available with narrow-band correlator, increases the continuum sensitivity and the instantaneous spectral coverage. Phase and amplitude calibrators were observed between on-source observations, every $\sim$ 20 minutes. Every $\sim$ 40 minutes the pointing and focus was re-adjusted.

Flagging is the act of rejecting certain data due to some effect that deems it unsuitable such as external interference (noise) or systematic problems.
2.2 Molecular line observations

The data were reduced and analysed by the author, using the GILDAS software. During this period, the reduction process showed that antenna one was underperforming on all observing runs, with an efficiency ranging from 0.48 on 03 May, to a more acceptable 0.88 on 24 April.

The observations of CO(5-4) were made on the 21 April under good conditions using PdBI’s 2 mm band centred on 152.986 GHz. One of the antennas suffered shadowing at the beginning of the run and was auto-flagged. The auto-flagging of this one antenna was overridden to compute the phase calibration using the calibrator 3C454.3. The flux calibrator used was MCW349. The quality assessment limits were set to flag data which exceeded maximum phase rms of 45 degrees and maximum amplitude loss of 23%. These thresholds were the default values given by the software for the 2 mm band. The number of observations flagged were 13% and 12% due to amplitude and phase differences, respectively. No observations were flagged due to pointing or focus errors.

The final step in the reduction process was to create a $uv$ table of the observations which is a table of all the visibilities measured at each baseline, and the $u$ and $v$ positions of each baseline. This was done using the WideX correlator which required us to resample the $uv$ table adding 500 MHz bluewards from the rest frequency of 152.986 GHz, to the spectrum (see Figure 2.1b). This is due to the large bandwidth of the WideX correlator. The internal frequency range of the receiver is from 4 to 8 GHz and the line was centred on the frequency which gives the best performance in the narrow-band configuration, 6500 MHz (note that the narrow-band and wide-band observations were simultaneous). The $uv$ table was created using the GILDAS CLIC software, with 230 channels and a resolution of 30 km s$^{-1}$ per channel. An rms noise of 0.95 mJy beam$^{-1}$ per 30 kms$^{-1}$ bin for the final data cube was reached.

The CO(7-6) line was detected in the 1.3 mm band, centred on 214.148 GHz. The

\footnote{http://www.iram.fr/IRAMFR/GILDAS}
2.2 Molecular line observations

observations were made over 24 – 25 April and 03, 05 May under variable weather conditions. The observations from 25 April were omitted from the analysis due to their poor quality with inclement weather conditions and system temperatures above 600 K. The remaining days’ observations were each affected by worsening weather conditions near the end of their runs. The weather conditions on 24 April deteriorated significantly as the run progressed.

The same data quality assessment criteria were used for all the 1.3 mm observations, with data flagged if they exceeded phase rms of 50 degrees and amplitude loss of 25%. Of the data from 24 April, 13% were flagged due to amplitude corrections, 3% flagged for exceeding maximum phase corrections and no data were flagged for the pointing and focusing adjustments. The observations made on 03 May suffered worsening weather conditions yet again and only ∼ 40 minutes of on-source observations were usable. There were no data flagged for exceeding the maximum corrections throughout the calibration. The observations on 05 May were made under better weather conditions. Around nine hours into the observing run the antenna elevation reached a height where antenna 5 rotated 360 degrees in order to keep tracking the target, this could be seen throughout the calibration. Using the same data quality assessment parameters, 11% of the data were flagged due to amplitude corrections and 7% for phase corrections.

The $uv$ tables created for each of these days were appended, creating a single table for the transition. The $WideX$ correlator data were used to make the $uv$ table. The number of channels in the table was 160 at a resolution of 30 km s$^{-1}$. The final data cube has an rms noise of 2.0 mJy beam$^{-1}$ per 30 kms$^{-1}$ bin.

The 1.3 mm spectrum made with the $WideX$ correlator also revealed the atomic carbon upper fine structure line $[\text{CI}](^3P_2 \rightarrow ^3P_1)$. With a rest frequency of 809 GHz at redshift $z = 2.7668$, the CI($^3P_2 \rightarrow ^3P_1$) line falls at 214.858 GHz, within the $WideX$ bandwidth centred on 214.148 GHz. The CO(7-6) and CI($^3P_2 \rightarrow ^3P_1$) detections
2.2 Molecular line observations

are shown in Figure 2.1c.

The emergence of the atomic carbon upper fine structure line $[^3P_2 - ^3P_1]$ prompted an investigation into the detection of the $[^3P_1 - ^3P_0]$ line in AMS12. These observations were carried out with the PdBI using 5 antennas in the D configuration over 7 days during July and August 2011 with the reduction and analysis done by the author. The observations were centred on 130.660 GHz ($\nu_{\text{rest}} = 492.161$ GHz) which falls in the 2 mm band. A total of 12.3 hours of on-source integration time was obtained with varying weather conditions. Data were flagged on the 19 July, 02 and 05 August due to receiver system temperatures exceeding $\sim 400$ K. The $uv$ table was created from the WideX correlator data with an rms noise of 0.577 mJy beam$^{-1}$ per 20 MHz bin (see Figure 2.1d).

The flux calibrator MCW349, used for all of the observations is accurate to $\sim 7\%$ in the 3mm, $\sim 10\%$ in the 2mm and $\sim 15\%$ in the 1.3 mm bands. These errors were also included when fitting the CO lines.

Data imaging

The GILDAS MAPPING software was used to image the $uv$ tables and perform deconvolution and cleaning routines to produce the datacubes used for the spectral line analysis. The imaging task (UVMAP), produces a dirty image and dirty beam via Fast Fourier Transform algorithms which require the regular gridding of visibilities in the $uv$-plane. Natural weighting was used in the mapping of all data to maximise the point source sensitivity. We used the software’s recommendation for the image and map sizes for the datacube set up. The recommended values ensured adequate sampling for producing the images. The gridding requires Nyquist sampling (i.e. at least twice the highest waveform frequency), in order to prevent information loss and to fully reconstruct the signal. We used the recommended gridding parameters given by the software.
2.2 Molecular line observations

Figure 2.1: Continuum subtracted CO molecular and [CI] atomic lines from the PdBI 3, 2 and 1.3 mm observations. (a): The 3 mm band PdBI spectrum of AMS12 with the narrow band correlator (1 GHz bandwidth) showing the CO(3-2) emission line. (b): Spectrum from the 2 mm band shows the CO(5-4) line. (c): Spectrum from 1.3 mm band shows [CI]($^3P_2 - ^3P_1$) and CO(7-6). (d): The [CI]($^3P_1 - ^3P_0$) line detection. See Table 2.1 for the line parameters derived from the Gaussian fits to these lines.
2.2 Molecular line observations

Figure 2.1: (continued).

(c) [CI]$^{3}P_{2} - ^{3}P_{1}$ and CO(7-6)

(d) [CI]$^{3}P_{1} - ^{3}P_{0}$
The images were cleaned using the Hög bom CLEAN algorithm details of which are given in Section 2.2.1 (Hög bom, 1974). Cleaning of the dirty image involves performing a deconvolution to derive meaningful measurements. The images were run through the cleaning algorithm without defining a support area around the source first, then run through the algorithm again using a defined support area to assist with the convergence. The synthesised clean beam sizes are 6.67′′ by 5.18′′ with a position angle 61.14°, 4.17′′ by 2.85′′ with 70.67° position angle and 2.59′′ by 2.04′′ with position angle 101.78° for the 3, 2, and 1.3 mm maps.

The output following the imaging and deconvolution algorithms are the cleaned datacubes which have flux density measured in Jy/beam. These are then fed into the CLASS software to extract the spectrum of the line from a single pixel centred on the source’s position. With the resolution offered by the compact configuration, the source is unresolved and so extracting the spectra from single pixels retains most, if not all, of the information from the source.

The continuum was extracted from each spectrum then the MINIMIZE routine performed a fit of a theoretical profile to the spectrum, in this case a Gaussian.

2.3 Spectral line fitting

We assume the emitted molecular lines which are rotationally broadened, can be approximated by a Gaussian shape, and parameterised by their peak flux and the full-width half-maximum (FWHM). The linewidths for the three CO and two [CI] lines are \( \sim 200 - 300 \) km s\(^{-1}\). The CO(3-2), CO(5-4) and CO(7-6) line emission towards AMS12 are detected at 11\( \sigma \), 17\( \sigma \) and 11\( \sigma \) significance respectively. To determine the characteristic CO emission from AMS12, we co-add (stack) the individual lines and find the average CO line profile. Figure 2.2 shows the stacked CO transition profiles with the continuum subtracted and weighted by the inverse
Spectral line fitting

squared noise. The averaged CO lines are well fit with a Gaussian of FWHM of 288 ± 19 km s\(^{-1}\) and offset from \(z = 2.7668\) by \(-15 \pm 9\) km s\(^{-1}\) agreeing with \(z_{\text{CO}}\) of 2.7672 ± 0.0003 as determined by the individual lines.

The results from the Gaussian fits to the lines, the velocity integrated flux densities, \(I_{\text{CO}}\), and the line luminosities are shown in Table 2.1. The line luminosities of both the CO transitions and the [CI] line are given in both \(L_\odot\) (\(L_{\text{CO}}\)), and K km s\(^{-1}\) pc\(^2\) (\(L^T_{\text{CO}}\)) units. The line luminosity in \(L_\odot\) can be expressed as (Solomon, Downes & Radford 1992)

\[
\left( \frac{L_{\text{CO}}}{L_\odot} \right) = 1.04 \times 10^{-3} \left( \frac{I_{\text{CO}}}{\text{Jy km s}^{-1}} \right) \left( \frac{\nu_{\text{rest}}}{\text{GHz}} \right) (1 + z)^{-1} \left( \frac{D_{\text{lum}}}{\text{Mpc}} \right)^2 \quad (2.2)
\]

where \(\nu_{\text{rest}}\) is the rest-frame velocity of the line in GHz, and \(D_{\text{lum}}(z)\) is the luminosity distance in Mpc.

The alternative line luminosities we use are the brightness temperature luminosities \(L^T_{\text{CO}}\) in units K km s\(^{-1}\) pc\(^2\). These are calculated via

\[
\left( \frac{L^T_{\text{CO}}}{\text{K km s}^{-1} \text{ pc}^2} \right) = 3.25 \times 10^7 \left( \frac{I_{\text{CO}}}{\text{Jy km s}^{-1}} \right) \left( \frac{\nu_{\text{obs}}}{\text{GHz}} \right)^{-2} (1 + z)^{-3} \left( \frac{D_{\text{lum}}}{\text{Mpc}} \right)^2 \quad (2.3)
\]

where \(\nu_{\text{obs}} = \nu_{\text{rest}}/(1 + z)\). \(L^T_{\text{CO}}\) is proportional to the brightness temperature. The ratio of the \(L^T\) for two lines (\(L^T_{\text{CO}(5-4)}/L^T_{\text{CO}(3-2)}\)), assuming they are from the same region, is a direct measure of the ratio of the intrinsic brightness temperatures. This is useful as we can gain information about the gas temperature from the brightness temperature luminosities when higher transition lines are thermalised. When lines are thermalised, their brightness temperature luminosities are equivalent to the lower transition’s, \(L^T_{\text{CO}(J+1)-J} = L^T_{\text{CO}(J-J-1)}\), and the kinetic temperature of the gas is equal to the \((J + 1) - J\) level’s excitation temperature.
Figure 2.2: The average CO profile from the CO(3-2), CO(5-4) and CO(7-6) transitions. The transitions are weighted by the inverse square of the noises. The profile is fit with a Gaussian centred on -15 km s$^{-1}$ and with a FWHM of 288 ± 19 km s$^{-1}$. 
### Table 2.1: Observed CO and CI line parameters towards AMS12 derived from the individual Gaussian fits.

<table>
<thead>
<tr>
<th>Line</th>
<th>ν_{obs} [GHz]</th>
<th>S_ν [mJy]</th>
<th>ΔV_{FWHM} [kms^{-1}]</th>
<th>I_{CO} [Jy kms^{-1}]</th>
<th>V^a [kms^{-1}]</th>
<th>L/10^8 [L_⊙]</th>
<th>L^T/10^{10} [K kms^{-1} pc^2]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO(3-2)</td>
<td>91.803</td>
<td>3.4 ± 0.3</td>
<td>300 ± 29</td>
<td>1.088 ± 0.101</td>
<td>-22.9 ± 14.3</td>
<td>0.552 ± 0.051</td>
<td>4.181 ± 0.388</td>
</tr>
<tr>
<td>CO(5-4)</td>
<td>152.898</td>
<td>6.9 ± 0.4</td>
<td>244 ± 21</td>
<td>1.781 ± 0.136</td>
<td>-4.8 ± 9.2</td>
<td>1.507 ± 0.115</td>
<td>2.465 ± 0.188</td>
</tr>
<tr>
<td>CO(7-6)</td>
<td>214.159</td>
<td>8.1 ± 0.7</td>
<td>318 ± 42</td>
<td>2.747 ± 0.310</td>
<td>-16.0 ± 17.8</td>
<td>3.254 ± 0.367</td>
<td>1.940 ± 0.219</td>
</tr>
<tr>
<td>[CI](^3P_1 - ^3P_0)</td>
<td>130.657</td>
<td>2.2 ± 0.27</td>
<td>279 ± 50</td>
<td>0.646 ± 0.141</td>
<td>10.5 ± 21.8</td>
<td>0.467 ± 0.102</td>
<td>1.217 ± 0.266</td>
</tr>
<tr>
<td>[CI](^3P_2 - ^3P_1)</td>
<td>214.837</td>
<td>5.9 ± 0.9</td>
<td>232 ± 29</td>
<td>1.451 ± 0.250</td>
<td>35.5 ± 21.0</td>
<td>1.724 ± 0.297</td>
<td>1.018 ± 0.175</td>
</tr>
</tbody>
</table>

^aVelocities are reported relative to a redshift of 2.7668.
2.4 PdBI continuum measurements

The CO(3-2) line in AMS12 was detected with the PdBI in 2009. The continuum from the PdBI observations was measured by Martínez-Sansigre et al. (2009). Further observations in the PdBI 2 and 1.3 mm bands were made providing further continuum measurements described below. See Section 2.2 for details on the calibration and reduction of the PdBI line observations.

The continuum measurements for the 3, 2, and 1.3 mm bands were made using GILDAS software. Firstly, the spectral feature(s) in each of the bands were excluded by setting a window around the areas, then a polynomial was fit to the remaining featureless spectrum. The polynomial fitting algorithm used a Chebyshev polynomial of one degree. This procedure was done before the Gaussian fitting of the spectral lines, which ensured the continuum emission was subtracted from each spectra before the Gaussian fits. Thus the peak line values are not enhanced by the underlying continuum.

The continuum towards AMS12 in the 3 mm band was found to be $S_{3\text{mm}} = 0.75 \text{ mJy}$, excluding the region from -396.9 to 250.7 km s$^{-1}$, which encompasses the CO(3-2) line. The total width in frequency terms of the remaining spectrum for continuum measurements is 684.4 MHz or 2053.8 km s$^{-1}$. This corresponds to 68.5 bins of 30 km s$^{-1}$ (10 MHz), leading to an rms noise of 0.08 mJy over the full 684.4 MHz range of continuum.

For the 2 mm band, the CO(5-4) line feature extended 505.1 km s$^{-1}$ and was excluded from the continuum measurements. The remaining 3.35 GHz of the spectrum was fit yielding a continuum measurement of $S_{2\text{mm}} = 1.07 \text{ mJy}$ with an rms noise of 0.06 mJy. The 1.3 mm band yielded the continuum measurement of $S_{1.3\text{mm}} = 2.16 \text{ mJy}$ over a spectral line free region spanning 2.572 GHz, giving an rms noise of 0.19 mJy.


## 2.5 The far-infrared dust

The PDBI continuum measurements are used along with existing Max-Planck Millimetre Bolometer Array (MAMBO), Spitzer and Herschel data in various wavelength bands to construct the FIR SED. The FIR SED work was completed by A. Martínez-Sansigre and details of the construction and fitting are given in Appendix A. The results from the fitting will be used in the proceeding analysis of the obscured quasar AMS12.

At FIR wavelengths dust is not optically thick, and the SED of the radiation can be described by a modified black body. If the absorption coefficient of the dust, $\kappa(\nu)$, is assumed to follow a law $\propto \nu^{\beta}$, the emission will be given by,

$$L_\nu = \frac{A\nu^{3+\beta}}{h(\epsilon^{\frac{h\nu}{kT_D}} - 1)},$$

where $A$ is a normalisation term given by:

$$A = \frac{L_{\text{FIR}}}{\zeta(\beta + 4)\Gamma(\beta + 4) \frac{h}{kT_D}}.$$  

The three variables are: the dust temperature $T_D$, the emissivity index $\beta$ and the FIR-luminosity $L_{\text{FIR}}$. Here $h$ and $k$ are Planck’s and Boltzmann’s constants, respectively, while $\zeta$ and $\Gamma$ are the Riemann zeta function and the Gamma function, respectively.

Figure A.2 shows the FIR SED for AMS12. The best fitting temperature is $T_D = 88$ K and emissivity index of $\beta = 0.6$. The best fitting $L_{\text{FIR}}$ determined from the single graybody model fitting of the FIR SED is $\log_{10}(L_{\text{FIR}}/L_\odot) = 13.5$. The $L_{\text{FIR}}$ can be used to determine the star formation rates (SFRs) of the galaxies. Assuming that the $L_{\text{FIR}}$ is solely due to star formation, with young OB stars the main
source of heating, we can determine the SFR from the Kennicutt (1998) conversion,

\[
\left( \frac{\text{SFR}}{[M_\odot \text{ yr}^{-1}]} \right) = 1.7 \times 10^{-10} \left( \frac{L_{\text{IR}}}{[L_\odot]} \right).
\] (2.6)

In order to obtain the total SFR, the assumption of an initial mass function (IMF) is required. Here, the Salpeter (1955) IMF is assumed.

The SFR for AMS12, given the \( L_{\text{FIR}} \) determined from the best dust fitting model, is \( \sim 5300 \, M_\odot \text{ yr}^{-1} \). SFR of this scale are seen in only the most extreme starburst galaxies (see Chapman et al. 2005, Tacconi et al. 2006, Coppin et al. 2008, Solomon & Vanden Bout 2003 for a review).

The typical temperature found for star forming galaxies in the local Universe and at high redshift is \( \sim 50 \, \text{K} \) (e.g. Farrah et al. 2003, Kovács et al. 2006, Elbaz et al. 2011). The temperature derived from the FIR SED of AMS12 is significantly higher than this. If a significant fraction of the \( L_{\text{FIR}} \) is due to heating from the AGN, this SFR is severely overestimated. We discuss the implications in Section 3.5.

### 2.5.1 Dust mass

The mass of the FIR dust can also be found from the \( L_{\text{FIR}} \) by (e.g. Beelen et al. 2006),

\[
M_D = \frac{L_{\text{FIR}}}{4\pi \int \kappa_{\nu_{\text{rest}}} B_{\nu_{\text{rest}}} (T_D) d\nu}, \tag{2.7}
\]

where \( \nu_{\text{rest}} \) is the rest frame frequency found by \((1 + z)\nu_{\text{obs}}\). The mass absorption coefficient \( \kappa_{\nu_{\text{rest}}} = \kappa_{\nu_{\text{obs}}} (\nu_{\text{rest}}/\nu_{\text{obs}})^\beta \) is given by a power law. The mass absorption coefficient is the main source of uncertainty of the dust mass. We assume two different reference values of \( \kappa_{\nu_{\text{obs}}} \); first, \( \kappa_{250 \, \text{GHz}} = 0.04 \, \text{m}^2\text{kg}^{-1} \) (i.e. Alton et al.)
Summary

With the PdBI we have detected three CO rotational transitions, (CO(3-2), (5-4) and (7-6)), and both of the atomic carbon fine structure lines, ([CI](3P_1 - 3P_0) and [CI](3P_2 - 3P_1)). These lines were all detected to between 7σ and 17σ significance. This is the first detection of both the [CI] fine structure lines in a high redshift, unlensed object. The continuum from the PdBI line observations was combined with observations from MAMBO, Spitzer and Herschel data to construct the dust FIR SED (Appendix A). We can use these observations of the CO ladder and the [CI] lines to model the molecular gas properties of AMS12 using LVG models. The details of the modelling will be given in the next chapter. Along with the molecular gas properties, we can use the FIR dust properties to characterise the heating of the gas and dust in AMS12.
Chapter 3

Modelling the molecular gas in AMS12

In this chapter we use the highly-excited molecular gas probed by CO(3-2), (5-4) and (7-6) to make the CO ladder which we model with large velocity gradient (LVG) models. The gas kinetic temperature $T_G$, density $n(H_2)$, and the characteristic size $r_0$, are determined first assuming flat priors on all parameters, and then using the dust temperature from the FIR SED as a prior for the gas temperature. The best fitting parameters using the dust temperature as a prior, are $T_G = 89.6$ K, $n(H_2) = 10^{3.9}$ cm$^{-3}$ and $r_0 = 0.8$ kpc. The ratio of the [CI] lines gives a [CI] excitation temperature of $43 \pm 10$ K, indicating the [CI] and the high-excitation CO are not in thermal equilibrium. The [CI] excitation temperature is below that of the dust temperature and the gas kinetic temperature of the high-excitation CO, perhaps because [CI] lies at a larger radius where there may also be a large reservoir of CO at a cooler temperature. We explore this idea further using the $[CI]^{3P_1 - 3P_0}$ line to estimate the CO(1-0) line strength and proposing a two component model to explain the discrepancy between the best fitting gas temperature from the high excitation CO lines and the [CI] excitation temperature.
3.1 Motivation

The detection of multiple rotational transition lines of CO towards AMS12 gives the CO ladder. We model the CO ladder with LVG models with which we can infer the physical characteristics of the gas, namely the temperature, the density and the characteristic size. Once we have characterised the gas we can then infer the location of the gas, i.e. its proximity to the nucleus of AMS12, and investigate the origins of the heating. For example, if the temperature of the gas from the modelling is very high and the emitting region small (assuming this indicates proximity to the nucleus), this would hint at possible heating by the AGN. Before fitting the LVG models to the CO ladder we can use the observations as they stand to determine whether the CO transitions are thermalised.

The excitation of CO to higher rotational levels depends upon the ambient temperature and the density of the gas. These lines are optically thick which makes it difficult to infer the physical properties of the gas directly. In the case where the CO transitions are thermalised, we may determine the temperature of the gas. Thus we can use the measured brightness temperature luminosities of the lines to gain an idea of the gas conditions before detailed modelling.

An initial interpretation of the CO lines detected in AMS12, reveal they are subthermally excited. The $L^{T}_{\text{CO}(5-4)}/L^{T}_{\text{CO}(3-2)}$ ratio is $0.59 \pm 0.05$, and the $L^{T}_{\text{CO}(7-6)}/L^{T}_{\text{CO}(3-2)}$ ratio $0.46 \pm 0.06$. This further warrants the more detailed investigation into the CO ladder via LVG modelling.

3.2 LVG modelling

In order to model the line intensities of the rotational transitions of CO, coupled equations of radiative transfer and statistical equilibrium must be solved. This
3.2 LVG modelling

involves solving for radiative excitation and de-excitation of the CO levels as well as collisional excitation and de-excitation. The rotational levels of CO are populated and de-populated via these processes, and the difference in the neighbouring levels’ population is used in the determination of the optical depth of the transition.

In the case where the radiative terms and the collisional terms are both significant to the net decay/growth of an excited level, it is necessary to evaluate the rate of change of the energy level populations at every point in the region. A simplification which is used to solve the coupled equations comes in the form of the LVG approximation. This assumes that there are large velocity gradients across the area of the gas which are significantly greater than local thermal velocities. The line photon from the de-excitation from the $J$ state to $(J - 1)$ within the CO molecule can be absorbed only within a short distance of where it was emitted due to the velocity gradient. It follows that radiation which has interacted with a molecule must either have originally been emitted by molecules a short distance away or be from the background radiation (e.g. Goldreich & Kwan, 1974; Scoville & Solomon, 1974; Genzel, 1992).

The probability that the photon will escape the region is given by a term called the escape probability. This is given by $\beta_\nu = [1 - e^{(\tau_\nu)}]^{\tau_\nu^{-1}}$ where $\tau_\nu$ is the optical depth of the line transition $J - (J - 1)$ represented here by the frequency $\nu$ (e.g. given by Equation 1.8). The line optical depth is dependent on the velocity gradient, the molecule number density and the level populations. When the transition has a large optical depth, the line photon of frequency $\nu$, locally trapped by the velocity gradient, will most likely be absorbed by a nearby CO molecule. Thereby, the same excitation takes place radiatively many times before the photon is lost. Photon trapping acts to raise the intensity of the line transition above the background intensity (i.e. from the cosmic microwave background radiation). Figure 3.1 shows a cartoon representation of the LVG approximation described here.
3.2 LVG modelling

The net transfer of molecules in the excited state $J$ to level $(J - 1)$, is determined by the photons which escape the region in which they were emitted. Thus, for large optical depths, the escape probability is less and the fraction of molecules in the higher $J$ state is more stable.

The intensity of the line emission above the background emission is determined from the spontaneous emission rate, the kinetic temperature of the gas, the optical depth of the transition lines (related to the level occupations, the velocity gradient and the density of the molecule), and the line collisional rate which depends upon the density and temperature of the gas. These are all locally determined properties of the gas when the LVG approximation is assumed. We refer the reader to Scoville & Solomon (1974) and Goldreich & Kwan (1974) for a complete derivation of the model in a collapsing spherical molecular cloud.

We have used a LVG code developed and kindly provided by C. Henkel. The LVG calculations require the collision coefficients of the molecules under consideration, in addition to a few input parameters. These are: an ortho-to-para ratio for $\text{H}_2$, the redshift of the source (needed to calculate the temperature of the background cosmic microwave background (CMB) radiation), a chemical abundance of the molecule relative to $\text{H}_2$ and the velocity gradient. The free parameters are the gas kinetic temperature, $T_G$, and the overall density of molecular hydrogen, $n(\text{H}_2)$. The outputs of the calculation include the occupation numbers, the excitation temperatures, $T_{\text{ex}}$, the Rayleigh-Jeans brightness temperatures, $T_b$, and the optical depths of the transitions of the observed molecule.

We have used single-component LVG models using the collision rates from Flower (2001), to investigate the CO excitation. In all calculations we used the $\text{H}_2$ ortho-to-para ratio of 3:1, and a cosmic microwave background temperature of $T_{\text{CMB}} = 10.28$ K (corresponding to the redshift $z = 2.767$). We adopted the fixed CO abundance per velocity gradient value of $[\text{CO}]/(\text{dv/dr}) = 1 \times 10^{-5}$ pc/kms$^{-1}$ (e.g. Weiß, ...
Figure 3.1: A cartoon depiction of the large-velocity gradient approximation. The velocity gradient within the cloud is shown as the vector $\mathbf{v}$, which under the LVG assumption allows the simplification of solving the coupled equations of radiative transfer and statistical equilibrium locally. The local region is shown by the black box within which are carbon monoxide molecules depicted by the dark grey and red circles. Molecules are excited by collisions, background radiation (depicted by photon $\nu_{\text{CMB}}$), and re-absorption of line photons as a result of photon trapping. The probability that a photon of a transition escapes the boxed region is given $\beta_\nu$ which depends on the optical depth of the transition. The line photon escaping this region is depicted by the photon $\nu_{\text{line}}$. Dust is interspersed within the molecular gas cloud and is involved in the production of the molecules and shielding from photodissociation of these molecules (see Section 2.1).
3.3 CO LVG model results

Walter & Scoville, 2005b; Weiß et al., 2007).

The LVG model provides the brightness temperatures $T_b$, of each rotational transition from $J = 1 \rightarrow J = 11$, which can be compared to the observed flux densities by,

$$S_{\text{CO}} = \Omega \frac{T_b}{(1 + z)} \frac{2k\nu_{\text{obs}}^2}{c^2}$$  \hspace{1cm} (3.1)

where $\Omega$ is the source solid angle (e.g. Weiß et al., 2007). Due to the fact that we have not resolved the source, $\Omega$ is kept as a free parameter. We use the equivalent source radius $r_0$ which is given by $r_0 = D_A\sqrt{\Omega/\pi}$, where $D_A$ is the angular diameter distance.

Therefore we have three variables $T_G$, $n$(H$_2$) and $r_0$ (‘unknowns’), and three CO transitions (‘knowns’).

With LVG models, there is degeneracy between the parameters $T_G$ and $n$(H$_2$) (e.g. Weiß et al., 2007; Ao et al., 2008). The degeneracy arises from the dependency of the line optical depth on the level populations which in turn depend upon the values of $T_G$ and $n$(H$_2$). In order to counteract this degeneracy and constrain the density and temperature further, we can use the information we have obtained from the continuum observations. We present the results from the LVG modelling below, firstly assuming no prior knowledge on the parameters, and then applying a prior on the temperature from the dust analysis.

3.3 CO LVG model results

To determine the best-fit parameters to the LVG models we use Bayes theorem (e.g. Sivia & Skilling, 2006):
CO LVG model results

\[ p(model|data) = \frac{p(data|model)p(model)}{p(data)}, \quad (3.2) \]

where the ‘model’ encompasses all parameters pertaining to the model in question. In our case, the LVG models brightness temperatures with parameters \( T_G, n(H_2) \), converted to flux densities with parameter \( r_0 \), are fitted. The term \( p(data|model) \) is otherwise known as the likelihood. The prior, \( p(model) \), reflects the degree of knowledge we have \textit{a priori} on the model parameters. The term in the denominator \( p(data) \) is known as the evidence, which acts as a normalisation term when we are estimating parameters. The data are the observed flux densities \( S_{CO} \).

The posterior probabilities \( p(model|data) \) may be split into each parameter of interest through marginalisation. This is done by integrating the posterior probability distribution function (PDF) over the other parameters resulting in the marginalised likelihood of the parameter of interest. Rewriting Bayes’ theorem with the parameters under consideration explicitly stated and \( S_{CO} \) as the observable data we get,

\[ p(T_G, n(H_2), r_0|S_{CO}) \propto p(S_{CO}|T_G, n(H_2), r_0)p(T_G)p(n(H_2))p(r_0) \quad (3.3) \]

where the evidence is treated as a normalisation term. The likelihood \( p(S_{CO}|T_G, n(H_2), r_0) \), is given by a Gaussian distribution:

\[ p(S_{CO}|T_G, n(H_2), r_0) \propto e^{-\sum_i \left( \frac{S_{CO,i} - S_{CO,m}(T_G, n(H_2), r_0)}{\sigma_i} \right)^2} \quad (3.4) \]

with \( S_{CO,m}(T_G, n(H_2), r_0) \) being the predicted flux given \( T_G, n(H_2) \) and \( r_0 \). Assuming Gaussian errors, maximising the likelihood is equivalent to minimising the \( \chi^2 \) statistic where \( \chi^2 = \sum_i \left( \frac{S_{CO,i} - S_{CO,m}(T_G, n(H_2), r_0)}{\sigma_i} \right)^2 \).

In order to get the posterior PDF for one particular parameter we can marginalise \( p(T_G, n(H_2), r_0|S_{CO}) \) over the other two parameters. For example, to get \( p(T_G|S_{CO}) \)
we integrate over $n(H_2)$ and $r_0$.

$$p(T_G|S_{CO}) = \int \int p(T_G, n(H_2), r_0|S_{CO}) \ dn(H_2) \ dr_0$$ (3.5)

and similarly for $p(n(H_2)|S_{CO})$ and $p(r_0|S_{CO})$.

### 3.3.1 No assumption on prior knowledge of the gas kinetic temperature

The results from the LVG modelling using flat priors for all the parameters, show there is no single conclusive region of parameter space which points to the best fitting region as can be seen in Figure 3.2. The marginalised PDFs of the individual parameters are also shown in Figure 3.2.

The minimum temperature “floor” in Figure 3.2(a) corresponds to the temperature of the background radiation at this redshift. The minimum density “wall” reflects the minimum density required to excite the higher CO lines to the observed levels. Figure 3.2(a) shows there are two regions which are within the 1σ contour; the region surrounding the triangle and the region around the square.

Taking just our current CO measurements, we cannot confidently rule out either region. However, we have an upper limit of $\sim 15$ kpc on the extent of the CO(7-6) emission from the unresolved PdBI 1.3 mm observations. The sizes which correspond to these temperatures and densities, are all physically possible, with the higher-temperature/lower-density solutions having lower sizes of the order of 1 kpc and less. The low-temperature/high-density solutions have radii of a few kpc. Figures 3.2(b), 3.2(c), and 3.2(d) show the marginalised probabilities of the unknown parameters. The peaks of these individual PDFs correspond to a $T_G = 12$ K, density of $n(H_2) = 10^{3.8}$ cm$^{-3}$ and $r_0 = 0.7$ kpc, the combination of these individually marginalised
Figure 3.2: (a): The 1, 2, and 3σ contours of the temperature and density PDFs marginalised over the size. The open triangle marks the model of least $\chi^2$, while the open square marks an arbitrary spot in the higher temperature-lower density/size region, within the 1σ confidence interval. The best values of the marginalised individual parameters are marked with an ‘X’. (b): The PDF of the gas temperature marginalised over the density and size parameters. (c): The PDF of the gas density parameter marginalised over the temperature and the size. (d): The PDF of the size of the emitting region of gas marginalised over the temperature and density.
Figure 3.2: (continued).
3.3 CO LVG model results

values are marked by an ‘X’ in Figure 3.2a. They do not provide a good fit to the data.

The LVG model solution which corresponds to the lowest $\chi^2$ value has a $T_G = 11.6$ K, a density of $n(H_2) = 10^{6.1}$ cm$^{-3}$ and the size of the emitting region $r_0 = 6.2$ kpc. This is marked with a triangle in Figure 3.2a.

The low kinetic temperature of the gas from this solution is only slightly above the temperature of the background radiation at this redshift. In order to get the line intensities above the background radiation level, this solution has a high gas density. The optical depths of the lines from $J = 1$ to $J = 7$ are $\gg 1$. At high optical depths, the photons emitted in the de-excitation of the levels remain in the region longer i.e. do not escape. They are able to interact further, driving up the number of molecules in these excited states.

While the opacity of the lines is high, the collisional excitation and de-excitation processes in this model dominate over the radiative processes and the level populations are in actual thermal equilibrium. The excitation temperatures of these line transitions is equal to the kinetic gas temperature since the system is in local thermodynamic equilibrium (LTE).

At higher transitions ($J \geq 8$) the line optical depths are $< 1$ and the system is no longer in LTE. The lines undergo collisional de-excitation and the intensities quickly decline to levels no longer detectable above the background radiation. This manifests itself in the sharp decline of the solid line in Figure 3.3 above $J = 7$.

The dashed line in this figure is the CO SED corresponding to a secondary minimum $\chi^2$ region within the 1$\sigma$ contours of Figure 3.2a, with $T_G = 135$ K, $n(H_2) = 10^{3.8}$ cm$^{-3}$ and $r_0 = 0.7$ kpc (the square in Figure 3.2a). Here both the radiative and collisional excitation and de-excitation have significant effects on the line intensities and the full statistical equilibrium analysis must be considered.
Figure 3.3: The CO SED of the CO(3-2), CO(5-4) and CO(7-6) lines towards AMS12, with the SED given by the LVG model giving the lowest $\chi^2$ with $T_G = 11.6$ K, $n(H_2) = 10^{6.1}$ cm$^{-3}$, and the CO region size of $r_0 = 6.2$ kpc (solid line). This corresponds to the open triangle in Figure 3.2a. A second line, showing a model from within the higher temperature-lower density/size region ($T_G = 135$ K, $n(H_2) = 10^{3.8}$ cm$^{-3}$ and $r_0 = 0.7$ kpc), corresponding to the open square in Figure 3.2a, is there for comparison (dashed line).

Figure 3.3 illustrates that the difference between this high-temperature/lower-density solution (the dashed line) and the previous low-temperature/high-density solution (solid line) is most pronounced at the higher $J$ levels. Observations of the $J \geq 8$ transitions are needed to further constrain the LVG modelling at these higher transitions.

High resolution imaging of the gas (and dust) in AMS12 would provide a measurement on the size and would constrain the model further.
3.3 CO LVG model results

3.3.2 Using the dust temperature PDF as a prior for $T_G$

Here we work under the assumption that the gas and dust arise from the same regions, and are therefore at the same temperature. Indeed dust shields the gas from the ultraviolet and optical radiation preventing it from dissociating the molecules.

Gas and dust studies in other high redshift galaxies have shown both the dust and gas are compact on scales of less than $\sim 4$ to 8 kpc supporting the assumption that they arise from the same region (Tacconi et al. 2006; Riechers et al. 2006; Weiß et al. 2007; Younger et al. 2008). The dust temperature is often used to either constrain or justify a particular gas kinetic temperature (see for example Weiß et al. 2007; Ao et al. 2008; Greve et al. 2009, among others).

Using the PDF of the dust temperature as the prior for the gas temperature, (shown in Figure A.1b where $T_D$ peaks at 89 K), i.e. $p(T_G) = p(T_D)$ in Equation 3.3 we can rule out the lower temperature regions. With the prior on the temperature distribution, the contour map in Figure 3.4a shows a tight convergence on the temperature and density. The values of the least $\chi^2$ in this model are approximately equivalent to the peaks of the marginalised PDFs of the parameters which are shown in Figure 3.4. Due to the small number of data points and the shapes of the likelihood and prior PDF, the prior is having a greater effect than the likelihood on the PDF. The values for the kinetic temperature, density of the gas and the size of the emission region are $89.6 \pm 8$ K, $10^{3.9\pm0.06}$ cm$^{-3}$, and $0.8 \pm 0.01$ kpc.

The LVG model corresponding to these values of temperature, density and emitting region as determined with the use of the dust temperature PDF as a prior for $T_G$, gives the CO ladder displayed in Figure 3.5. The combination of these parameter values fall within the $1\sigma$ contours of Figure 3.2a, making them an acceptable fit to the CO SED assuming no prior knowledge.
Figure 3.4: (a): The 1, 2, and 3σ contours of the temperature and density PDFs marginalised over the size. The dust temperature is used as a prior for the gas temperature. The best values of marginalised temperature, and density are shown in each figure marked with an ‘X’. (b): The marginalised PDF of the gas temperature using the prior of $p(T_D)=p(T_G)$. Best fit with $T_G = 89.6 \pm 8$ K. (c): The marginalised PDF of the gas density parameter $\log_{10}[n(H_2)/\text{cm}^{-3}]=3.9 \pm 0.06$. (d): The marginalised PDF of the size of the emitting region of gas best fit with $\log_{10}[r_0/\text{pc}]=2.9 \pm 0.02$. 
Figure 3.4: (continued).
Figure 3.5: The CO SED of the CO(3-2), CO(5-4) and CO(7-6) lines towards AMS12, with the SED given by the LVG model with $T_G = 89.6$ K, $n(H_2) = 10^{3.9}$ cm$^{-3}$, and the CO region size of $r_0 = 0.8$ kpc, these have the highest probabilities once the prior on the temperature, $p(T_D)$, is applied.

3.4 Atomic carbon

The $[\text{CI}](^3P_2 - ^3P_1)$ line was detected in AMS12 to a $7\sigma$ significance which prompted a search for the $[\text{CI}](^3P_1 - ^3P_0)$ line. The $[\text{CI}](^3P_1 - ^3P_0)$ line was detected to an $8\sigma$ significance, these lines can be seen in Figure 2.1 and the line properties are given in Table 2.1. AMS12 is the first unlensed, high redshift galaxy with both the $[\text{CI}](^3P_1 - ^3P_0)$ and $[\text{CI}](^3P_2 - ^3P_1)$ lines detected.

With both the upper and lower fine structure atomic carbon lines detected, we may directly determine physical properties of $[\text{CI}]$ in AMS12.
3.4 Atomic carbon

3.4.1 Excitation temperature and mass of atomic carbon

The excitation temperature of [CI] can be directly determined, assuming the lines are optically thin, from the ratio of $I_{[\text{CI}](^3P_2-^3P_1)}^T$ to $I_{[\text{CI}](^3P_1-^3P_0)}^T$ (e.g. Stutzki et al. 1997). In order to relate the excitation temperature of [CI] to the gas kinetic temperature, we require the [CI] excitation to be in LTE (much in the same way we can estimate the $T_G$ directly from CO in LTE). If this is not the case, the excitation temperature of [CI] could be lower than the kinetic temperature.

The line column densities may be found using the integrated brightness temperatures of the lines (see Appendix A of Schneider et al. 2003, for a complete derivation). The excitation temperature is given by the ratio of column densities ($N_{ul}$) expressed by the ratio of the statistical weights ($g_{ul}$) of the levels and the Boltzmann factor,

$$\frac{N_{21}}{N_{10}} = \frac{g_{21}}{g_{10}} e^{-\frac{h\nu_{21}}{kT_{\text{ex}}}}. \quad (3.6)$$

Rearranging and equating constants the expression for the excitation temperature is found to be,

$$\frac{T_{\text{ex}}}{[\text{K}]} = \frac{38.8}{ln(\frac{R_{\text{CI}}}{R_{\text{CI}}} )} \quad (3.7)$$

where $R_{\text{CI}}$ is the ratio of the line brightness temperatures; these may be replaced with the $L^T$ of the lines as $L^T \propto T_b$, $R_{\text{CI}} \equiv \frac{L_{[\text{CI}](^3P_2-^3P_1)}^T}{L_{[\text{CI}](^3P_1-^3P_0)}^T}$.

The $T_{\text{ex}}$ can then be used to find the [CI] total column density and mass. Weiß et al. (2003) derive the beam averaged [CI] column density in the optically thin limit and use this and the area of the emitting region (given by the solid angle subtended by the source convolved with the beam and multiplied by the angular
distance squared, $\Omega_{ssb}D_A^2$), to derive the mass of [CI]. Solomon, Downes & Radford (1992) express the line brightness luminosity related to the emitting area, $L_T = 23.5\Omega_{ssb}D_{\text{lum}}I_{\text{CO}}(1 + z)^{-3}$, where the luminosity distance is related to the angular distance via $D_A = D_{\text{lum}}/(1 + z)^2$. Thus, Weiβ et al. (2003, 2005a), use this expression of $L_T$ and Equation 2.3 in order to determine the mass of [CI] in an unresolved source using $L_T^{[\text{CI}]}(^{3}\text{P}_1-^{3}\text{P}_0)$ via,

$$
\left( \frac{M_{\text{CI}}}{M_\odot} \right) = 5.706 \times 10^{-4} Q(T_{\text{ex}}) \frac{1}{3} e^{T_1/T_{\text{ex}}} \left( \frac{L_T^{[\text{CI}]}(^{3}\text{P}_1-^{3}\text{P}_0)}{[K \ \text{km}\ \text{s}^{-1}\ \text{pc}^2]} \right) 
$$

(3.8)

where $Q(T_{\text{ex}}) = 1 + 3e^{-T_1/T_{\text{ex}}} + 5e^{-T_2/T_{\text{ex}}}$ is the partition function for [CI]. $T_1 = 23.6$ K and $T_2 = 62.5$ K are the energies above the ground state.

### 3.4.2 The [CI] temperature and mass towards AMS12

The relationships defined above require the [CI] lines to be optically thin. We can test this requirement assuming a $T_{\text{ex}}$ and size (see equations A6 and A7 in the appendix of Schneider et al. 2003). Firstly, assuming $T_{\text{ex}} = T_G$ and a size of 0.8 kpc (i.e. from the LVG model) we infer line optical depths of $\sim 0.3$ for both the $[\text{CI}](^{3}\text{P}_1-^{3}\text{P}_0)$ and $[\text{CI}](^{3}\text{P}_2-^{3}\text{P}_1)$ lines. If we assume a larger size (i.e. 2 kpc), the optical depths are $\sim 0.05$.

From the line luminosities of the atomic carbon in AMS12 (see Table 2.1), we determine the line ratio $R_{\text{CI}} = 0.85 \pm 0.24$ and hence a [CI] excitation temperature of $42.7 \pm 10$ K. This yields a [CI] mass of $(1.54 \pm 0.49) \times 10^7 M_\odot$.

The [CI] excitation temperature determined does not fall within either $1\sigma$ temperature regions in Figure 3.2a. Thus when we fit the LVG model with no prior on the temperature, the $T_{\text{ex}}$ of [CI] is not in agreement with the gas temperatures indicated by high-excitation CO. It is in even less agreement with the gas kinetic
temperature we determine using the dust temperature as a prior. This could be either suggesting that the [CI] is not in LTE, or alternatively, the [CI] emission arises from a more spatially-extended, cooler molecular gas component than the gas probed by the higher-excitation CO lines.

The optical depth of the [CI] increases with decreasing temperature - while using $T_{\text{ex}} = 42.7 \pm 10$ K the line optical depths are calculated to be $\sim 1$ when assuming the compact size of 0.8 kpc. However, assuming the size is more extended, for example 2 kpc, the line optical depths drop to $\sim 0.1$. It is likely that our assumption of optically-thin lines is appropriate as even with the more compact size we are at the limit of optically-thin lines, and the possibility that the [CI] emission region is more extended would lower the line optical depths.

The LVG model analysis of the high-excitation gas has indicated this to be relatively compact with a radius $\sim 1$ kpc. The low-excitation temperature of [CI] which in other high redshift sources is broadly in agreement with the dust temperature \cite{Walter2011}, could be alluding to a second, cooler, more extended region of gas. Since we expect the CO(1-0) to trace the same region as the [CI] emission, if we were to observe CO(1-0) we could test if this low-excitation gas component exists in AMS12.

This has been seen in high redshift submillimetre galaxies (SMGs). High resolution imaging of CO(1-0) has revealed the CO(1-0) to be more extended than the higher $J$ emission (i.e. CO(3-2) or (4-3)), \cite{i2010,2011,2010,2011,Riechers2011b}. However, resolved CO(1-0) in strongly lensed quasar hosts has shown the CO(1-0) to be compact and similar to the CO(3-2) emission, justifying the assumption that the low-excitation gas has the same magnification factor as the high-excitation gas in these sources \cite{Riechers2006,2011a}.

\cite{Gerin2000} observed a relationship between $L_{\text{CO}(1-0)}^T$ and $L_{\text{[CI](1-0)}}^T$ from a survey of low redshift galaxies; $L_{\text{[CI](1-0)}}^T = 0.2 \pm 0.2 \times L_{\text{CO}(1-0)}^T$. \cite{Walter2011}
observed that their sample of (mostly lensed), high redshift sources support the relationship from Gerin & Phillips (2000). For the majority of their sources, CO(3-2) was the lowest transition observed, in order to get an estimate on CO(1-0) they used $L_{\text{CO}(1-0)}^T = 0.9 L_{\text{CO}(3-2)}^T$. From this, they determined that $L_{\text{[CI]}(1-0)}^T = 0.29 \pm 0.13 L_{\text{CO}(1-0)}^T$.

However, in the case of one SMG (SMM J163658+4105), which Ivison et al. (2011) observed in CO(1-0), the ratio of the brightness temperature luminosities of CO(3-2) and CO(1-0) is $0.54 \pm 0.12$, significantly lower than the assumed 0.9 that Walter et al. (2011) used. Using the CO(1-0) strength from Ivison et al. (2011), the ratio of $L_{\text{CO}(1-0)}^T/L_{\text{[CI]}(1-0)}^T$ then becomes 0.14, in disagreement with the high redshift relationship from Walter et al. (2011) and closer to the low redshift relationship. The ratio of $L_{\text{CO}(1-0)}^T/L_{\text{[CI]}(1-0)}^T$ at high redshift might well be similar to that at low redshift, once differential magnification of high- and low-J CO lines in lensed sources, has been taken into account (e.g. Deane et al. 2013).

3.5 Discussion

3.5.1 Gas mass

From the measured brightness temperature luminosity of the CO gas ($L_{\text{CO}}^T$), the H$_2$ mass can be found using the relation:

$$\left( \frac{M_{\text{H}_2}}{M_\odot} \right) = \left( \frac{\alpha}{M_\odot (\text{K km s}^{-1} \text{ pc}^2)^{-1}} \right) \left( \frac{L_{\text{CO}}^T}{\text{K km s}^{-1} \text{ pc}^2} \right) \quad (3.9)$$

(see Solomon & Vanden Bout 2005), where $M_{\text{H}_2}$ includes He and is therefore $\approx M_{\text{gas}}$. Downes & Solomon (1998) determined the constant $\alpha$ empirically from a study of a sample of ultra-luminous infrared galaxies (ULIRGs) and high redshift
galaxies to be $\alpha = 0.8$. We assume this value of $\alpha = 0.8$ hereafter, although we note that the acceptable values are in the range 0.3 - 1.3 (see Table 9 of Downes & Solomon 1998).

The brightness temperature luminosity of CO is best represented by $L_{\text{CO}(1-0)}^T$. This requires a measurement of CO(1-0) that we do not have, but we can calculate estimates of the CO(1-0) flux via various methods (A, B, C and D, detailed below). Three of these methods presented here assume that the molecular gas is represented by a single highly-excited component. We can also use the [CI]($^{3}P_{1} - ^{3}P_{0}$) line and the relationship from Gerin & Phillips (2000) to estimate the CO(1-0) brightness temperature luminosity without this assumption.

Method A

Assuming the transitions lower than CO(3-2) are thermalised, we can derive a total gas mass, $M(H_2)$, from $L_{\text{CO}(1-0)}^T$ using the relationship in Equation 3.9.

Under the assumption that CO(3-2) is thermalised, $L_{\text{CO}(3-2)}^T = L_{\text{CO}(1-0)}^T = (4.2 \pm 0.4) \times 10^{10}$ K km s$^{-1}$ pc$^2$. This gives a gas mass of $M(H_2) = (3.3 \pm 0.3) \times 10^{10}(\frac{\alpha}{0.8})M_\odot$.

Method B

Method B uses the best fit LVG model shown in Figure 3.3 (solid line, $T_G = 11.6$ K, $n(H_2) = 10^{6.1}$ cm$^{-3}$). The CO(1-0) line brightness temperature $T_b$ is given by the LVG model, and following the use of Equation 3.1 to convert to observed flux, we calculate $L_{\text{CO}(1-0)}^T = (4.8 \pm 0.8) \times 10^{10}$ K km s$^{-1}$ pc$^2$. The gas mass determined from Method B is $M(H_2) = (3.9 \pm 0.6) \times 10^{10}(\frac{\alpha}{0.8})M_\odot$. The uncertainty on the mass estimate from this method comes from the uncertainty on $L_{\text{CO}(1-0)}^T$ which is determined by the uncertainty on the model in Figure 3.3.
3.5 Discussion

Method C

We have used the dust information we have for this source and applied the prior distribution for the gas temperature $p(T_G) = p(T_D)$. The best fit model is shown in Figure 3.5 ($T_G = 89.6$ K, $n(H_2) = 10^{3.9}$ cm$^{-3}$). We can again use this LVG model’s CO(1-0) brightness temperature and convert it to a luminosity of $L_{CO(1-0)} = (3.2 \pm 0.3) \times 10^{10}$ K km s$^{-1}$ pc$^2$. This gives $M_{H_2} = (2.6 \pm 0.2) \times 10^{10} (\frac{\alpha}{0.8}) M_\odot$.

Method D

Using the relationship relating $L_{CO(1-0)}^T$ to $L_{[CI](1-0)}^T$ from [Gerin & Phillips (2000)], we estimate the $L_{CO(1-0)}^T = (6.1 \pm 6.1) \times 10^{10}$ K km s$^{-1}$ pc$^2$ (this large uncertainty comes from the uncertainty in the [Gerin & Phillips (2000)] relationship). We use the low redshift relationship as the results from [Walter et al. (2011)] are estimated from the higher CO(3-2) transition and not CO(1-0). The gas mass inferred from this CO(1-0) strength is $M_{H_2} = 4.9 \times 10^{10} (\frac{\alpha}{0.8}) M_\odot$. This is larger than the masses determined from the previous three methods, which is expected if the [CI] observations are revealing a low-excitation, more diffuse region of gas, such as the low-excitation gas reservoirs seen in the CO(1-0) observations of some SMGs ([Ivison et al. 2010, 2011; Carilli et al. 2010; Riechers et al. 2011b]).

The results are displayed in Table 3.1. We obtain the mean value of the methods A through to C (as they assume the high-excitation CO traces the total molecular gas), with the variance between the values yielding an estimate of the uncertainty. We find a mean value of $\langle M_{H_2} \rangle = (3.3 \pm 0.9) \times 10^{10} (\frac{\alpha}{0.8}) M_\odot$ from these three methods.

Studies of high redshift SMGs, show the resolved CO(1-0) evidently arise from extended, low-excitation gas (see for example [Ivison et al. 2010, 2011; Carilli et al. 2010] and references therein). [Ivison et al. (2011)] found that using CO(1-0) to determine the gas mass in four SMGs at $z \gtrsim 2$ gave masses $\sim 2$ times higher than
3.5 Discussion

<table>
<thead>
<tr>
<th>Method</th>
<th>$L^{T}_{\text{CO}(1-0)}/10^{10}$ [K kms$^{-1}$ pc$^2$]</th>
<th>$M_{\text{H}<em>2}/10^{10}$ [M$</em>\odot$]</th>
<th>$L^{T}<em>{<a href="1-0">\text{CI}</a>}/L^{T}</em>{\text{CO}(1-0)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$4.2 \pm 0.2$</td>
<td>$3.3 \pm 0.3$</td>
<td>$0.29$</td>
</tr>
<tr>
<td>B</td>
<td>$4.8 \pm 0.8$</td>
<td>$3.9 \pm 0.6$</td>
<td>$0.25$</td>
</tr>
<tr>
<td>C</td>
<td>$3.2 \pm 0.3$</td>
<td>$2.6 \pm 0.2$</td>
<td>$0.37$</td>
</tr>
<tr>
<td>D</td>
<td>$6.1 \pm 6.1$</td>
<td>$4.9 \pm 4.8$</td>
<td>$0.20^\ast$</td>
</tr>
</tbody>
</table>

Table 3.1: The calculated CO(1-0) line strength determined four ways. A: from the assumption that CO(3-2) is thermalised; B: from the unconstrained best fitting LVG model; C: from the LVG model assuming $p(T_D) = p(T_G)$; and D: from the $[\text{CI}](^3P_1-^3P_0)$ strength using the Gerin & Phillips (2000) relationship. The $[\text{CI}]/\text{CO}$ relationships for each CO(1-0) are determined and displayed. * This is the Gerin & Phillips (2000) relationship that we have assumed to derive the CO(1-0) mass which is responsible for the high uncertainty.

masses determined from the CO(3-2) or higher $J$ lines, i.e. the assumption that the higher $J$ transition (CO(3-2) or CO(4-3) in most cases), is thermalised was incorrect.

Single component LVG models of SMGs have been shown to underestimate the CO(1-0) line (i.e. Carilli et al., 2010; Riechers et al., 2011b). Direct observations of the CO(1-0) line in AMS12 are needed to test whether a significant component of low-excitation gas is present.

We shall use the gas mass determined from the CO(1-0) strength estimated from the $[\text{CI}](^3P_1-^3P_0)$ line, $M_{\text{H}_2} = 4.9 \times 10^{10} (\alpha/0.8) M_\odot$, although we acknowledge there is considerable uncertainty in this estimate due to the large scatter in the $L^{T}_{\text{CO} (1-0)}/L^{T}_{[\text{CI}](1-0)}$ relationship.

3.5.2 Discriminating between LVG models

We discuss various methods we use to discriminate between the LVG models. First, we consider the models’ outputs and the constraints future observations can make. Secondly, by estimating the mass given by the models assuming the gas is distributed in a thin disc, we compare these with the gas mass from Section 3.5.1. Then we explore the possibility of a two component LVG model inspired by the
excitation temperature of [CI].

From the LVG models’ output line brightness temperatures

The sharp decline in the intensity of the higher-\(J\) transitions in the low-temperature/high-density model can be used with future observations to discriminate between the models. Using the output brightness temperatures we can predict the possible line strengths of the higher \(J\) transitions, particularly using the CO(8-7) and CO(9-8) transitions which happen to be in observable windows. The flux densities would be \(S_{\text{CO}(8-7)} = 1.13\) mJy and \(S_{\text{CO}(9-8)} = 0.028\) mJy respectively (see Figure 3.3). However, if we instead use the brightness temperatures from the solution where we used the dust temperature as a prior for the gas temperature the line strengths are \(S_{\text{CO}(8-7)} = 5.9\) mJy and \(S_{\text{CO}(9-8)} = 3.1\) mJy respectively (see Figure 3.5).

The significant differences in these line strengths mean that observations of higher CO transitions in this object would conclusively constrain the region of parameter space of the LVG model which describes the conditions of the gas.

From mass estimates using volume arguments

Another way we may be able to rule out one of the models is by considering the spatial distribution of the gas. Consider the LVG model which gives the lowest \(\chi^2\) value, \(T_G = 11.6\) K, \(n(H_2) = 10^{6.1}\) cm\(^{-3}\) and \(r_0 = 6.2\) kpc, and assume the molecular gas is distributed in a thin disc with height \(H\), of uniform density. We can calculate the volume of the gas given the density. We use the findings of Downes & Solomon (1998) that the molecular gas in the central regions of ULIRGs is not a collection of separate clouds undergoing self-gravitation, but rather clouds fused together to form a disc of more or less constant density. They modelled the structure of these gas discs and found the average height of the discs to be \(H \sim 58\) pc. Taking this value
as the height of the assumed disc of gas for AMS12, we can calculate the volume of the disc given \( r_0 = 6.2 \) kpc.

From this volume, and the best fitting density of the gas \( n(\text{H}_2) = 10^{6.1} \) cm\(^{-3}\), we can estimate the gas mass in this volume to be \( \sim 4.1 \times 10^{14} \) M\(_{\odot}\). This is four orders of magnitude higher than the gas mass derived from our observations (see Section 3.5.1). Table 3.1 shows the gas mass \( M_{\text{H}_2} \), for the four methods described in Section 3.5.1.

On the otherhand, the LVG model giving the parameters \( T_G = 89.6 \) K, \( n(\text{H}_2) = 10^{3.9} \) cm\(^{-3}\) and \( r_0 = 0.8 \) kpc, gives a mass of \( 4.6 \times 10^{10} \) M\(_{\odot}\) assuming the same disc height of 58 pc. This is very similar to the masses in Table 3.1 determined by Equation 3.9.

Though we note we have made various assumptions to estimate the gas masses in Section 3.5.1, this crude mass estimate from the volume argument, distinguishes the LVG model solutions from each other. The estimate from the volume of the low-temperature/high-density solution does not agree with the gas masses determined in Section 3.5.1, while the mass from the solution using the dust temperature as a prior does agree.

Two-component model

Riechers et al. (2011b) found in the two SMGs they studied that the CO ladders were best fit with two component LVG models: a dense, high-excitation component to fit the higher-\( J \) CO lines, and another low-excitation component to fit the low-\( J \) CO measurements.

To investigate the possibility that the [CI] gas in AMS12 is from a lower-temperature, more diffuse region, we fit a second component to the LVG model. We used the CO(1-0) estimate from the [CI](\(^3P_1 - \(^3P_0\)) line, and for the high-excitation
3.5 Discussion

component we keep the previous LVG model solution using the dust prior (from Figure 3.5).

We fix the temperature of the second component at 42.7 K (the [CI] excitation temperature), and tie both components to the estimate of the CO(1-0) strength. Since we only have one data point, we cannot fit for both the size and the density at the same time, so we fit each independently keeping the other fixed at a fiducial value. Firstly, we keep the density fixed to $10^{2.5} \text{ cm}^{-3}$ and fit for the size. This density is chosen as a representative of a diffuse component and corresponds to the critical density of CO(1-0). The best fit to the radius is 2.8 kpc and the resulting LVG model is shown in Figure 3.6a.

Next, we fix the size of the low-excitation gas, while continuing to hold the temperature at 42.7 K, and fit for the density. We choose two arbitrary sizes; 1 kpc, which is similar to the best fitting size of the single component LVG model solution to the high-$J$ CO lines, and 6 kpc, which is significantly more extended. For the second component solution with a fixed size of 1 kpc, the best fitting density is $n(\text{H}_2) = 10^{3.4} \text{ cm}^{-3}$. The two component LVG model (comprising of this solution and the best fitting solution of the high-$J$ CO lines with the dust temperature prior applied), is shown as the solid line in Figure 3.6b. It does not provide a good fit to the observations.

With the second component size fixed at 6 kpc, (and temperature remaining at 42.7 K), the best fitting density is $n(\text{H}_2) = 10^{2.1} \text{ cm}^{-3}$. The corresponding two component LVG model is shown in Figure 3.6c. This second component solution provides a better fit to the observed CO lines. The estimates shown in Figures 3.6a and 3.6c are consistent with a more diffuse and extended gas which is only detectable via the low-$J$ CO lines and the [CI] lines, while still providing good fits to the high-$J$ observations.

We note that by construction, our method will only select solutions with weak
Figure 3.6: Two component LVG models fitted to the CO observations and an estimated CO(1-0) line strength of 0.57 mJy. The low-excitation component is given by the dashed line, while the high-excitation component (from Figure 3.5) is given by the dot-dot-dash line. The combined model is given by the solid line. (a): The low-excitation component with $T_G = 42.7$ K, $n(H_2) = 10^{2.5}$ cm$^{-3}$ and size of $r_0 = 2.8$ kpc. (b): The low-excitation component in this figure has the solution $T_G = 42.7$ K, $n(H_2) = 10^{2.4}$ cm$^{-3}$ and a fixed size of $r_0 = 1$ kpc. (c): The low-excitation component here has $T_G = 42.7$ K, $n(H_2) = 10^{2.1}$ cm$^{-3}$ and size of $r_0 = 6$ kpc.
emission of the high-$J$ lines. Since we have begun by fitting CO(3-2), CO(5-4) and CO(7-6) with a single component model, the extra component used to fit CO(1-0) must produce negligible flux in the CO(3-2), CO(5-4) and CO(7-6) lines, otherwise their predicted fluxes will be higher than those observed.

### 3.5.3 Dynamical mass

We can estimate the dynamical mass of the system from the FWHM of the CO lines, assuming a characteristic radius via (Neri et al., 2003);

$$\left( \frac{M_{\text{dyn}} \sin^2 i}{M_\odot} \right) = 4 \times 10^4 \left( \frac{\Delta V_{\text{FWHM}}}{\text{kms}^{-1}} \right)^2 \left( \frac{r}{\text{kpc}} \right)$$

Studies have shown the virial mass estimate is reasonable even if the gas is clumpy (see Daddi et al., 2010).
We can use the CO linewidths to estimate the dynamical mass; however, if the [CI] traces a low-excitation region of gas, we should use the [CI] linewidths. Given the large uncertainties in the linewidths (Table 2.1), the CO and [CI] linewidths are very similar and agree within 2σ. Using both the [CI] linewidths, we use $\Delta V_{\text{FWHM}} \approx 260$ km s$^{-1}$. This will give an approximate estimate on the dynamical mass. Assuming two arbitrary radii encompassing a range of sizes, i.e. the radii from Section 3.5.2, $r_1 = 1$ kpc and $r_2 = 6$ kpc, we can estimate two values of the dynamical mass of AMS12. Using $r_1$, the dynamical mass is $M_{\text{dyn}} \sin^2 i = 2.7 \times 10^9$ M$_\odot$, while using $r_2$ gives $M_{\text{dyn}} \sin^2 i = 1.6 \times 10^{10}$ M$_\odot$.

The dynamics are dominated by the molecular gas and stars, with the dark matter and ionized hydrogen sub-dominant \cite{Daddi2010}. The stellar mass in AMS12 is found from the mid-/near-infrared SED by \cite{Lacy2011} to be $M_* \approx 3 \times 10^{11}$ M$_\odot$. The dynamical mass estimates make it difficult to account for the stellar mass. It may be that the radius estimates we have used are not adequate to encompass all the stellar mass. The constraints on the inclination angle assuming the radii above are severe when considering the stellar and gas mass, for example, if $r_1 = 1$ kpc, then the inclination angle $i \approx 5^\circ$. While using the radius estimate $r_2 = 6$ kpc, the constraint on the inclination angle relaxes slightly to $i \approx 13^\circ$. These arguments suggest the host galaxy is seen face-on.

The radio spectrum of AMS12 has been investigated in \cite{Martinez-Sansigre2006a} and \cite{Kloekner2009}. This object has a steep extended radio spectrum, and narrow optical emission lines pointing to torus obscuration, i.e. the orientation of the central engine and its obscuring material is closer to edge-on to the observer. The low inclination angle of the host from the dynamical mass estimates suggest the central AGN region and the galaxy’s stellar, gas and dust regions are not aligned. Together with the observed narrow emission lines, this suggests AMS12 is obscured by the torus and not by dust in the host galaxy \cite{Martinez-Sansigre2006a}. 


3.5 Discussion

The caveats to using this dynamical mass estimate are many. The high-excitation CO lines may be tracing a separate gas component than the [CI] lines, and therefore the radius estimates based on the high-excitation CO LVG modelling are tenuous for the possible low-excitation component. Though the [CI] linewidths are similar to the CO linewidths, this does not immediately place the [CI] at the same region of the high-excitation CO, i.e. if [CI] is alluding to a more massive while more extended low-excitation gas reservoir. Resolved CO(1-0) in this object would significantly improve the dynamical mass estimate in this object.

3.5.4 [CI] abundance and cooling contribution in AMS12

The abundance of [CI] in AMS12 can be determined by $X_{\text{[CI]}}/X_{\text{H}_2} = M_{\text{CI}}/(6 M_{\text{H}_2})$, where $M_{\text{CI}} \sim 1.5 \times 10^7 M_\odot$. The abundance, assuming the high-excitation gas can be used to estimate the gas mass, is $7.8 \times 10^{-5}$. While if there is a lower temperature, more diffuse component of the gas, the $M_{\text{H}_2}$ would be higher. The $M_{\text{H}_2}$ estimated from the [CI]($^3P_1 - ^3P_0$) line is $\approx 5 \times 10^{10} M_\odot$, giving a [CI] abundance of $X_{\text{[CI]}}/X_{\text{H}_2} = 5.2 \times 10^{-5}$. This indicates the molecular gas is already enriched at this redshift, supporting findings from Walter et al. (2011).

The ratio of $L_{\text{[CI](1-0)}}/L_{\text{FIR}}$ provides a measure of the cooling contribution of [CI]. For AMS12, $L_{\text{[CI](1-0)}}/L_{\text{FIR}} = 1.5 \times 10^{-6}$, which appears to be typical for the quasar sources of Walter et al. (2011). Their sources are split into quasars and SMGs, and while quasars have $L_{\text{[CI](1-0)}}/L_{\text{FIR}}$ ratios similar to that of AMS12, the SMGs ratios are around an order of magnitude higher. This could be due to the AGN contribution to the $L_{\text{FIR}}$ in the quasars, as discussed earlier. Overall, the [CI] lines are not major coolants, in fact they are negligible compared to the cooling by the dust continuum.
3.5.5 AGN bolometric luminosity and the scale of AGN heating

From our current observations of the dust and high-excitation gas in AMS12 the temperatures we determine are shown to be beyond the capabilities of heating by star formation alone. Empirical observational evidence and radiative models of star-forming galaxies show that the dust temperatures reach $\approx 50$ K (see for example Kovács et al. [2006]; Siebenmorgen & Krügel [2007]). The typical dust temperatures derived from FIR SED fitting in studies of high redshift star-forming galaxies, range from 30-60 K with the average being $\sim 35$ K (e.g. Kovács et al. [2006]; Coppin et al. [2008]; Elbaz et al. [2011]). These temperatures are typical of local galaxies where heating of the dust is dominated by young stars (e.g. Farrah et al. [2003]; Elbaz et al. [2011]).

However, in the hosts of luminous AGN, the FIR emission could also be heated by the AGN. APM 08729+5255, F10214, BR 1202-0725 and Cloverleaf for example, have hotter FIR dust than typical star-forming galaxies and other SMGs. Detailed studies of the dust and gas in these objects have revealed the dust is compact which supports the possibility of significant AGN heating (Solomon et al. [2003]; Riechers et al. [2006]; Weiß et al. [2007]; Ao et al. [2008]).

We can estimate the scale of the heating from the AGN using the bolometric luminosity of AMS12 from the broad-band data between 3.6 and 24 $\mu$m (as done by Martínez-Sansigre et al. [2009]). For AMS12, an AGN of $L_{\text{bol}} = 2 \times 10^{13} L_\odot$, we assume that $L_{\text{UV}}$ is $\sim 0.25L_{\text{bol}}$, and using the FIR SED fitted parameters ($T_D = 88$ K and $\beta = 0.6$), we find the scale of AGN heated dust out to 88 K is 2.7 kpc. This result is obtained assuming that the ultraviolet photons travel unhindered up until the radius where the dust becomes self-shielding (i.e. at large optical depths), thus 2.7 kpc may be thought of as the characteristic radius to which the dust is heated.
to this temperature by the AGN (see Barvainis 1987 for details). Clearly, since the characteristic scale of the dust is indeed found to be $\sim 2$ kpc in many objects (Greve et al. 2005; Tacconi et al. 2006; Younger et al. 2008), the dust temperature observed in AMS12, $T_D = 88$ K, could be achieved through heating from the AGN.

Note, this estimate of the bolometric luminosity does not take the $L_{\text{FIR}}$ into account, and therefore may be considered a lower limit. If we account for the $L_{\text{FIR}}$ which we believe is also attributed to the AGN ($3 \times 10^{13} L_\odot$), our new estimate of the $L_{\text{bol}}$ is $5 \times 10^{13} L_\odot$ for AMS12, which could heat the dust to 88 K out to around $\sim 4$ kpc.

### 3.5.6 Black hole, stellar, gas and dust masses

We estimate the dust mass to be between $9.2 \times 10^7 M_\odot$ and $1.6 \times 10^9 M_\odot$ using the $L_{\text{FIR}}$ and two different mass absorption coefficients (this range is typical for high redshift sources, see for example Solomon & Vanden Bout 2005). The dust masses vary greatly due to the process of extrapolating the mass absorption coefficients to the rest frequencies, a power law which depends upon $\beta$. The gas mass from the [CI] observations is determined to be $4.9 \times 10^{10} M_\odot$. This yields a gas-to-dust mass ratio between $\sim 30 - 530$.

We can estimate the Eddington limited black hole mass of this system given the $L_{\text{bol}}$. If accreting at $\lesssim 100\%$ of the Eddington rate (reasonable for quasars at high redshifts, see McLure & Dunlop 2004), the black hole mass estimated from the $L_{\text{bol}}$ determined from the mid-infrared SED is $M_\bullet \gtrsim 6 \times 10^8 M_\odot$. The revised $L_{\text{bol}}$ (including the $L_{\text{FIR}}$), leads to an estimate of the black hole mass $M_\bullet \gtrsim 1.5 \times 10^9 M_\odot$. We have assumed that all of the $L_{\text{FIR}}$ is attributed to heating from the AGN; though we acknowledge that star formation is likely to contribute, we do not have a measure of the extent of this contribution.
The stellar mass of AMS12 has been estimated by Lacy et al.\cite{2011} to be $M_\star = (3.2 \pm 0.3) \times 10^{11}\, M_\odot$. We assume this stellar mass is located entirely within the bulge as AMS12 is probably the progenitor of a present-day elliptical galaxy. This means AMS12 has already the stellar mass of a present-day 2L* galaxy.

Given the gas mass of $M_{\text{H}_2} = (4.9 \pm 0.9) \times 10^{10}\, L_\odot$ we derive from the CO observations, were this to be converted into stars with 100% efficiency, it would still only increase the stellar mass by $\sim 15\%$. Observations of CO(1-0) are needed to probe the lower-excitation gas and give a more accurate value of the gas mass. However, it seems unlikely that the stellar mass will increase significantly, unless we have underestimated the gas mass by a factor of $\sim 10$.

Assuming $M_\star = M_{\text{bulge}}$, we have a $M_\star/M_{\text{bulge}}$ ratio for AMS12 of $\gtrsim 0.005$. This is significantly higher than the relationship determined from nearby ($z \sim 0$) galaxies by Marconi \& Hunt\cite{2003} and H"aring \& Rix\cite{2004} where $M_\star/M_{\text{bulge}} \sim 0.002$ with a scatter of $\sim 0.3$ dex.

The $M_\star - M_{\text{bulge}}$ relationship has been investigated at higher redshifts up to $z \sim 4$. For example, McLure et al.\cite{2006} investigated the relationship in radio-loud galaxies at $z \sim 2$, Decarli et al.\cite{2010} studied a sample of 96 quasars out to $z \sim 3$, Peng et al.\cite{2006a,2006b} used both gravitationally lensed and non-lensed galaxies to study the relationship out to $z \sim 4.5$, while Targett et al.\cite{2011} studied $z \sim 4$ quasars. The high redshift studies all agree that the $M_\star/M_{\text{bulge}}$ relationship appears to be evolving with redshift. This evolution was seen to be independent of radio-loudness and quasar luminosity (see Decarli et al.\cite{2010}, who studied both radio-loud and radio-quiet quasars, as well as investigated possible biases). The results of these studies agree with one another and imply that the black holes at high redshifts are more massive for a given bulge mass than their local counterparts.

While our results have considerable uncertainty, they are consistent (within the scatter) with what these other groups have found (Peng et al.\cite{2006b}; McLure et al.\cite{2006})
For AMS12 to evolve to the local relationship, the bulge would have to grow \( \sim 3 \) times as much as the central black hole from \( z = 2.8 \) to \( z = 0 \).

Given the amount of molecular gas implied by the CO observations, if AMS12 were to evolve secularly, it would, at most, only increase its bulge mass by \( \sim 15\% \). Mergers could add more gas for star formation while adding stellar and black hole mass. It is expected that massive galaxies \( (M_\bullet \gtrsim 10^{11}M_\odot) \), undergo > 1 mergers from \( z \sim 3 \) to present day (Conselice et al. 2003, 2007; Bluck et al. 2009; Hopkins et al. 2010; Robaina et al. 2010). However, it is not clear how mergers affect the black hole mass and whether it is possible to achieve the necessary growth of the bulge relative to the growth of the black hole.

Alternatively, Decarli et al. (2010) addressed the possibility that the remnants of the high redshift quasars they studied are high-mass outliers to the local relationship. These high redshift quasars are progenitors to present day massive ellipticals, and keeping their \( M_\bullet/M_{\text{bulge}} \) value to \( z = 0 \), they become outliers, rather than evolve, to the local relationship.

We must note that there was significant bias towards selecting a powerful quasar, giving a large \( M_\bullet \), while demanding a faint 3.6 \( \mu \)m flux, limiting the host galaxy’s luminosity (Martínez-Sansigre et al. 2005, 2006b). In addition, the search for CO in this object was initiated by selecting the brightest MAMBO detection from the obscured quasar sample (Martínez-Sansigre et al. 2009).

There is also a possibility that by biasing ourselves towards such a high \( L_{\text{bol}} \), AMS12 is super-Eddington, in which case we would be overestimating the black hole mass. If AMS12 were accreting at super-Eddington rates, the black hole mass would be overestimated by the amount by which the bolometric luminosity exceeds the Eddington-limited luminosity.
3.5.7 Comparison to other galaxies

AMS12 is the first unlensed, high redshift source detected in both [CI] lines. This eliminates any ambiguity on the effects of possible differential magnification from lensing.

The characterisation of the gas and dust in AMS12 is consistent with the observations of other high redshift galaxies. This includes the value of $\log(L_{\text{FIR}}/L_{\text{CO}}^T)$, known as the star formation efficiency, see Figure 8 in Solomon & Vanden Bout (2005), where AMS12 lies within the scatter of the high redshift galaxies. The gas masses in these sources are similar to AMS12 (i.e. studies of SMGs, ULIRGs and quasars at high redshifts; Solomon & Vanden Bout 2005; Greve et al. 2005; Riechers et al. 2006; Coppin et al. 2008; Yan et al. 2010; Lacy et al. 2011).

A study of CO(1-0) in two $z\sim 2.8$ obscured quasar hosts found similar gas and dust masses to AMS12 (Lacy et al. 2011). The studies of obscured quasars and their hosts also indicate mature systems, with dust and gas masses low compared to the stellar mass estimates (e.g. Lacy et al. 2011).

However, comparing to other high redshift quasars which are strongly lensed is difficult, as is illustrated with F10214. The gas and dust properties of F10214 are very similar to AMS12, with approximately equivalent [CI] abundances, cooling rate, and line ratios in terms of both CO and [CI]. Ao et al. (2008) modelled the CO emission in F10214 with LVG models, and found a similar dust temperature to AMS12 (80 K), and determined a range of gas kinetic temperatures from 45-80 K (note their Figure 7 has a similar shape to the high temperature region in Figure 3.2), with $T_{\text{ex}}$ of [CI] $\sim 42$ K. Riechers et al. (2011a) detected CO(1-0) in F10214, and found there was no evidence for an extended, low-excitation gas component. For their analysis they assumed a constant magnification factor for the CO(1-0) (the magnification given by the higher-$J$ CO lines), hence if there is differential
magnification of the gas components, their results may be affected.

[Deane et al. (2013)] have revised the lens model in F10214 and have indeed found differential magnification on frequency and spatial scales. Deane et al. (2013, in prep.), study resolved CO(1-0) in this object and find preferential magnification between individual channels and predict distortion of the CO SED. Thus, AMS12 which is unlensed, offers, so far, a unique opportunity to study the gas and dust in an obscured quasar host without the added complication of gravitational lensing.

### 3.6 Conclusion and future work

We have presented new observations of the obscured quasar AMS12 and, along with previous mm observations, we have investigated the gas properties of this object.

Modelling of the FIR dust observations in Appendix A have shown that the FIR SED is well fit by a single component graybody model with dust temperature \( T_D = 88 \) K, emissivity index \( \beta = 0.6 \), and \( L_{\text{FIR}} = 3.16 \times 10^{13} \) \( L_\odot \), implying heating by the AGN.

The CO SED was fit with LVG models, and we used the marginalised PDF of the dust temperature as the prior distribution of the gas kinetic temperature to constrain the parameters. This yielded the gas kinetic temperature of \( T_G = 89.6 \) K, and density \( n(H_2) = 10^{3.9} \) cm\(^{-3}\), suggesting that SF is not the sole heating source.

The atomic carbon fine structure lines \([\text{CI}](^3P_1 - ^3P_0)\) and \([\text{CI}](^3P_2 - ^3P_1)\) were observed and the \([\text{CI}]\) excitation temperature was determined to be \( 42.7 \pm 10 \) K. This is significantly lower than \( T_G \), indicating either \([\text{CI}]\) is not in LTE, or it is from a more extended, lower temperature gas component.

The gas mass determined from the CO(1-0) estimate, is \( \sim 4.9 \times 10^{10} \) \( M_\odot \). The
3.6 Conclusion and future work

dynamical mass was calculated from the CO linewidth to be $M_{\text{dyn}} \sin^2\theta = (2.7 \pm 0.2) \times 10^9 \, M_\odot$ assuming $r = 0.8 \, \text{kpc}$, giving a limit to the host galaxy’s inclination $i \lesssim 13^\circ$.

The stellar mass in this object is estimated at $M_* = 3.2 \times 10^{11} \, M_\odot$. It follows that the gas and dust mass are only a fraction of the current stellar mass. The $M_*/M_\star$ ratio is $\gtrsim 0.005$, which is higher than in the local Universe.

The system has already amassed the majority of its stellar mass and is host to a massive black hole, indicating a mature system. It is not clear how the system will evolve to the present-day $M_*/M_\star$ relation, or whether the extreme value is due to a selection bias. Future observations that would help with our characterisation of this object include a detection, or even mapping of the CO(1-0). Mapping the low-excitation CO would reveal whether this system consists of multiple components; a diffuse, low-excitation gas component with a more concentrated, high-excitation component from which arises the strong high-excitation CO lines. Observing even higher CO transitions, such as CO(8-7), CO(9-8) and above could constrain the high-excitation end of the CO ladder. Also possible is the revelation of strong, secondary peaks in the CO ladder at these transitions reflecting an X-ray dominated region leading to the excitation of these lines.
Part II

Radio-quiet quasars
Chapter 4

Searching for a radio-quiet lower envelope

In this chapter we investigate the existence of a lower radio envelope in the “radio loudness” property of QSOs. We use QSOs selected in a narrow redshift range from the Sloan Digital Sky Survey (SDSS) and match them with the NRAO Very Large Array Sky Survey (NVSS). We select the QSOs that are undetected in NVSS and stack these in samples of different optical luminosities to search for evidence of a lower envelope, i.e. a minimum radio power for a given accretion rate.

4.1 Introduction

Quasars were first discovered via their radio emission and it has remained an important diagnostic tool in their classification. The first radio surveys revealed strong radio emission from sources and when matched with optical photometry the optical counterparts appeared as point-like objects (e.g. Matthews & Sandage, 1963). Before tools to measure their distance were implemented these point-like objects seemed indistinguishable from stars except for their radio emission. It was only when spectra were taken of these objects that the non-stellar nature and the cosmological distances of these objects was realised (Schmidt, 1963) and that there were many more optical point sources that were not associated with radio emission.
4.1 Introduction

(i.e. QSOs Sandage, 1965).

4.1.1 Radio loudness parameter

The radio loudness of quasars is either defined by a radio loudness parameter which is the ratio of the radio to optical flux densities (or luminosity densities) ($R = L_{\text{rad}}/L_{\text{opt}}$) such as in Kellermann et al. (1989), or as Miller, Peacock & Mead (1990) defined it, by the radio luminosity. The shape of the radio loudness distribution has alternatively been found to be bimodal (the $R$ values of quasars appear to cluster around two values) or a smoother distribution. Whether bimodality is a real effect or a selection effect is an area that is still under investigation (e.g. Cirasuolo et al. 2003b; Ivezic et al. 2002; Sikora, Stawarz & Lasota, 2007; Broderick & Fender, 2011; Baloković et al. 2012; Mahony et al. 2012). If the distribution in radio loudness is an intrinsic property of QSOs, then the theories of the production of radio jets should be able to reproduce the distribution. Bimodality implies two separate populations, one radio-loud (RLQ) and one radio-quiet (RQQ) possibly powered by different processes or at least representing two distinct modes of the same process.

However, obtaining the distribution of $R$ is complicated by the processes involved in matching two unrelated observations in different wavelengths, optical and radio. There are always more faint sources than bright ones in any wavelength, so if a population is selected in one wavelength the majority of these sources will lie close to the survey’s detection limit. If the sources detected in this wavelength are then compared to observations in another wavelength, the chance of detecting an object in both wavelengths is smaller than the object being detected in either wavelength. For example, the optically selected QSOs are dominated by “radio-quiet” sources, or more precisely the majority of optically selected sources are undetected in the cross-matched radio survey (e.g. Kellermann et al., 1989; Miller, Peacock & Mead, 1990).
4.1 Introduction

Cirasuolo et al. [2003b] among others). This effect is even more pronounced when one of the surveys is relatively shallow compared to the survey in the wavelength the sources were selected from. The term ‘shallow’ refers to the sensitivity limit reached in the survey, i.e. given a restriction on the time a survey is allocated to observe a field (integration time), the survey can either cover a large area spending less time on each observation, or longer integration times on a smaller area of the sky (‘deeper’ observations). The choice between a shallow, wide-field survey or a deeper, smaller area survey depends on the targets of the survey; brighter objects are rarer, and so require a search on a larger area to accumulate many of them, while the fainter objects are more common but they require longer integration times (greater sensitivity) to detect them, so a deeper survey is required.

The radio loudness parameter requires a luminosity measurement from two wavelengths - radio and optical. Since there are more faint objects there will be more (detected) objects close to the optical survey’s detection limit. These are matched with the radio survey where most sources are close to the flux limit. Thus, an artificial correlation is introduced; the majority of doubly-detected sources will have an $R$ value that is determined by the flux limits of the surveys. Therefore it is important to consider the sources which are detected in one wavelength but not in the other. Since these sources are not detected simultaneously in both bands, we need to find a way of estimating their luminosities in the wavelength they are not detected. In our case, we have used the Sloan Digital Sky Survey (SDSS) and selected QSOs from their optical colours and cross-matched their positions with the NRAO VLA Sky Survey (NVSS). We are interested in the radio undetected QSOs, which we will stack and measure their flux densities from the stacked image.

Observations of the radio-loud and radio-quiet dichotomy hinted at host morphological differences between the two populations (e.g. Sikora, Stawarz & Lasota [2007]; Volonteri, Sikora & Lasota [2007]); radio-loud quasars predominantly were at
the centre of very large elliptical galaxies with the most luminous associated with largest galaxies, while the radio-quiet objects were preferentially associated with spiral galaxies. \cite{Wilson & Colbert 1995} suggested that the radio loudness dichotomy could be explained by the difference in the spin of the black holes, the radio-loud objects having fast spinning black holes which they assume had been spun up only by recent major mergers and not by ordinary accretion, hence the difference in galaxy morphology. They argued that the rarity of an object being radio loud reflects the rarity of major mergers leading to a “bottom heavy” distribution of spins. However this distribution of spins is also possible if the evolution of black hole spin is dominated by accretion episodes (e.g. \cite{Moderski, Sikora & Lasota 1998}, and references within). The notion that the radio dichotomy was related to the morphology of the host was confronted by Hubble Space Telescope imaging of the hosts of luminous QSOs which found the hosts to be predominantly bulge-dominated galaxies (e.g. elliptical galaxies) regardless of the radio luminosity (see \cite{Floyd et al. 2004}, for example).

A further study which looked at the radio loudness parameter with respect to the Eddington ratio (where the Eddington ratio is defined as the ratio of the bolometric luminosity with the Eddington luminosity $L_{bol}/L_{Edd}$) was done for a collection of AGN samples including radio-loud AGN, Seyferts, radio galaxies and optically selected quasars that had black hole mass information \cite{Sikora, Stawarz & Lasota 2007}. They found that the AGNs followed two distinct tracks in the radio loudness - Eddington ratio plane. For the same Eddington ratio the objects in their sample followed two distinct groups. The gap between these may have been due to selection effects since their sample of AGN was incomplete; nevertheless the difference in radio loudness remained prominent. The upper sequence in their sample consisted of the objects with high black hole masses. The mass of the black hole is related to the growth history (i.e. mergers and/or accretion episodes), hence also the host galaxy morphology and the black hole spin. Therefore \cite{Sikora, Stawarz & Lasota 2007}
proposed a revised spin paradigm which takes into account the growth history of the black hole (accretion episodes and mergers) and suggests that the mode of accretion also plays a significant role in the radio power. [Volonteri, Sikora & Lasota (2007)] investigated the morphology related bimodality of the spin distribution, and determined that mergers contribute to the most massive black holes and that these have the highest spins; they are also observed to have the highest radio luminosities. Black holes hosted by spirals are believed to grow predominantly by accretion episodes which result in typically lower spins (see the conclusions in Stawarz, Sikora & Lasota (2008) for a succinct presentation of the revised spin paradigm).

A recent investigation of the sample of AGN that Sikora, Stawarz & Lasota (2007) studied, using the core radio luminosities was carried out by Broderick & Fender (2011). They found that there was a difference between using the total radio luminosity (which includes core and extended emission) and just the core radio luminosities in the radio loudness - Eddington ratio plane. In addition to the core luminosities they applied a mass correction term to the radio luminosity to account for any mass dependence. With the mass correction applied, the gap between the two tracks of radio loudness with Eddington ratio that Sikora, Stawarz & Lasota (2007) found was significantly diminished. Therefore they suggest that the radio loudness dichotomy seen in the total radio luminosity comparison (e.g. Sikora, Stawarz & Lasota 2007) could be due to several additional parameters to black hole spin such as the environment, the age of the radio source and/or the dependence on the mass of the black hole. Therefore, while in these scenarios the black hole spin is an important factor in determining the jet power, the appearance of a strong radio jet is believed to depend on a number of factors such as the host morphology, growth history of the black hole, and in turn the environment (e.g. mergers are more likely to occur in dense environments).

Bearing these factors in mind, there is still the requirement for the jet production
mechanism to be able to produce a wide variation of jet power given a certain accretion rate and host galaxy morphology. One such mechanism is that the radio jets in quasars are produced by magnetic outflows driven by differentially rotating magnetic field lines in the accretion disc or near the black hole. The rotation of the field lines may be due to the angular momentum of the accretion disc or the spinning black hole (e.g. Meier, 2002). Theoretically the spin of the black hole can reproduce the observations through the Blandford-Znajek mechanism extracting energy from the angular momentum of the black hole (Blandford & Znajek, 1977). However, to reproduce the most powerful radio jets observed requires a larger power of the black hole angular momentum than the Blandford-Znajeck mechanism (e.g. Tchekhovskoy, Narayan & McKinney, 2010). With the improvements of magneto-hydrodynamic (MHD) simulations in the relativistic framework, simulations of jet powers have provided promising results in modelling the observations (Lagos, Padilla & Cora, 2009; Krolik & Hawley, 2010; Tchekhovskoy, Narayan & McKinney, 2010; Fanidakis et al., 2011). The recent work by Martinez-Sansigre & Rawlings (2011) constructed different spin distributions given different models of the jet efficiencies; they found the spin distributions to be bimodal, with a population of sources at near maximal spin, and a population at low spins. Assuming that jet power scales with spin, the observed radio power should appear bimodal also, although this may be affected by selection effects.

The simulations that take the black hole spin into account are able to reproduce the radio-loud end of the QSO population well, producing jet efficiencies high enough to account for the observations of the extreme end of radio loudness (Fernandes et al., 2011; McNamara, Rohanizadegan & Nulsen, 2011). These observations show that the most powerful radio-loud quasars produce jets which probably require most, if not all of the available energy from the accretion rate to be converted to power the jet. Hence, a radio-loud upper envelope has been observed and theoretically investigated for quasars which shows a maximum jet efficiency given an accretion rate (e.g.
Observationally, there have been hints of a corresponding radio-quiet counterpart to this upper envelope such as in the optically selected sample of QSOs matched with the FIRST survey of Cirasuolo et al. (2003b). However, Cirasuolo et al. (2003b) point out that the low radio luminosity data is incomplete and the hence the apparent correlation between the low luminosity radio and optical luminosities (i.e. lower envelope) may be due to selection effects since it lies close to the sensitivity limit of major radio surveys. Therefore this lower envelope has not been confirmed.

### 4.1.2 Stacking

Constraining the radio characteristics of radio faint QSOs is difficult. Large area radio surveys are not sensitive enough to detect faint radio sources and deeper surveys do not amass large numbers of objects. Attempts to use the major radio surveys to study faint radio emitting QSOs have been done via the method of stacking (e.g. Glikman et al., 2004; Wals et al., 2005; White et al., 2007; Falder et al., 2010; Mahony et al., 2012). Wals et al. (2005) stacked blank sky Faint Images of the Radio Sky at Twenty centimetres survey (FIRST) images at the positions of over 8000 QSOs in the 2df quasar redshift survey (2QZ) and found the median flux levels to be 20-40 $\mu$Jy. The typical rms sensitivity of the FIRST survey is 0.15 mJy (Becker, White & Helfand, 1995), so through stacking Wals et al. (2005) were able to lower the rms in the stacked image to $\sim$ 6 $\mu$Jy.

White et al. (2007) stacked over 40000 FIRST images to investigate the radio-quiet population of QSOs from the SDSS Data Release 3 (DR3) spectroscopic catalogue. They advocated the use of the median statistic rather than the mean, as it is less susceptible to influence from outliers, although the relationship with noise in the stacked image is more onerous to determine. Median stacking recovers the arithmetic mean value as the underlying distribution of the stacked sample approaches a
4.1 Introduction

Gaussian and, implementing thresholds for outlier detection, the mean result from stacking is a good representation and easily accessible in terms of the behaviour of the noise. White et al. (2007) found the radio luminosity to be correlated with the optical luminosity, where the radio was increasing as the 0.85 power of the optical luminosity. This implied that the radio loudness parameter decreases with increasing optical luminosity. They found the radio loudness distribution to show bimodality; however, the shape may be due to selection effects. Falder et al. (2010) stacked 71 quasars undetected in the FIRST survey to obtain a weighted mean flux of 100 µJy.

Mahony et al. (2012) looked at higher frequency radio emission at 20 GHz in 874 X-ray selected QSOs to investigate whether bimodality of the radio luminosities appears in this frequency regime. Mahony et al. (2012) recognised that the mean flux result obtained through stacking does not provide information on the underlying distribution of flux density from the individual sources and used the additional information given by the flux density PDF. They also constructed a flux density PDF from random positions to determine if the undetected quasar PDF was due to noise and found it to be appreciably different.

In this Chapter we take radio-undetected, optically selected QSOs, which are expected to make up ~ 90% of the total QSO population (e.g. Kellermann et al., 1989), and use these to extract information at the faint end of the radio luminosity distributions. In this chapter we will investigate the radio-quiet QSO population, undetected in the 1.4 GHz NVSS, using optically selected QSOs from the SDSS, and use the method of stacking to go beyond the sensitivity of the NVSS. In Chapter 5 we go beyond the stacked results, and model the underlying distribution of the radio flux densities within our stacked samples. Then in Chapter 6 we relate the modelling results to the jet efficiency; from the distribution of jet efficiencies of our QSO population we investigate if the theoretical models of jet efficiency can be constrained at the faint end of the luminosity function.
4.2 The SDSS and NVSS surveys

We have selected a volume-limited sample of QSOs from the Sloan Digital Sky Survey (SDSS) Data Release 6 (DR6) and matched their positions with their NRAO VLA Sky Survey (NVSS) images.

SDSS is a large, multiple filter optical redshift survey taking both photometric and spectroscopic data. It began taking data in 2000 with the 2.5 metre telescope at the Apache Point Observatory, New Mexico, USA. The 2008 DR6 was the first release of the SDSS-II extension to the original SDSS with a total imaging footprint of 9583 square degrees.\(^1\)

The NVSS is a 1.4 GHz continuum survey covering the whole sky north of -40 degrees declination. The survey was conducted at the Very Large Array (VLA) in New Mexico, in its most compact configuration. It consists of 2326 image tiles each 4 × 4 degrees in Stokes I, U and Q. The survey’s beam size has FWHM of 45″ and reaches an rms sensitivity of 0.45 mJy/beam (see Condon et al. [1998] for further details).

4.3 Radio emission from our sources

We are using the optical luminosity of the QSOs as a proxy to the accretion rate onto the central super-massive black hole, and the radio luminosities as an indicator to the radio jet power. This is a fair assumption provided there is not significant dust obscuration of the central region, and assuming that the radio emission is dominated by processes related to the QSO. We note that there are a few possible sources that may contaminate the radio flux densities we measure at each QSO position which we address below.

\(^1\)http://www.sdss.org/dr6/
4.3 Radio emission from our sources

4.3.1 Multiple sources

The likelihood that two sources occupy the same position on the sky is low. The source density of the Richards et al. (2009) catalogue is 140 QSO candidates per square degree spanning all redshifts. We have selected a specified photometric redshift band (0.3 < z_{phot} < 0.4, see Section 4.4.1), sampled from the whole sky. The FWHM of the NVSS beam is 45″ so per square degree there are ~ 6000 beam areas on the sky. With only 140 objects per square degree in the whole catalogue, it is unlikely that multiple sources would be in the same position. Indeed, the total SDSS QSO sample after the selection criteria are applied numbers 53156 and the number of sources that are within 45″ of another source is 73 (determined from their SDSS positions). Thus possibly ~ 0.1% of our radio flux density measurements may be contaminated by other sources within our sample. This does not account for sources at redshifts outside of the narrow band we are investigating. However, we assume this form of contamination to affect small percentage of our sources (~ 0.1%).

4.3.2 Missing lobes

The FWHM of the NVSS beam corresponds to 240 kpc at z = 0.4. Therefore we would only be missing extended emission that is > 200 kpc and lacking any significant core emission. Best et al. (2005) cross-matched the SDSS spectroscopic main galaxy sample to the NVSS and FIRST radio samples and found ~ 6% of the radio-loud sources had multiple components. The sources with two NVSS components were typically offset by 20″ to 50″ from the centre. Given the beam FWHM is 45″ we would expect to recover some flux, and our sources are radio-quiet so there could be fewer incidences of multiple component sources than in the radio-loud population.
4.3 Radio emission from our sources

4.3.3 Radio emission from star formation

Radio emission in normal galaxies is an indicator of recent star formation. The emission below 30 GHz is thought to be dominated by synchrotron radiation either produced by the nuclear activity or in the star forming case, driven by supernovae (e.g. Condon 1992). There have been many studies on the origin of radio emission within galaxies hosting radio-quiet QSOs and there is no solid consensus on the source of this emission. Kukula et al. (1998) determined a significant fraction of the radio emission in radio-quiet quasars originates in the nucleus. While Leipski et al. (2006) compared several radio-quiet quasars observed with deep 1.4 GHz, high resolution images to Seyferts, and found the RQQs to be consistent with scaled up Seyferts. Smolčić et al. (2008) investigated the μJy radio population in the COSMOS field observed at 1.4 GHz, and found at the redshift we are probing, the majority of sources are AGN as determined by their colours supporting observations from Simpson et al. (2006).

We follow Kukula et al. (1998) to determine the supernova rates per year (γSN) given the 1.4 GHz luminosity density where \( L_\nu = 2.3 \times 10^{23} (\nu/\text{GHz})^{-\alpha_R} \) γSN, and \( \alpha_R \) is the spectral index where \( s_\nu \propto \nu^{-\alpha_R} \) and \( \alpha_R = 0.8 \). For our sources, the average radio luminosity densities give supernova rates that are close to 1 year\(^{-1} \), which is reasonable in galaxies undergoing star formation. Kimball et al. (2011) carried out deep 6 GHz observations with the Expanded VLA of QSOs within a narrow redshift range. They found a two component radio luminosity function (RLF) which explained their observations with the lower radio luminosity end dominated by SF radio emission and the higher end dominated by AGN with the transition between the two components occurring at \( \log_{10}(L_{6\text{GHz}}) \sim 23.5\text{WHz}^{-1} \).

We are comparing the radio emission from optically selected QSOs and from previous in depth studies of radio emission from RQQs, there is a high chance that a significant proportion of the radio emission is related to the QSO. In addition, the
4.6 Stacking procedure

SDSS QSO photometric catalogue specifically selected point sources, and classified them as QSO candidates through their colours. With this in mind, we make the, perhaps overly, simplistic assumption that the radio emission we observe can be attributed to synchrotron radiation generated by the AGN rather than supernovae.

4.4 Stacking procedure

We stack the individual undetected QSOs to reveal the mean radio flux density of the stacked sample. Stacking is a technique which allows us to get below the sensitivity of a survey by the principle that, when one takes the mean of a set of \( N \) images, the noise of the mean image decreases by \( 1/\sqrt{N} \) and hence allowing for a detection of initially undetected \((< 3\sigma)\) sources. This technique can be used to detect a mean signal which in individual images appears as blank sky. Stacking is used across all wavelengths of observational astronomy and can be very successful provided the astrometry of the images is accurate.

According to previous stacking studies of radio-undetected quasars, it is estimated that we need to reach to the \( \mu \)Jy sensitivity level, e.g. [Falder et al. (2010)] stacked 71 quasars undetected at the 5\( \sigma \) level in the FIRST survey, to find their average flux density of 100 \( \mu \)Jy, with an rms noise of 20 \( \mu \)Jy (a 5\( \sigma \) detection). The level of rms sensitivity of the NVSS is 0.45 mJy; therefore in order to reach the \( \mu \)Jy level of sensitivity in the stacked images (i.e. for a \( > 10\sigma \) detection of a 100 \( \mu \)Jy source), we are required to stack \( \sim 10000 \) images.

The optical QSOs are from [Richards et al. (2009)] photometric quasar catalogue which used the SDSS DR6 data. The photometric catalogue is used for the sheer number of sources, over one million, as opposed to the [Schneider et al. (2010)] SDSS spectroscopic quasar catalogue which has \( \sim 105000 \) sources at the time of this analysis.
4.4.1 SDSS QSO selection criteria

The [Richards et al. (2009)] photometric catalogue reaches a limiting magnitude of $m_I = 21.3$ and spans a large redshift range out to $z \sim 6$. The majority of the photometric redshifts are correct to a fractional uncertainty of $\Delta z \pm 0.3$ (where $\Delta z = \delta z / (1 + z)$). The catalogue only chooses point sources as a rule, so any faint quasar residing in a bright host galaxy where it could appear more extended in the optical wavelengths is omitted.

The SDSS QSOs were selected to have photometric redshifts $0.3 < z_{\text{phot}} < 0.5$ centred around $z_{\text{phot}} = 0.4$ which corresponds to an SDSS I-band rest-frame frequency $\sim 5500\text{Å}$. This region in the quasar spectrum (e.g. see Figure 5 of [Vanden Berk et al., 2001]) is relatively featureless and devoid of strong emission lines. Thus the SDSS I-band PSF magnitudes are not dominated by strong spectral features. This is important as we are using the I-band magnitudes as a proxy for the optical luminosity.

The redshift range specified is relatively narrow, though it reflects a large fraction of the uncertainty in the photometric redshifts ($\delta z = 0.2$), while also eliminating the need for any uncertainty arising from k-corrections. If we did not restrict the redshift of the QSOs in our sample, we would have to apply k-corrections to convert the measured/observed flux densities at the different redshifts (or magnitudes, whichever is the equivalent observed quantity) to the rest-frame equivalent flux density to compare them with each other.

At the redshift range we are looking at, the completeness of the survey is $\sim 95\%$ (as determined by the recovery of the training set quasars used in [Richards et al., 2009]). The redshift uncertainty at $z_{\text{phot}} \sim 0.4$ is lower than the $83\%$ average of the whole sample, which is partly due to host galaxy contamination, though with the fraction correct lying at 0.68 (see Figures 14-16 of [Richards et al., 2009]).
4.5 Matching with the NVSS catalogue

We need to match our optically selected QSO sample to the NVSS catalogue in order to compare the radio and optical luminosity densities.

The NVSS catalogue covers the sky northwards of declination -40°. The 1.4 GHz observations are given in Stokes I, Q and U with an rms error of $\sigma_{I\text{rms}} = 0.45$ mJy \cite{Condon1998}.

The 62424 QSOs we obtain from the SDSS quasar catalogue are split into six samples of 10404 sources. The sample size is chosen so that the noise in the stacked NVSS images, reduced by $\sqrt{N}$, reaches the theoretical noise level of $\sim 4.5$ $\mu$Jy, about a factor of 100 lower than the NVSS survey rms ($\sigma_{I\text{rms}} = 0.45$ mJy). The SDSS QSOs are first sorted in I-band magnitudes ($m_I$), and then they are split into samples by increasing $m_I$, i.e. decreasing accretion rates onto the SMBH.

The optically selected quasar positions are used to extract postage stamp FITS files via the online NVSS Postage Stamp Server\footnote{http://www.cv.nrao.edu/nvss/postage.shtml} using an excerpt from the OBIT OSurvey.PNVSSFetch \textsc{Python} code with guidance from Bill Cotton\footnote{http://www.cv.nrao.edu/~bcotton/Obit.html}. The postage stamp images are 0.05° by 0.05° in size, which corresponds to four times the synthesized beam FWHM of 45", and are centred on the optical positions. The pixel size was 15" per pixel, which resulted in $13 \times 13$ pixel images. Given the astrometry rms error in Stokes I at the NVSS detection limit (2.25 mJy) is of the order of 7", and the SDSS astrometry is expected to have milli-arcsecond precision, we expect the FWHM of the NVSS beam to fully incorporate any error in the positioning of the sources \cite{Condon1998, Pier2003}.

We develop a set of criteria to ensure we retain only the undetected QSOs, discarding any radio-loud sources. This is done by numerically integrating over the

\cite{Condon1998, Pier2003}
4.5 Matching with the NVSS catalogue

NVSS synthesized beam, a two dimensional Gaussian with 45" FWHMs in both dimensions,

\[ S_{\text{int}} = \frac{\sum_{\text{beam}} S_i}{2\pi \sigma_x \sigma_y} \]  

(4.1)

where, \( S_i \) is the peak flux in Jy beam\(^{-1} \), and \( \sigma_x \) and \( \sigma_y \) are the widths of the two dimensional Gaussian. Both the FWHMs in each dimension are given by \( 2\sqrt{(2\ln2)}\sigma \).

As the integration is done over the FWHM of a two dimensional Gaussian, Equation 4.1 should only yield 46% of the total flux.

We remove all sources with an integrated flux above \( 3\sigma_{I_{\text{rms}}} = 1.35 \) mJy. This is below the \( 5\sigma_{I_{\text{rms}}} \) detection threshold for NVSS, which ensures that only non-detected sources are included in our population. This threshold is below the classical definition of the cut-off between radio-loud and radio-quiet AGN. \cite{Miller, Peacock & Mead 1990} defined objects to be radio-loud with \( L_{5 \text{ GHz}} > 10^{25} \) W Hz\(^{-1} \) sr\(^{-1} \) and radio-quiet when \( L_{5 \text{ GHz}} < 10^{24} \) W Hz\(^{-1} \) sr\(^{-1} \). At the redshift we are probing (\( z = 0.4 \)), the 1.4 GHz observed radio flux density corresponds to rest-frame 2.5 GHz (assuming flux density \( s_{\nu} \propto \nu^{-\alpha_R} \) and \( \alpha_R = 0.8 \)). The cut-off from \cite{Miller, Peacock & Mead 1990} corresponds to AGN classified as radio-quiet when \( L_{2.5 \text{ GHz}} < 10^{25.2} \).

Since the NVSS detection threshold corresponds to \( L_{2.5 \text{ GHz}} = 10^{23.84} \) W Hz\(^{-1} \) sr\(^{-1} \) there are many radio-quiet sources (classified by their radio luminosity densities) which are not included in our stacked samples. We have chosen the low cut-off threshold of \( 3\sigma_{I_{\text{rms}}} \) in order to probe the very radio-faint, optically selected AGN with the NVSS.

In addition to removing bright sources from our population of NVSS undetected QSOs, we also avoid bright outliers, by discarding an image that has any pixels with values greater than \( 10\sigma_{I_{\text{rms}}} \). This ensures there is no contamination from bright sources located off-centre in our images, and effectively lowers the noise in
4.6 Stacking

the stacked images (although bright sources lying just outside of our postage stamp image can still affect the rms noise of the image). To simplify the subsequent manipulation of these images, any image without 13 pixels in each dimension is discarded (i.e. if the SDSS quasar positions are too close to the edge of the NVSS sky coverage to be able to construct a 0.05° × 0.05° image then the image is discarded). In total only 436 images from 62424 are discarded due to this criterion (corresponding to < 1%).

The undetected sources from the NVSS catalogue have not gone through the deconvolution process (e.g the cleaning algorithm as mentioned in Chapter 2), and so there are negative pixels in many of the images as a result of sampling the dirty image (true sky image convolved with the dirty beam which has negative sidelobes).

Table 4.1 gives the final numbers of QSOs in each $m_1$ sample, along with the range of I-band magnitudes and the theoretical rms given by $\sigma_{\text{rms}}/\sqrt{N}$. The measured rms of each image is measured by excluding the central region in each image where the stacked source lies, and using the Common Astronomical Software Applications (CASA) imstat task to calculate the rms. That is, the measured rms is $\sigma_{\text{rms}} = \sqrt{\sum_i S_i / N}$, where $S_i$ is the flux in pixel $i$, and $N$ is the number of pixels used.

4.6 Stacking

A technique used across multiple wavelength observations to extract a measurement from a number of undetected sources is called stacking (introduced in Section 4.1). Here we stack the radio images centred on the positions of our optically selected QSOs to get the average flux density of the images in each optical magnitude sample.

\footnote{http://casa.nrao.edu/}
### 4.6 Stacking

<table>
<thead>
<tr>
<th>Sample</th>
<th>( (m_I) ) (range)</th>
<th>( N )</th>
<th>Theoretical rms ( \sigma_{\text{rms}} / \sqrt{N} ) (( \mu )Jy beam(^{-1} ))</th>
<th>Measured rms ( \sigma_{\text{rms}} ) (( \mu )Jy beam(^{-1} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19.09 (14.495-19.767)</td>
<td>8374</td>
<td>4.9</td>
<td>10.4</td>
</tr>
<tr>
<td>2</td>
<td>20.05 (19.767-20.279)</td>
<td>8837</td>
<td>4.8</td>
<td>10.6</td>
</tr>
<tr>
<td>3</td>
<td>20.44 (20.279-20.575)</td>
<td>8871</td>
<td>4.8</td>
<td>8.0</td>
</tr>
<tr>
<td>4</td>
<td>20.69 (20.575-20.795)</td>
<td>9004</td>
<td>4.7</td>
<td>7.1</td>
</tr>
<tr>
<td>5</td>
<td>20.89 (20.795-20.978)</td>
<td>9047</td>
<td>4.7</td>
<td>8.2</td>
</tr>
<tr>
<td>6</td>
<td>21.13 (20.978-21.326)</td>
<td>9023</td>
<td>4.7</td>
<td>8.2</td>
</tr>
</tbody>
</table>

**Table 4.1:** Our QSO population split into samples sorted in increasing optical I-band magnitudes. The mean \( m_I \) is given in the second column along with the range of magnitudes in each sample (in brackets). The measured image rms is obtained by taking the root mean square of all the pixels in the image outside of the central region (a box with bottom left corner pixel coordinates \([3,3]\), and top right corner pixel coordinates \([9,9]\)). The difference between the measured and theoretical rms will be discussed in Section 4.7.

The NVSS images of non-detected QSOs are co-added, with each pixel combined and then divided by the number of images to obtain the mean value for each pixel,

\[
\frac{\sum_{i=1}^{N_i} \text{pix}_i}{N_i},
\]

where \( N_i \) is the number of images in each \( m_I \) sample. Theoretically, the noise in each sample’s stacked image decreases by \( \sqrt{N_i} \); therefore with \( N \approx 10000 \) in each stacked image, we expect the noise to reach a theoretical limit of \( \approx 5 \mu \)Jy (see Table 4.1). However, practically, the rms noise in each stacked image is larger than the predicted value as there could be additional flux from nearby bright sources. We test the behaviour of the noise in the stacked images in Section 4.7.

The final radio-quiet QSO images in each \( m_I \) sample are combined and the mean image is produced (e.g. mean stacking). The resulting images are shown in Figure
4.6 Stacking

4.1 which clearly show detections in at least five of the images (Figures 4.1a to 4.1e), and a tentative (2σ) detection in the faintest \( m_I \) = 21.13 sample (Figure 4.1f). Table 4.2 gives the integrated flux using Equation 4.1.

The CASA task imfit is used to fit the stacked images’ sources to obtain a measure of the total flux in each stacked source. imfit fits a two-dimensional Gaussian to sources within either a whole image, or a specified region of an image. It is not used to measure the total flux of each individual QSO within each of the \( m_I \) samples as it has difficulty fitting sources with very low flux densities, and even more difficulty with non-detections (e.g. where the signal is hidden by the noise in the image). Therefore Equation 4.1 is used instead, and we apply the scaling factor 0.46, to account for the missing flux.

The stacked images are fit with the CASA routine imfit with a fixed box region specified to help with the fit. The results from the fits, including the errors on the flux from the imfit task, are also shown in Table 4.2.

<table>
<thead>
<tr>
<th>( m_I )</th>
<th>FWHM integrated flux (( \mu \text{Jy} ))</th>
<th>Peak flux (( \mu \text{Jy beam}^{-1} ))</th>
<th>imfit integrated flux (( \mu \text{Jy} ))</th>
<th>Significance ( \text{imfit}/\sigma_{\text{rms}} )</th>
<th>Ratio FWHM/imfit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:</td>
<td>19.09</td>
<td>77.2</td>
<td>113.8</td>
<td>167.1 ± 4.8</td>
<td>16.0σ_{\text{rms}}</td>
</tr>
<tr>
<td>2:</td>
<td>20.05</td>
<td>51.7</td>
<td>75.4</td>
<td>128.0 ± 3.8</td>
<td>12.1σ_{\text{rms}}</td>
</tr>
<tr>
<td>3:</td>
<td>20.44</td>
<td>36.9</td>
<td>53.3</td>
<td>88.7 ± 5.1</td>
<td>11.1σ_{\text{rms}}</td>
</tr>
<tr>
<td>4:</td>
<td>20.69</td>
<td>21.7</td>
<td>32.8</td>
<td>45.5 ± 4.2</td>
<td>6.4σ_{\text{rms}}</td>
</tr>
<tr>
<td>5:</td>
<td>20.89</td>
<td>18.8</td>
<td>28.9</td>
<td>41.2 ± 2.7</td>
<td>5.0σ_{\text{rms}}</td>
</tr>
<tr>
<td>6:</td>
<td>21.13</td>
<td>11.9</td>
<td>19.1</td>
<td>19.0 ± 2.6</td>
<td>2.3σ_{\text{rms}}</td>
</tr>
</tbody>
</table>

Table 4.2: The flux measurements of the stacked images. The errors on the imfit integrated fluxes are given by the SNR (the peak model flux/rms of the residual image) found by the imfit routine in CASA, http://casa.nrao.edu/docs/TaskRef/imfit-task.html. The significance of each detection is shown in the fifth column. The errors on the ratios are calculated from the errors on the imfit integrated fluxes.

The ratio of the FWHM integrated and the imfit fluxes are approximately what one would expect from a 2-D Gaussian. Of course, in reality the VLA beam is not a
Figure 4.1: The mean stacked radio-quiet quasar images in each $m_I$ sample. The NVSS synthesized beam is shown in the bottom left hand corner in each plot.
perfect Gaussian (see Figures 16 & 17 in [Condon et al. 1998]), and this accounts for the slight variation from the expected recovery of 46.3% of the total beam integrated flux when only integrating over the FWHM. The faintest optical luminosity sample (Figure 4.1f) yields only a tentative ($2\sigma$) detection; the flux recovered by IMFIT is appreciably lower than expected which could be due to the lack of a strong detection. It is worth mentioning that IMFIT had difficulty deconvolving the source size from the clean beam in this image (the only image where the source could not be deconvolved), which suggests this stacked “source” was too faint to be have a reliable fit from IMFIT.

### 4.6.1 Bootstrap resampling

In order to test the distribution of each of the $m_1$ samples’ populations (i.e. to ensure our mean flux values are not dominated by a few outliers), and to get an approximate error on the integrated flux using the FWHM (Equation 4.1), we perform bootstrap resampling of the flux densities of each source in every sample.

The bootstrap resampling procedure consists of 10000 randomly generated index arrays of size $N$, where $N$ is the number of sources in each bin (shown in Table 4.1). The data (or reference) array contains the FWHM integrated fluxes for every source within each $m_1$ sample. This reference array of size $N$, is then resampled 10000 times with the random indice arrays populated by corresponding values of the reference array. The mean of each of these resampled arrays is computed, and a histogram comprising of 100 bins is constructed from the results. The resulting number distributions are shown in Figure 4.2.

The distributions of the resampled means are fit with Gaussians in order to determine the mean and standard deviations of each $m_1$ samples’ distribution. We are not concerned about the tails of this distribution, where the small number counts in
Figure 4.2: The results from the bootstrap resampling of the original data integrated fluxes in each $m_I$ sample, the Gaussian fit to each distribution is shown by the grey hatched area, and outlined in red. The mean ($\mu$), standard deviation ($\sigma$) and reduced-$\chi^2$ statistics from the Gaussian fits are shown in each samples’ plot. The red vertical lines are the mean values from the Gaussian fits, while the blue vertical lines are the flux values from the stacked images (i.e. the mean values from the original data).
these (histogram) bins affect the symmetry of the errors (in fact using the approximations from Gehrels [1986], the errors on bins with small \( n \) are highly asymmetric). Therefore we fit the Gaussian to the distribution where \( n \geq 10 \), where \( n \) is the number of counts within each histogram bin. We use IDL’s \texttt{gaussfit} routine which performs a \( \chi^2 \) goodness-of-fit test (reliable when \( n \gtrsim 10 \)). The error on each histogram bin is assumed to be Poissonian \( (1/\sqrt{n}) \) above \( n = 10 \). Since our main concern is the fit to the mean and standard deviation of the distribution, and bins with \( n < 10 \) occur only in the tails of the distribution (out to widths of at least \( \sim 4\sigma \)), omitting these bins does not affect the results.

The standard deviations of these fits are used as an estimate on the error of the mean stacked images’ flux densities. For the brightest \( \langle m_I \rangle = 19.09 \) sample, the detection has a signal to noise ratio of \( \sim 20 \), which indicates we should be able to split this sample into smaller sub-samples, and still detect the stacked sources. Table 4.3 lists the errors on the FWHM integrated fluxes in each mean stacked image, and gives the corresponding detection significance. Note the standard deviations from the bootstrap resampling are significantly smaller than the measured errors of the stacked images. In general bootstrapping tends to underestimate the variance in parameters.

<table>
<thead>
<tr>
<th>Sample</th>
<th>FWHM integrated flux (( \mu \text{Jy} ))</th>
<th>Error (( \mu \text{Jy} ))</th>
<th>SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>77.2</td>
<td>4.1</td>
<td>19( \sigma )</td>
</tr>
<tr>
<td>2</td>
<td>51.7</td>
<td>3.8</td>
<td>13( \sigma )</td>
</tr>
<tr>
<td>3</td>
<td>36.9</td>
<td>3.8</td>
<td>10( \sigma )</td>
</tr>
<tr>
<td>4</td>
<td>21.7</td>
<td>3.8</td>
<td>6( \sigma )</td>
</tr>
<tr>
<td>5</td>
<td>18.8</td>
<td>3.8</td>
<td>5( \sigma )</td>
</tr>
<tr>
<td>6</td>
<td>11.9</td>
<td>3.7</td>
<td>3( \sigma )</td>
</tr>
</tbody>
</table>

Table 4.3: The flux measurements of the stacked images. The errors are from the fits of the bootstrap resampling distributions.
4.6.2 Splitting the optically brightest bin

As can be seen from Table 4.3, the brightest \( m_I \) sample achieves a signal to noise ratio (SNR) of around 20. In order to get more out of the data we have (i.e. create more stacked detection measurements for our analysis) we split this sample into three sub-samples. This is achieved by specifying a desired theoretical SNR of \( \sim 10 \), and co-adding each individual postage stamp until that limit is reached. Each sub-sample required enough images to ensure an adequate number of sources in each histogram bin (i.e. for statistical purposes) hence the different SNRs for each sub-sample. For the brightest, in order to get a reasonable number of sources, the theoretical SNR was set to 11. The next bin, required a theoretical SNR of 8, and the remaining sub-sample contained all the rest of individual QSOs from the \( \langle m_I \rangle = 19.09 \) sample. Table 4.4 gives the new \( m_I \) sub-samples’ properties, while Figure 4.3 shows the results from the stacking.

The bootstrap resampling was repeated for these three sub-samples, and the results are shown in Figure 4.3.

Averaging the FWHM/imfit ratio for all the \( m_I \) samples, i.e. the three sub-samples (1(a-c) in Table 4.4) along with the original samples (2-6 in Table 4.2), yields an overall ratio of 0.48 ± 0.3 (where the error comes from the variance in the ratio values), supporting the approximation of using a 2-D Gaussian for the NVSS synthesised beam, and applying the scaling factor 0.46 to the FWHM integrated fluxes.

4.7 Testing the behaviour of the noise

We perform a simple test on the stacked noise to see if it follows the expected \( 1/\sqrt{N} \) trend. We did this by measuring the rms noise in each individual NVSS
Testing the behaviour of the noise

Figure 4.3: The three further mean stacked radio-quiet QSO images from the brightest $\langle m_1 \rangle = 19.09$ sample.
4.7 Testing the behaviour of the noise

Figure 4.4: The results from the bootstrap resampling on the additional optical magnitude samples, with Gaussian fits to each distribution. The mean, standard deviation and $\chi^2$ statistic of the Gaussian fit is shown in each figure. The red vertical line shows the mean from the Gaussian fit, while the blue shows the mean from the original data and is illustrating the closeness of the Gaussian fit to the original data.

(a) $\langle m_1 \rangle = 17.88$

(b) $\langle m_1 \rangle = 18.72$

(c) $\langle m_1 \rangle = 19.43$
Testing the behaviour of the noise

<table>
<thead>
<tr>
<th>Property</th>
<th>Sub-sample 1(a)</th>
<th>Sub-sample 1(b)</th>
<th>Sub-sample 1(c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>1221</td>
<td>1357</td>
<td>5796</td>
</tr>
<tr>
<td>( (m_I) ) (range)</td>
<td>(17.88, 18.463)</td>
<td>(18.72, 18.944)</td>
<td>(19.43, 19.767)</td>
</tr>
<tr>
<td>Theoretical rms ( \left( \sigma_{\text{rms}} / \sqrt{N} \right) ) (( \mu )Jy beam(^{-1} ))</td>
<td>12.9</td>
<td>12.2</td>
<td>5.9</td>
</tr>
<tr>
<td>Measured rms ( (\mu )Jy beam(^{-1} ))</td>
<td>19.5</td>
<td>13.2</td>
<td>11.0</td>
</tr>
<tr>
<td>FWHM integrated flux ( (\mu )Jy) )</td>
<td>141.7</td>
<td>97.7</td>
<td>58.8</td>
</tr>
<tr>
<td>Peak flux ( (\mu )Jy beam(^{-1} ))</td>
<td>216.8</td>
<td>142.9</td>
<td>85.4</td>
</tr>
<tr>
<td>IMFIT integrated flux ( (\mu )Jy) )</td>
<td>266.2 ± 8.7</td>
<td>202.0 ± 13.0</td>
<td>143.0 ± 5.7</td>
</tr>
<tr>
<td>Ratio ( \text{FWHM/IMFIT} )</td>
<td>0.53 ± 0.02</td>
<td>0.48 ± 0.03</td>
<td>0.41 ± 0.02</td>
</tr>
<tr>
<td>Significance</td>
<td>( 14\sigma_{\text{rms}} )</td>
<td>( 15\sigma_{\text{rms}} )</td>
<td>( 13\sigma_{\text{rms}} )</td>
</tr>
</tbody>
</table>

**Table 4.4:** The brightest optical luminosity bin split into three bins (a-c). The significance of each detection is given by the IMFIT integrated flux and the measured stacked image rms (from IMSTAT).

postage stamp cut-out image and in each cumulative stacked image. The rms in the cumulative image is compared to the theoretical noise \( \sigma_{\text{rms}} / \sqrt{N} \) where \( \sigma_{\text{rms}} = 0.45 \) mJy. Figure 4.5 shows the measured rms for each optical magnitude sample; the blue line is the rms measured for the cumulative stacked image, while the red line is the theoretical rms noise. Figure 4.5 reveals the noise in the stacking process behaves as expected, and does not suffer effects from saturation of the field, i.e. the confusion limit where the noise fails to behave as \( 1/\sqrt{N} \). The final measured rms is slightly greater than the theoretical rms; however, this is not unexpected as there could be contamination to the measured rms from nearby bright sources for example.
Testing the behaviour of the noise

Figure 4.5: The measured rms errors on the individual NVSS postage stamp cutouts (black dots), the measured rms of each successive stacked image (blue line), and the theoretical rms noise where the noise follows $\sigma_{\text{rms}}/\sqrt{N}$ (red line). The blue and red lines follow the same trend, although the final measured rms in the stacked images of all the samples is above the theoretical noise, which is not unexpected.
Testing the behaviour of the noise

\[ \langle m_1 \rangle = 20.44 \]

\[ \langle m_1 \rangle = 20.69 \]

\[ \langle m_1 \rangle = 20.89 \]

\[ \langle m_1 \rangle = 21.13 \]

Figure 4.5: (continued).
4.8 Searching for a radio-quiet quasar envelope

To determine whether there is a hint of a lower envelope in the radio luminosity given a certain optical luminosity, we use the average optical and radio luminosity densities from the stacked samples. Our optically selected QSOs are all undetected in the NVSS survey, which means that we have gone beyond the sensitivity limitations of the NVSS survey to look at the radio low-luminosity sources without introducing a possible false correlation between the optical and radio by only using the detected objects.

To compare the radio and optical luminosity densities, we must first convert the optical magnitudes into flux densities, and the convert the optical and radio flux densities to their rest-frame, k-corrected luminosity densities with Equation 4.3:

\[
\frac{L_\nu}{[\text{WHz}^{-1}]} = \left(\frac{4\pi}{(1+z)(1-\alpha)}\right) \left(\frac{D_{\text{lum}}(z)}{[\text{m]}}\right)^2 \left(\frac{s_\nu}{[\text{Wm}^{-2}\text{Hz}^{-1}]}\right),
\]

where \(D_{\text{lum}}(z)\) is the luminosity distance at redshift \(z\), and \(s_\nu\) is the flux density in \(\text{Wm}^{-2}\text{Hz}^{-1}\) (1 Jy = \(10^{-26}\) Wm\(^{-2}\)Hz\(^{-1}\)). The flux density is assumed to follow a power-law with frequency, \(s_\nu \propto \nu^{-\alpha}\), where we have assumed \(\alpha = 0.8\) for the radio and \(\alpha = 0.5\) for the optical.

The SDSS I-band PSF magnitudes are converted into AB magnitudes following the instructions given by the SDSS website\(^5\). The SDSS I-band is approximately on the AB magnitude system where the flux may be determined via the AB conversion

\[
\left(\frac{s_\nu}{[\text{Jy}]}\right) = \left(\frac{3631}{[\text{Jy}]}\right) \times \left(\frac{f}{f_0}\right)
\]

where the zero-point flux in the AB system is 3631 Jy. The photometric zero-point

\(^5\)http://www.sdss.org/dr6/algorithms/fluxcal.html
count rate is the \( f/f_0 \) term and is determined by inverting the equation for asinh magnitudes,

\[
m = -(2.5/\ln(10)) \times \text{asinh}
\left(\frac{f/f_0}{2b} + \ln(b)\right),
\]

(4.5)

where \( b \) is the softening parameter which depends on the filter. These fluxes are then converted into luminosities via Equation 4.3. The rest-frame SDSS I-band luminosities (460 nm) fall close to the traditional B-band wavelength (445 nm).

Table 4.5 gives the luminosity densities for the eight stacked detections. The radio versus optical rest-frame luminosity densities are shown in Figure 4.6 with the mean stacked images from the eight stacked radio-undetected samples shown as squares.

<table>
<thead>
<tr>
<th>( \langle m_1 \rangle )</th>
<th>( \log_{10}(L_{2.5} \text{ GHz}/[\text{WHz}^{-1}\text{sr}^{-1}]) )</th>
<th>( \log_{10}(L_B/[\text{WHz}^{-1}\text{sr}^{-1}]) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.88</td>
<td>22.98</td>
<td>23.01</td>
</tr>
<tr>
<td>18.72</td>
<td>22.82</td>
<td>22.61</td>
</tr>
<tr>
<td>19.43</td>
<td>22.60</td>
<td>22.33</td>
</tr>
<tr>
<td>20.05</td>
<td>22.54</td>
<td>22.08</td>
</tr>
<tr>
<td>20.44</td>
<td>22.39</td>
<td>21.92</td>
</tr>
<tr>
<td>20.69</td>
<td>22.16</td>
<td>21.82</td>
</tr>
<tr>
<td>20.89</td>
<td>22.10</td>
<td>21.74</td>
</tr>
<tr>
<td>21.13</td>
<td>21.90</td>
<td>21.64</td>
</tr>
</tbody>
</table>

Table 4.5: The stacked mean radio and optical luminosity densities for the eight optical magnitude samples. The SDSS I-band optical luminosity densities are converted to their AB magnitudes and then to flux densities using Equation 4.4. The flux densities from both the optical and radio bands are converted to their rest-frame luminosity densities via Equation 4.3. The scale factor of 0.46 has been applied to all the radio fluxes which were determined by integrating over the beam’s FWHM (see Section 4.6).

Figure 4.6 shows the relationship between optical luminosity density and radio luminosity density for the radio-quiet stacked QSOs, whereby, the brightest optical luminosity density bin yields the brightest radio luminosity density. In order to get a clear picture of the radio quiet quasar envelope we are trying to determine, we
Figure 4.6: The rest-frame radio luminosity density vs. the rest-frame optical luminosity density for each mean stacked optical luminosity bin (squares), and for each NVSS detected quasar in our SDSS $z_{\text{phot}} \sim 0.4$ original population. The grey, dashed, horizontal line indicates the $5\sigma_{\text{rms}}$ detection limit of NVSS. The limits of each optical luminosity bins is shown by the black lines extending left and right from each of the squares. The $3\sigma$ errors on the stacked radio luminosities from the bootstrap resampling are plotted as red error bars. (We use $3\sigma$ purely for illustrative purposes). The scale factor of 0.46 has been applied to all the radio fluxes which were determined by integrating over the beam’s FWHM (see Section 4.6). The black asterisks at $\log_{10}(L_{2.5\text{GHz}} > 23.8$ are the NVSS detected QSOs in the redshift range we selected.
have included the NVSS detected QSOs from our SDSS photometric sample. The radio flux densities for each source which was equal to or above the NVSS detection threshold of $5\sigma_{\text{rms}}$ (2.25 mJy), is included in Figure 4.6.

4.9 Discussion

We have investigated the 1.4 GHz radio flux density of undetected QSOs using a sample of QSOs selected from the SDSS survey to be within the photometric redshift range $0.3 < z_{\text{phot}} < 0.5$. The position of these QSOs are matched with the NVSS survey and images extracted from online database of NVSS. The QSOs that are undetected in the NVSS survey are kept and split into eight samples of increasing optical magnitude. The radio flux density of these otherwise undetected sources was detected via stacking. Stacking allowed us to see beyond the noise and measure the mean flux density of each optical magnitude sample.

We have assumed the radio emission is predominantly due to the AGN and not star formation. The radio luminosity densities ($L_{1.4\,\text{GHz}}$) determined from the stacked images can be used to infer rough estimates of the star formation rate via the relation (assuming a constant burst of star formation over 100 Myr; [Sullivan et al., 2001]):

$$\text{SFR}(L_{1.4\,\text{GHz}}) = \frac{L_{1.4\,\text{GHz}}}{8.85 \times 10^{20}},$$

(4.6)

where the star formation rate (SFR) is in M$_{\odot}$ yr$^{-1}$. Our stacked $L_{1.4\,\text{GHz}}$ give SFRs ranging from $\sim 200$ M$_{\odot}$ yr$^{-1}$ in the brightest sample, to $\sim 10$ M$_{\odot}$ yr$^{-1}$ in the faintest bin. These estimates are assuming that all radio emission is from star formation.

Figure 4.6 shows the radio luminosity densities versus the radio luminosity densities for the eight optical magnitude samples. In addition for comparison, the QSOs within the $z_{\text{phot}}$ range that are detected by the NVSS survey are shown as black
asterisks. The detected sources at similar optical luminosity densities span a wide range of radio luminosity densities. This is further illustrating the range of radio luminosity densities QSOs can emit while maintaining similar optical luminosity densities which we have taken as a proxy to bolometric luminosity. If we assume the bolometric luminosity is related to the accretion rate via \( L_{\text{bol}} = \epsilon \dot{m} c^2 \), where \( \epsilon \) is the radiative efficiency (the efficiency of converting the rest-mass energy due to the accreting matter into radiation), assuming \( \epsilon \) varies only a little so it can be approximated to a constant, we can infer that a larger \( L_{\text{bol}} \) implies a larger accretion rate \( \dot{m} \). Therefore we are assuming sources that share a similar optical luminosity density, share a similar bolometric luminosity and hence accretion rate.

The stacked radio 1.4 GHz detections show an apparent correlation with the optical luminosity density; as the optical luminosity density increases so does the radio luminosity density. Therefore from the assumption of the optical luminosity density being an indicator of the accretion rate via the bolometric luminosity, the radio luminosity density increases with accretion rate. This correlation can be interpreted as a minimum radio luminosity (jet power) for a given optical luminosity (accretion rate). The overall correlation between the radio luminosity (jet power) and optical luminosity (accretion rate) suggests that jet power grossly correlates with accretion rate. However, the scatter (known as the radio loudness) covers several decades, which means a second process must cause this scatter, i.e. it must be able to cause, for given accretion rate, huge variation in jet power.

However, stacking is only demonstrative of the true underlying distribution of the radio flux if the distribution can be fairly represented by the mean value, e.g. a Gaussian distribution. In Section 4.6.1 we tested the shape of the distribution of the flux densities from the individual QSOs and found their distributions were Gaussian-like in all eight optical magnitude samples. This suggests the mean stacked flux density is representative of the overall distribution. That being said however,
the images were all non-detections, i.e. the signals were hidden or comparative to the noise, and it would be useful to remove the NVSS noise from each image and then measure the underlying radio flux density from the QSOs. Then we could parameterise the distribution of the underlying QSO radio flux density. If the shape of this distribution was different, e.g. could not be well approximated by the mean, then the stacked detections which appear to be correlated with the optical measurements, are giving us a false view of the nature of the radio undetected QSOs. The appearance of lower envelope where the radio luminosity density is correlated with the optical luminosity density in Figure 4.6 may not be what the true underlying QSO radio flux density distribution shows, but rather due to the noise, the measured flux density distribution takes the shape similar to a Gaussian.

4.10 Conclusion

From the stacked detections we see the radio luminosity increases with optical luminosity, thus there is a hint of correlation which could be interpreted as evidence for a lower envelope. The lower envelope scenario has repercussions on our understanding on the mechanism for the processes responsible for radio emission from QSOs, e.g. indicating the accretion rate has an effect on the possible minimum jet efficiency. To investigate this apparent correlation further, we conduct an investigation into the distribution of the QSO flux density distribution. This will reveal whether the stacked mean values are representative of the distribution of the individual sources’ radio flux densities within each stacked sample. The next chapter details the modelling the underlying QSO flux density distributions.
Chapter 5

Beyond stacking: modelling the distributions of QSO flux densities

In this chapter we look beyond the mean stacked results for each optical magnitude sample by constructing flux density distributions for the populations in each sample. We model these QSO flux density distributions with three models: an empirical distribution of the noise alone, a power-law plus the noise and a single Gaussian plus the noise. The best-fitting parameters of the power-law and the Gaussian are determined by calculating the maximum likelihood of these models fit to the observed data. These three models are compared using the odds ratio with the power-law plus noise model preferred for all the samples. Thus, the power-law model best describes the underlying QSO flux density distribution indicating that there is no lower-envelope, i.e. no minimum radio power for a given accretion rate.

5.1 Motivation

The results from the mean stacking in Chapter 4 hint at an apparent lower envelope, i.e. there is a minimum radio luminosity for a given optical luminosity. However, such a conclusion would be driven purely by the appearance of Figure 4.6 based on the detected radio flux density from the stacking of the optical magnitude samples. The detections from the stacking analysis may not represent the distribu-
5.1 Motivation

tion of the individual source flux densities within the stacked sample. This is true if
the mean (or median) value is not drawn from a distribution that is clustered around
the mean (e.g. a Gaussian distribution) but rather drawn from a distribution which
spans many orders of magnitude, or is dominated by objects at one end (e.g. the
faint end) and only a very few at the other end of the distribution (e.g. a few bright
objects) such as a power-law distribution.

In Chapter 4 we tested the distribution of the measured flux densities in each
optical magnitude sample via multiple resampling of the distribution in each optical
magnitude sample (e.g. the bootstrapping done in Section 4.6.1). The results from
these tests reveal the mean value of the measured flux density is a good representa-
tion of the sample’s flux density. However, we have selected QSOs that are below the
NVSS survey’s detection limit and are often indistinguishable from the noise. The
distribution of the noise in the NVSS could be affecting the measured flux density
distribution shape, and the true shape of the distribution of the underlying QSO
flux densities would be hidden. If the true underlying QSO flux density distribution
cannot be well represented by the mean, our physical interpretation of the lower
envelope seen in Figure 4.6 would be incorrect. There could be many QSOs with
radio flux densities fainter than this mean value, which would indicate that there
is no minimum radio luminosity density for a given optical luminosity density. For
this reason we undertake further investigations into the properties of these stacked
samples by modelling their distributions, in order to determine whether the lower
envelope scenario is indeed correct. The following sections describe in detail the
steps taken in modelling the QSO flux density distributions in each of the eight
optical magnitude samples and the model selection used to discriminate between
different models before interpreting the results.
5.2 Modelling

The radio flux density we measure for each individual QSO in each optical magnitude sample has at least two major contributions: 1) noise that is inherent in the NVSS, and 2) the underlying flux density emitted by the QSO itself. Since we have deliberately discarded any detected or partially detected sources, the contribution from the individual QSO is less than $3\sigma_{\text{rms}}$ of the noise. In other words, the QSO signal is hidden by the noise. Therefore, in order to model the QSO flux density distribution, we must take the noise into account. To do this we need to construct the flux density distributions of each optical magnitude sample and then also a distribution profile for the noise.

From now on we are dealing with the flux densities measured from the NVSS postage stamp cutouts. Since we are considering the non-detections we cannot use imfit as explained earlier in Section 4.6, so we are using the FWHM integrated flux densities from Equation 4.1. Also, we are omitting the scale factor of 0.46 (this is the amount of flux the FWHM integrated flux recovers as opposed to integrating over the whole beam) since we are not comparing the radio data with the optical anymore. The scale factor has little impact on the physical interpretation of the distribution of the underlying QSO flux density in our modelling since we are looking for the shape of the radio flux density distribution only, and the most important thing we need to take care of is to be consistent with our handling of the individual measure fluxes both for the QSO positions and the blank sky positions (noise distribution).

5.2.1 Quasar flux density distributions

In each sample selected in SDSS I-band magnitude ($m_I$), the integrated radio flux densities of the individual quasars are determined using Equation 4.1. The flux densities are then split into bins of equal width, with the lowest flux density value
equal to the minimum quasar radio flux density in any of the \(<m_1\) samples (-4 mJy in the \(<m_1\) = 20.89 sample), and with the upper limit as the maximum quasar radio flux density (≈ \(3\sigma_{\text{rms}}\) due to the non-detection threshold explained in Section 4.5 where \(\sigma_{\text{rms}} = 0.45\) mJy beam\(^{-1}\) is the rms of the NVSS survey).

The resolution of these flux density bins is chosen to be several times the measured rms of the \(<m_1\) samples. The width of the bins is chosen such that the uncertainty associated with each bin is dominated by the vertical uncertainty (e.g. the number of sources in each bin), rather than the horizontal uncertainty (e.g. the measurement error on the flux density). Hence we adopt the resolution of 61.25 \(\mu\)Jy which is approximately \(3\sigma_{\text{rms}}\) of the brightest \(<m_1\) = 17.63 sample (this sample has has the largest \(\sigma_{\text{rms}}\), equal to 19.5 \(\mu\)Jy).

The QSO flux density distributions are converted into probability densities and normalised by

\[
p(x) = \frac{n_x}{\sum_x n_x}
\]

where \(p(x)\) is the probability of bin \(x\), \(n_x\) is the number count of bin \(x\). Each \(<m_1\) sample is binned to the same resolution for the subsequent manipulations of different samples to be the same for each. This gives the QSO radio flux density distributions shown in Figure 5.1.

The QSO flux density distributions are, by construction, mostly ‘below the noise’ of the NVSS survey; hence the distributions shown in Figures 5.1 are a combination of some underlying source flux density distribution from the QSOs with the noise contribution. In order to isolate the QSO radio flux density distribution from the noise, we are first going to construct an empirical noise distribution. We will then construct different models of the emitted flux density from the QSOs, which will be combined with the noise distribution and compared to the data (shown in Figure
Figure 5.1: The radio flux density distributions in each of the $\langle m_I \rangle$ samples. The flux densities have been binned to the same resolution ($x_{\text{res}} = 61.25 \, \mu\text{Jy}$), with the same range of values. The x-axes have been cut at the low end to -2000 $\mu$Jy in order to see the shape of the distributions, (though the minimum quasar radio flux density is -4471 $\mu$Jy in the $\langle m_I \rangle = 20.89$ sample, only one source resides in this bin). At the high flux density end, the non-detection criteria is seen as each distribution is cut at $3\sigma_{\text{rms}} = 1350 \, \mu\text{Jy}$ shown by the dashed vertical line.

(a) $\langle m_I \rangle = 17.88$

(b) $\langle m_I \rangle = 18.72$

(c) $\langle m_I \rangle = 19.43$

(d) $\langle m_I \rangle = 20.05$
Figure 5.1: (continued).
5.2 Modelling

The noise distribution is constructed from random sky positions in the NVSS and is handled in the same way as the QSO NVSS cut-outs (e.g. Section 4.5). This ensures there are no radio sources contaminating our noise distribution. We will compare the QSO flux density distributions (Figure 5.1) to three models: a) the noise itself, with no underlying distribution of QSO radio flux densities added to it, b) a single power-law model of the underlying QSO flux density plus the noise, c) a single Gaussian model plus the noise. This is detailed in Section 5.3. The models will be evaluated against each other using the odds ratio (or Bayes factor, e.g. Sivia & Skilling, 2006) presented in Section 5.4.1 and the preferred model is then interpreted as our best representation of the underlying distribution of radio flux densities for our QSOs.

The power-law model has one varying parameter, which is its slope the index; which we consider to range from -5 to 5. The negative indices indicate the majority of the sources would have increasingly smaller flux densities, while the positive indices would indicate the opposite, an increasing number of sources with greater flux densities. The Gaussian model has two variable parameters, the mean and the width. A Gaussian model would suggest the majority of the sources have fluxes close to the mean of the Gaussian, and the width determines the spread of the QSO flux density from this mean.

The notion of a radio-quiet envelope suggests a minimum radio flux, and therefore a minimum jet power, for a given accretion rate. If the Gaussian model is preferred, and its mean is positioned at some point of flux density greater than zero, this could point to a minimum radio flux, and in this case the mean value (e.g. from stacking) would be meaningful. A power law with a negative index, would instead show that there is an increasing number of sources with less flux, and the flux density that the stacking exercise has given us, is the average of this power-law distribution but does
5.2 Modelling

not include a ‘minimum’ flux density in the sense of the radio-quiet envelope.

5.2.2 Building a noise profile

We wish to characterise the underlying QSO flux density distribution beyond the noise level of the NVSS survey. Consequently, a model of the noise profile of the NVSS must be constructed, including enough points of measurement to ensure a good statistics. The latter is necessary to ensure the number of points in each bin is large enough to drive the uncertainty of that bin’s population down to a negligible level; this uncertainty is based on Poisson noise $\sqrt{n}$ where $n$ is the number in each bin. To achieve this goal, we generate 150000 random positions in RA$_{J2000}$ and $\delta_{J2000}$. To reconstruct a uniform distribution of random positions on the surface of the sky, we use the cosine of the declination. In spherical coordinates the area element is given by $\sin \theta d\theta d\phi$ where $\theta \in [0, \pi]$ and is analogous to the declination, while $\phi \in [0, 2\pi]$ is analogous to the right ascension. The random positions are then used to fetch NVSS postage stamps of $13 \times 13$ pixels in size (the same size as the QSO postage stamps).

The same selection criteria used for the QSOs are applied to the random position images in order to discard any chance detections our 150000 random positions may yield. Therefore, according to the criteria set out in Section 4.5, the size of the image must be $13 \times 13$ pixels, the integrated flux density calculated using Equation 4.1 must be below the $3\sigma_{I_{\text{rms}}}$ threshold we set for the QSOs, and there must be no bright outliers (pixels greater than $10\sigma_{I_{\text{rms}}}$).

Following the selection process we are left with 129110 random position noise images. The flux densities in each image are determined by integrating over the beam FWHM (Equation 4.1), and then a histogram of the noise distribution is constructed. The bounds of the histogram are determined by the minimum and
maximum QSO flux density in the whole population, just as in Section 5.2.1. The size of each bin is equal to the resolution of the QSO distributions (61.25 $\mu$Jy).

The resulting NVSS noise profile is shown in Figure 5.2, and is approximately Gaussian in shape, slightly positively skewed, though cut off at the high flux density end (an artifact of the $3\sigma_{\text{rms}}$ cut off we imposed to ensure non-detections). Since the mean is approximately zero there are negative regions which are artifacts of the imaging process and the fact that none of our random positions (or the SDSS QSOs) are detected in the NVSS and therefore were not CLEANed (Section 2.2.1).

**Figure 5.2:** The noise radio flux density probability distribution constructed from the random positions in the NVSS sky. The flux densities are binned with equal resolution (61.25 $\mu$Jy). The dashed vertical line shows the $3\sigma_{\text{rms}}$ cutoff imposed to make sure there are no detected sources. The distribution is positively skewed with a mean flux density very close to zero (of the order of $\sim 10^{-9}$ $\mu$Jy).
5.3 The models of the underlying QSO flux density distributions

The models of the intrinsic underlying QSO flux density distributions were chosen to be a power-law and a single Gaussian. The two models are constructed using radio flux density bins with widths, again, equal to the resolution of the observed samples’ distribution. The models cover a range of flux densities, from 0 $\mu$Jy to $5\sigma_{I_{\text{rms}}}$. We have extended the upper limit out to $5\sigma_{I_{\text{rms}}}$ despite the cut off of our measured NVSS radio flux densities at the $3\sigma_{I_{\text{rms}}}$ level. This is to allow a higher intrinsic QSO radio flux density than the $3\sigma_{I_{\text{rms}}}$ cut off which could be affected by noise to bring it back to $\leq 3\sigma_{I_{\text{rms}}}$. Since we are modelling the emitted QSO radio flux, we do not allow the flux density in these models to be negative. Negative flux, however, is seen in the noise distribution which is approximately Gaussian in shape centred on 0 $\mu$Jy (see Figure 5.2).

The final radio flux density distribution, which hereafter we refer to as the model distribution, is compared to the observed optical magnitude sample distributions. The model distribution is built using the noise profile and the QSO underlying flux models. The probability at the flux density bin $x_i$ ($p(x_i)$) is equal to the sum of all possible combinations of the probability of the QSO underlying flux density at bin $p(x_m)$ multiplied by the probability of the noise at flux density bin $p(x_N)$, since $p(x_m)$ and $p(x_N)$ are independent. The subscripts $m$ and $N$ stand for ‘underlying QSO flux density model’ and ‘noise’ respectively.

At each flux density bin $x_i$ (where $x_i = x_m + x_N$), the probability is constructed as follows:

$$p(x_i) = \sum_m p(x_m)p(x_N) \quad (5.2)$$
where \( N = (i - m) \), i.e. the flux density bin in question of the noise profile \( x_N \), is the required flux density value for the model \( x_i \) minus the underlying QSO flux density bin \( x_m \). Figure 5.3 provides an illustrative aid as to how \( p(x_i) \) is constructed; however it is not shown explicitly in Figure 5.3 that for the particular flux value of \( x_i \), the \( p(x_i) \) will be given by the sum in Equation 5.2.

\[
\begin{align*}
X_i &= X_N + X_m \\
p(x_i) &= \sum_m p(x_m) p(x_i)
\end{align*}
\]

**Figure 5.3:** An illustration of the construction of the models used to compare to the stacked sample distributions. The top left distribution corresponds to the noise profile, and the top right, the underlying QSO flux density model (in this case a power-law). The bottom distribution is the model we compare to the observed distribution, with \( p(x_i) \) for the shaded bin at \( x_i = x_m + x_N \) determined via Equation 5.2.

Underlying QSO flux density distributions are made for all allowed variations in the parameters (power-law: index \( \alpha \), Gaussian: mean \( \mu \) and width \( w \)). Then the comparison models are constructed as above, using the noise profile, and these are fitted to each optical magnitude sample flux density distribution to get the best-fitting parameters.
5.3 The models of the underlying QSO flux density distributions

5.3.1 Power-law with noise

The first model we fit is the single power-law model of the underlying QSO radio flux density with only one parameter, the index $\alpha$. The power-law model is combined with the noise to produce the model we compare to the observed flux density distribution. Our prior on $\alpha$ is flat over the parameter space we sample ($[-5,5]$), and zero elsewhere:

$$p(\alpha) = \begin{cases} 
\text{constant}, & \text{for } -5 \leq \alpha \leq 5, \\
0, & \text{otherwise.}
\end{cases} \quad (5.3)$$

The indices are sampled with $\Delta \alpha = 0.025$ which samples the parameter space finely while not becoming too computationally expensive; the power-law models combined with the noise distribution are calculated at each $\alpha$ value. We will determine the best-fitting model to each optical magnitude sample by finding the parameters which maximise the posterior probability distribution function (PDF) described in Section 5.4.

5.3.2 Single Gaussian with noise

The other model we fit is a single Gaussian with two variable parameters. The model is again combined with the noise distribution. The two variables of the Gaussian model are the mean ($\mu$) and the width (which we denote as $w$ since we have used $\sigma$ earlier to refer to the rms error). The priors on these two variables are flat over the range we sample, i.e. over $(0 \geq \mu \geq 1219) \ \mu\text{Jy}$ and $(61.25 \geq w \geq 6125) \ \mu\text{Jy}$, and
Parameter estimation and model comparison

\[ p(\mu) = \begin{cases} 
\text{constant}, & \text{for } (0 \leq \mu \leq 1219) \text{ } \mu \text{Jy}, \\
0, & \text{otherwise}; 
\end{cases} \] 

(5.4)

\[ p(w) = \begin{cases} 
\text{constant}, & \text{for } (61.25 \leq w \leq 6125) \text{ } \mu \text{Jy}, \\
0, & \text{otherwise}. 
\end{cases} \] 

(5.5)

The range of widths sampled are in the same resolution as the observed QSO distributions \((x_{\text{res}} = 61.25 \text{ } \mu \text{Jy})\), with a minimum value equal to the \(x_{\text{res}}\). However, we have sampled the mean of the model with a finer resolution \((x_{\text{res}}/10)\). Although we are not sensitive to features below \(x_{\text{res}}\), we have allowed the centre of the Gaussian model to lie within the distribution’s resolution in order to sample a larger parameter space. Again, the best fitting parameters will be determined from the maximum of the posterior PDF, details of this technique are outlined in Section 5.4.

5.4 Parameter estimation and model comparison

The best-fitting parameters for the models are determined via the maximum likelihood using Bayes theorem (Equation 3.2). We are modelling here a discrete distribution, hence the likelihood is given by a Poissonian distribution. The likelihood in Equation 3.2 becomes,

\[ p(N_k|D_k) = \prod_{k}^{k} \frac{D_k^{-N_k}e^{-D_k}}{N_k!}, \] 

(5.6)

where \(N_k\) is now the data, \(D_k\) is the model, and \(k\) is the bin. For computational reasons we are using the natural logarithm of the likelihood function which we denote \(L\) (also referred to as the log-likelihood). In order to calculate \(\ln(N_k!)\) at large values of \(N\) \((N \geq 170)\), we use Stirling’s approximation where \(\ln(N!) = \)
5.4 Parameter estimation and model comparison

\[ N \ln(N) - N + \ln(\sqrt{2\pi N}). \] Equation 5.6 in log form is then,

\[ \mathcal{L} = \sum_k N_k \ln(D_k) - D_k - \ln(N_k!) \] (5.7)

where \( N_k \) and \( D_k \) are the same as above.

5.4.1 Model Comparison

We compare two models using the odds ratio to see which model is preferred. The odds ratio is given by the ratio of the evidences e.g.

\[ B_{a/b} = \frac{\text{p}(\text{data}|M_a)}{\text{p}(\text{data}|M_b)} = \frac{\int \text{p}(\text{data}|M_a : \{\theta_a\}) \text{p}(M_a : \{\theta_a\}) \text{d}\theta_a}{\int \text{p}(\text{data}|M_b : \{\theta_b\}) \text{p}(M_b : \{\theta_b\}) \text{d}\theta_b}, \] (5.8)

where the prior terms are cancelled out (\( \text{p}(\text{data}) \)). The odds ratio compares models \( M_a \) and \( M_b \), that each have a set of parameters \( \{\theta_a\} \) and \( \{\theta_b\} \). The term \( \int \text{p}(\text{data}|M : \{\theta\}) \text{p}(M : \{\theta\}) \text{d}\theta \) is called the marginal likelihood and it causes the odds ratio to incorporate the “Occam’s razor”, which favours simpler models with fewer parameters. Model \( M_a \) is decisively preferred over model \( M_b \) i.e. when \( B_{a/b} > 100 \), according to the Jeffreys criterion (e.g. [Jeffreys 1961]).

As a further reference, a measure of the goodness-of-fit of the models to the observed data is performed using the \( \chi^2 \) goodness-of-fit test. The \( \chi^2 \) test statistic is determined by,

\[ \chi^2 = \sum_k \frac{(N_k - D_k)^2}{D_k}, \] (5.9)

where we are using consistent expressions for the data \( N_k \) and model \( D_k \) probabilities. The \( \chi^2 \) test statistic is then compared to a \( \chi^2 \) distribution dependent on the number of degrees of freedom. The probability (or p-value) is then the probability of obtaining a \( \chi^2 \) statistic at least as extreme as the value calculated. The null-
Results

5.5 Results

5.5.1 Parameter estimation results

Model: power-law with noise

In order to obtain the best-fitting index, $\alpha$, for the power-law model of the underlying QSO flux density, we maximise the likelihood given by Equation 5.6. This is repeated for each of the eight optical magnitude sample distributions, which will allow us to see evolution (if any) of the radio flux density distributions with optical magnitude/luminosity. The best-fitting power-law indices for each sample are shown in Table 5.1. Figure 5.4 shows the best-fitting models (power-law plus noise) for each of the eight magnitude samples in red, with the actual samples’ distributions shown in black.

The posterior probability density functions (PDFs) for each $\alpha$ variable in each of the optical magnitude samples are shown in Figure 5.5. For visualisation reasons, the posterior PDFs were made by choosing the $L_{\text{max}}$ from all $\mathcal{L}$ evaluated
5.5 Results

Figure 5.4: The best-fitting power-law plus noise models for each optical magnitude sample in red, with the distributions of the observed radio flux densities shown in black. The best-fitting parameters were determined by finding the maximum likelihood given by Equation 5.6.
5.5 Results

\begin{figure}[h]
\centering
\begin{subfigure}{0.49\textwidth}
\centering
\includegraphics[width=\textwidth]{figure5_4_e}
\caption{\(\langle m_1 \rangle = 20.44\)}
\end{subfigure} \hfill
\begin{subfigure}{0.49\textwidth}
\centering
\includegraphics[width=\textwidth]{figure5_4_f}
\caption{\(\langle m_1 \rangle = 20.69\)}
\end{subfigure}
\begin{subfigure}{0.49\textwidth}
\centering
\includegraphics[width=\textwidth]{figure5_4_g}
\caption{\(\langle m_1 \rangle = 20.89\)}
\end{subfigure} \hfill
\begin{subfigure}{0.49\textwidth}
\centering
\includegraphics[width=\textwidth]{figure5_4_h}
\caption{\(\langle m_1 \rangle = 21.13\)}
\end{subfigure}
\caption{(continued).}
\end{figure}
Table 5.1: The best-fitting power-law indices ($\alpha$) for each optical magnitude sample with their 95% confidence intervals (CIs).

<table>
<thead>
<tr>
<th>$\langle m_1 \rangle$ sample</th>
<th>$\alpha$</th>
<th>$\alpha$ 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.88</td>
<td>-1.700</td>
<td>-1.8 : -1.6</td>
</tr>
<tr>
<td>18.72</td>
<td>-1.900</td>
<td>-2.1 : -1.8</td>
</tr>
<tr>
<td>19.43</td>
<td>-2.300</td>
<td>-2.4 : -2.2</td>
</tr>
<tr>
<td>20.05</td>
<td>-2.350</td>
<td>-2.5 : -2.3</td>
</tr>
<tr>
<td>20.44</td>
<td>-2.625</td>
<td>-2.8 : -2.5</td>
</tr>
<tr>
<td>20.69</td>
<td>-2.950</td>
<td>-3.3 : -2.8</td>
</tr>
<tr>
<td>20.89</td>
<td>-2.975</td>
<td>-3.2 : -2.8</td>
</tr>
<tr>
<td>21.13</td>
<td>-3.375</td>
<td>-4.3 : -3.1</td>
</tr>
</tbody>
</table>

at each $\alpha$, and then inverted and normalised to obtain the posterior PDF. The 95% confidence intervals were evaluated by determining the limits $\alpha_1$ and $\alpha_2$, where $p(\alpha_1 \leq M : \{\alpha\} \leq \alpha_2) = 0.95$, where $M : \{\alpha\}$ is the power-law model over the range of values of the parameter $\alpha$. The posterior PDFs shown in Figure 5.5 are asymmetric and we cannot assume that the Gaussian approximation holds in these cases ($\ln(L) = \text{const.} - \chi^2/2$). Therefore an integration over the normalised posterior PDF is necessary to find the limits $\alpha_1$ and $\alpha_2$ where $\int_{\alpha_1}^{\alpha_2} p(M : \{\alpha\}|\text{data})d\alpha \approx 0.95$. The 95% confidence intervals are given in Table 5.1 and shown as red, dashed vertical lines in Figure 5.5.

Model: Gaussian with noise

The likelihood, $p(\text{data}|\text{model})$, is given by Equation 5.6 and the best-fitting parameters are again found by maximising the likelihood. Table 5.2 gives the best fitting parameters and their 95% confidence intervals.

The Gaussian model parameters for the last three optical magnitude samples are restricted by the boundaries of the parameter space modelled. The minimum width is constrained by the resolution of the distributions, while the mean is not allowed to go below zero. This is because we are modelling the QSO emitted flux, not the
Figure 5.5: The PDFs of the power-law indices $\alpha$ with the 95% confidence intervals shown as red dashed, vertical lines. For aesthetic purposes only the range of [-5,0] for $\alpha$ is shown.
Figure 5.5: (continued).
### 5.5 Results

<table>
<thead>
<tr>
<th>$m_i$ sample</th>
<th>$\mu$ (µJy)</th>
<th>$\mu$ 95% CI (µJy)</th>
<th>$w$ (µJy)</th>
<th>$w$ 95% CI (µJy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.88</td>
<td>42.9</td>
<td>0.0 : 73.5</td>
<td>183.75</td>
<td>183.75</td>
</tr>
<tr>
<td>18.72</td>
<td>49.0</td>
<td>0.0 : 75.5</td>
<td>122.25</td>
<td>122.25</td>
</tr>
<tr>
<td>19.43</td>
<td>49.0</td>
<td>42.88 : 61.25</td>
<td>61.25</td>
<td>61.25</td>
</tr>
<tr>
<td>20.05</td>
<td>36.75</td>
<td>30.63 : 42.88</td>
<td>61.25</td>
<td>61.25</td>
</tr>
<tr>
<td>20.44</td>
<td>12.25</td>
<td>6.125 : 24.50</td>
<td>61.25</td>
<td>61.25</td>
</tr>
<tr>
<td>20.69</td>
<td>0.0</td>
<td>0.0</td>
<td>61.25</td>
<td>61.25</td>
</tr>
<tr>
<td>20.89</td>
<td>0.0</td>
<td>0.0</td>
<td>61.25</td>
<td>61.25</td>
</tr>
<tr>
<td>21.13</td>
<td>0.0</td>
<td>0.0</td>
<td>61.25</td>
<td>61.25</td>
</tr>
</tbody>
</table>

**Table 5.2:** The best-fitting Gaussian parameters ($\mu$ and $w$) for each optical magnitude sample with their 95% confidence intervals (CIs). Where there is only one value for the confidence interval, only the best-fit value is above the 95% requirement.

noise which has an average of zero and negative regions.

The best-fitting models (Gaussian model plus noise) are shown in Figure 5.6 for each of the optical magnitude samples. The observed radio flux density distributions from the SDSS QSOs are shown in black, while the models are shown in red.

The best-fitting mean values for each optical magnitude sample are below the resolution of the flux density distribution (61.25 µJy). In other words, they are consistent with zero since we are not sensitive to the flux density below our resolution limit. This essentially causes each Gaussian model of the underlying QSO flux density to be reminiscent of a power-law (for flux densities > 0 µJy). For example, Figure 5.7 shows the best-fitting underlying QSO flux density model for the $\langle m_i \rangle = 17.88$ sample. The shape of this model is very similar to the power-law models fitted in Section 5.3.1.

The posterior PDFs for the Gaussian model are shown in Figure 5.8. As the optical magnitude becomes fainter, the posterior PDF hugs the bottom left corner of the parameter space as discussed earlier. The best-fitting parameters are marked with a red asterix in each figure.

The marginalised posterior PDFs for the Gaussian parameters are shown in
5.5 Results

Figure 5.6: The best-fitting Gaussian plus noise models for each optical magnitude sample in red, with the distribution of the observed radio flux densities shown in black.

(a) $\langle m_I \rangle = 17.88$

(b) $\langle m_I \rangle = 18.72$

(c) $\langle m_I \rangle = 19.43$

(d) $\langle m_I \rangle = 20.05$
\textbf{5.5 Results}

\begin{figure}[h]
\centering
\subfloat{(e) $\langle m_1 \rangle = 20.44$}{
\includegraphics[width=0.45\textwidth]{figure5_6_e}
}\hspace{1cm}
\subfloat{(f) $\langle m_1 \rangle = 20.69$}{
\includegraphics[width=0.45\textwidth]{figure5_6_f}
}

\subfloat{(g) $\langle m_1 \rangle = 20.89$}{
\includegraphics[width=0.45\textwidth]{figure5_6_g}
}\hspace{1cm}
\subfloat{(h) $\langle m_1 \rangle = 21.13$}{
\includegraphics[width=0.45\textwidth]{figure5_6_h}
}
\caption{(continued).}
\end{figure}
Figure 5.7: The best-fitting Gaussian model from \( \langle m_i \rangle = 17.88 \) sample (the optically brightest), of the underlying QSO flux density distribution. The resolution of the flux density bins is the same as the resolution in Figures 5.1 and 5.2 \((x_{\text{res}} = 61.25 \mu \text{Jy})\). Since by construction the minimum allowable flux density in this model is 0 \( \mu \text{Jy} \), the Gaussian is not permitted to go into negative fluxes, and so the Gaussians have a similar shape to the power-law models.

Figures 5.9 and 5.10 for each optical magnitude sample. These are produced by marginalising the total PDF over the other parameter in order to isolate the parameter of interest. From Figure 5.9f to Figure 5.9h, the best-fitting mean is at the lower limit of the range of \( \mu \) sampled, 0.0 \( \mu \text{Jy} \), and the marginalised PDF is very narrow around the best-fit; this is seen also in the 95% confidence intervals in Table 5.2. As discussed earlier, the lower limit on the mean flux density is constrained by the definition of the underlying QSO flux density distribution; since we are modelling the radio flux density emitted by the QSOs a negative flux density would be unphysical. In the three faintest optical magnitude samples, the mean and the width of the Gaussian model is the lower limit of each parameter, suggesting the parameter space of the Gaussian model severely truncates the PDF. The lower limit of the range in each parameter space sampled has physical motivations, making it
Figure 5.8: The two dimensional posterior PDF of the Gaussian model for each optical magnitude sample. The best-fitting parameters are shown by a red asterix in each figure. As the samples become optically fainter, the best-fitting Gaussian model parameters are severely truncated by the prior of the both the mean and the width.
Figure 5.8: (continued).
difficult to justify moving the lower bounds further in order to find a solution which is not truncated by the prior parameter space.

As the optical magnitude increases (optical luminosity decreases), the Gaussians become narrower, emulating the trend seen in the power-law models’ indices. The Gaussians are all centred on values that are within the bin resolution $x_{\text{res}}$ meaning their centres are essentially $0 \mu Jy$, thus imitating the steepening power-law models. This trend continues until the best-fitting parameters are constrained to the lower limits and hence are truncated by the bounds of the parameters.
Figure 5.9: \( \langle m_1 \rangle = 20.44 \)

\( \langle m_1 \rangle = 20.69 \)

\( \langle m_1 \rangle = 20.89 \)

\( \langle m_1 \rangle = 21.13 \)
5.5 Results

Figure 5.10: The marginalised posterior PDF of the Gaussian width $\mu$ in each optical magnitude sample. The PDFs are sharply peaked at their maximums in each sample due to the resolution of our distributions $x_{\text{res}} = 61.25 \, \mu\text{Jy}$). The six optically faintest samples show severe truncation of the width PDFs due to the minimum width value set at one bin-width (61.25 $\mu\text{Jy}$).
Figure 5.10: (continued).
5.5 Results

As is clear from Table 5.2, the width parameter PDFs are very narrow around the best-fitting values. Therefore the marginalised PDFs in Figure 5.10 show sharp peaks at the best-fitting values. The best-fitting width for Figures 5.10c to 5.10h is the lower limit of the range of parameter space, set at the resolution of the distributions. Thus the Gaussian models cannot become narrower once this limit is reached.

5.5.2 Noise as a model

The QSO and noise PDFs are shown in Figure 5.11. In the brightest \( \langle m_1 \rangle \) samples there is an obvious deviation from the noise by the QSO distribution. This difference lessens as the \( \langle m_1 \rangle \) becomes fainter, as can be seen in Figure 5.11. In order to determine whether each QSO distribution is statistically different from the noise we perform a Kolmogorov-Smirnov test between the QSO \( \langle m_1 \rangle \) samples and the noise distributions. The results are shown in Table 5.3.

<table>
<thead>
<tr>
<th>( \langle m_1 \rangle ) sample</th>
<th>D-statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.88</td>
<td>0.16</td>
<td>( 1.3 \times 10^{-28} )</td>
</tr>
<tr>
<td>18.72</td>
<td>0.12</td>
<td>( 4.3 \times 10^{-17} )</td>
</tr>
<tr>
<td>19.43</td>
<td>0.08</td>
<td>( 1.7 \times 10^{-29} )</td>
</tr>
<tr>
<td>20.05</td>
<td>0.07</td>
<td>( 1.6 \times 10^{-31} )</td>
</tr>
<tr>
<td>20.44</td>
<td>0.05</td>
<td>( 1.2 \times 10^{-17} )</td>
</tr>
<tr>
<td>20.69</td>
<td>0.03</td>
<td>( 2.3 \times 10^{-8} )</td>
</tr>
<tr>
<td>20.89</td>
<td>0.03</td>
<td>( 5.3 \times 10^{-7} )</td>
</tr>
<tr>
<td>21.13</td>
<td>0.03</td>
<td>( 2.3 \times 10^{-6} )</td>
</tr>
</tbody>
</table>

Table 5.3: The results from the Kolmogorov-Smirnov test comparing the QSO sample distributions with the noise distribution. The D-statistic in the second column is the maximum difference between the cumulative distributions. The third column contains the probability which corresponds to the probability of obtaining a value for the D-statistic as large as the one given and whether you would reject the null hypothesis; the probability that the two compared distributions are the same. The probability is very close to zero, much less than our assumed significance level of 0.01, which means the null hypothesis is rejected; the distributions are significantly different.

The Kolmogorov-Smirnov test computes the D-statistic which is the maximum...
5.5 Results

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.11.pdf}
\caption{The QSO (black) and noise (red) radio flux density probability density distributions for each \langle m_1 \rangle sample. The flux densities are binned with equal resolution in both the noise and the QSO radio flux density distributions ($x_{\text{res}} = 61.25 \ \mu\text{Jy}$). The noise differs significantly from the QSO distribution in all bins as the results from the Kolmogorov-Smirnov test show (Table 5.3).}
\end{figure}

\begin{itemize}
\item[(a)] \langle m_1 \rangle = 17.88
\item[(b)] \langle m_1 \rangle = 18.72
\item[(c)] \langle m_1 \rangle = 19.43
\item[(d)] \langle m_1 \rangle = 20.05
\end{itemize}
\[\langle m_1 \rangle = 20.44\]

\[\langle m_1 \rangle = 20.69\]

\[\langle m_1 \rangle = 20.89\]

\[\langle m_1 \rangle = 21.13\]

*Figure 5.11: (continued).*
difference between the two cumulative distributions, in our case the noise and the QSO radio flux density distributions. The probability is the probability of obtaining a D-statistic value as large as the one determined and gives the significance of the agreement between the null hypothesis that the two distributions are the same. Our assumed significance level is 0.01, and if the probability is less than this value the null hypothesis is rejected. The difference between the two distributions is easy to see already by eye in the brighter optical samples (Figures 5.11), but becomes harder to discern in the optically faintest samples.

Although we have seen the noise distribution is significantly different from the QSO flux density distributions, we need to perform the model selection since we cannot distinguish which model is preferred in terms of modelling the underlying QSO radio flux density.

5.6 Results from the model comparison

Each of our three models of the QSO flux density distributions are compared with each other to determine the preferred model. This is done using the odds ratio as described in Section 5.4.1.

The noise vs Gaussian

The Kolmogorov-Smirnov test shows the noise distribution is significantly different to the QSO flux density distribution in Section 5.5.2. However, the noise distribution should also be tested as a model for the data compared to the other models via the odds ratio (also known as the Bayes factor). The odds ratio is the ratio between the marginal likelihoods of the models since we have assumed equal priors for all the models.
The noise function is constructed empirically from random blank sky positions in the NVSS. The noise distribution can be defined as a function describing the shape of the noise distribution \( f(x, y) \) (e.g. a Gaussian), which is integrated over priors for the mean \( (x) \) and width \( (y) \); these are given by a delta function at particular values \( x_0 \) and \( y_0 \) which have been determined by the empirical distribution. This fixes the mean and width of the function \( f(x, y) \) to \( x_0 \) and \( y_0 \). Mathematically this is expressed as:

\[
\int_x \int_y \delta(x - x_0, y - y_0) f(x, y) dx dy = f(x_0, y_0). \tag{5.10}
\]

Since the marginal likelihood is the likelihood function integrated over the prior given by a delta function, the marginal likelihood is the likelihood at \( x_0 \) and \( y_0 \). Or explicitly, the marginal likelihood is given by the likelihood (e.g. from Equation 5.6) evaluated where \( D_k \) is given by the noise distribution.

The marginal likelihood for the Gaussian plus noise model which has two parameters \( (M_G(\mu, w)) \), is given by the likelihood and the priors integrated over both \( \mu \) and \( w \). The integral is then \( \int \int p(data|M_G(\mu, w))p(\mu)p(w)dw\,d\mu \). To evaluate this numerically, we set the upper and lower limits of integration in one dimension, \( \mu_u = 1219 \, \mu\text{Jy} \) and \( \mu_l = 0 \, \mu\text{Jy} \) respectively, and find the limits in the other dimension as a function of \( \mu \), i.e. the width given a value of \( \mu \) \( (w(\mu)) \). We have assumed flat priors on the parameters (equal unity when integrated over the parameter space, i.e. \( \int_{\mu_{\text{min}}}^{\mu_{\text{max}}} p(\mu)\,d\mu = 1 \) and \( \int_{w_{\text{min}}}^{w_{\text{max}}} p(w)\,dw = 1 \)). The integral can then be written as

\[
\int_{\mu_l}^{\mu_u} d\mu \int_{w(\mu)}^{w_{\text{max}}} p(data|M_G(\mu, w))\,dw \tag{5.11}
\]

and evaluate the inner integral first, then the outer integral.
Table 5.4 shows the odds ratio comparing the best-fitting Gaussian plus noise model with the noise alone, \( B_{G/N} \). The odds ratio is given in its natural logarithmic form as is standard for model comparisons.

<table>
<thead>
<tr>
<th>( \langle m_i \rangle ) sample</th>
<th>( \ln(B_{G/N}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.88</td>
<td>77</td>
</tr>
<tr>
<td>18.72</td>
<td>33</td>
</tr>
<tr>
<td>19.43</td>
<td>68</td>
</tr>
<tr>
<td>20.05</td>
<td>76</td>
</tr>
<tr>
<td>20.44</td>
<td>31</td>
</tr>
<tr>
<td>20.69</td>
<td>-1.6</td>
</tr>
<tr>
<td>20.89</td>
<td>-10</td>
</tr>
<tr>
<td>21.13</td>
<td>-23</td>
</tr>
</tbody>
</table>

**Table 5.4:** The odds ratio comparing the Gaussian plus noise and the noise alone for the model of the QSO underlying flux. The preference for model 1 over model 2 is decisive when \( \ln(B_{1/2}) > 5 \). The Gaussian models are preferred in the five optically brightest samples. However in the three optically faintest samples the preference switches to the noise indicated by the negative \( \ln(B_{G/N}) \).

The odds ratio intrinsically penalises more complex models since it involves integration over all the variable parameters in the model. In this case, the Gaussian model has two parameters while the noise has no free parameters since it is an empirical distribution. The odds ratio comparing the Gaussian plus noise models and the noise, \( B_{G/N} \), shows that the Gaussian models in the five optically brightest samples are decisively favoured over the noise; however the preference switches towards to the noise from \( \langle m_i \rangle = 20.69 \) with \( \ln(B_{G/N}) \) becoming negative. In this sample there is no strong preference for either model, but in the two optically faintest samples the noise as a model is favoured decisively over the Gaussian models.

**The noise vs power-law**

Given the results from the comparison between the Gaussian models and the noise distribution, which show that in the optically fainter samples the noise is preferred over the Gaussian, we proceed to compare the noise with the power-law mod-
Results from the model comparison

The power-law has a single varying parameter, less complex than the Gaussian model, but more complex than the noise. The marginal likelihood for the power-law involves integrating the likelihood over the parameter space for the index. The only parameter for the power-law $\alpha$, has a flat prior ($\int_{\alpha_{\text{min}}}^{\alpha_{\text{max}}} p(\alpha) d\alpha = 1$), and the marginal likelihood is,

$$p(\text{data}|M_{\text{PL}}) = \int_{\alpha_{\text{min}}}^{\alpha_{\text{max}}} p(\text{data}|M_{\text{PL}}: \{\alpha\}) d\alpha$$

(5.12)

where $\alpha \in [-5, 5]$.

<table>
<thead>
<tr>
<th>$\langle m_I \rangle$ sample</th>
<th>$\ln(B_{\text{PL/N}})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.88</td>
<td>85</td>
</tr>
<tr>
<td>18.72</td>
<td>53</td>
</tr>
<tr>
<td>19.43</td>
<td>78</td>
</tr>
<tr>
<td>20.05</td>
<td>94</td>
</tr>
<tr>
<td>20.44</td>
<td>47</td>
</tr>
<tr>
<td>20.69</td>
<td>19</td>
</tr>
<tr>
<td>20.89</td>
<td>16</td>
</tr>
<tr>
<td>21.13</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 5.5: The odds ratio comparing the power-law plus noise and the noise alone for the QSO underlying flux density model. The preference for model 1 over model 2 is decisive when $\ln(B_{1/2}) > 5$. The power-law models are preferred in all the samples.

The odds ratio comparing the power-law plus noise and the noise $B_{\text{PL/N}}$, is decisive in its favouring of the power-law model of the underlying QSO flux density distribution in all the optical magnitude samples. The power-law plus noise model is preferred to the noise alone, even after penalising for the additional parameter. Hence the data prefer the power-law plus noise model over the noise alone.

The power-law vs Gaussian

Although the power-law is favoured over the noise as a model in all of the samples, we must also compare the power-law with the Gaussian since the Gaussian model...
was preferred over the noise in the optically brightest five samples.

The marginal likelihood ratios for each optical magnitude sample are given in Table 5.6. The odds ratio is calculated as power-law/Gaussian, therefore if $B_{PL/G} > 5$, the power-law is favoured decisively.

<table>
<thead>
<tr>
<th>$\langle m_I \rangle$ sample</th>
<th>$\ln(B_{PL/G})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.88</td>
<td>7</td>
</tr>
<tr>
<td>18.72</td>
<td>20</td>
</tr>
<tr>
<td>19.43</td>
<td>10</td>
</tr>
<tr>
<td>20.05</td>
<td>19</td>
</tr>
<tr>
<td>20.44</td>
<td>15</td>
</tr>
<tr>
<td>20.69</td>
<td>21</td>
</tr>
<tr>
<td>20.89</td>
<td>27</td>
</tr>
<tr>
<td>21.13</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 5.6: The odds ratio comparing the power-law model and the Gaussian model for the QSO underlying flux density. In all optical magnitude samples the power-law is preferred.

The odds ratio for each optical magnitude sample decisively favours the power-law model for the underlying QSO flux density distribution compared to both the Gaussian and the noise alone. **This indicates the radio-quiet envelope scenario is not supported up to our sensitivity limit of 61.25 $\mu$Jy.** Below this flux density we cannot say there is not a minimum radio flux density for a given optical magnitude since we are not sensitive to flux densities below 61.25 $\mu$Jy. All of the power-law models have negative indices indicating there are increasing numbers of fainter sources, and comparatively few bright sources. In addition, since our power-laws have no cut off we expect even more faint sources below the sensitivity limit (61.25 $\mu$Jy). Therefore from our observed QSO flux density distributions there is no minimum required radio flux density for a given optical luminosity.

This being the case, the stacked detections shown in Figure 4.6 are simply the means of distributions with large ranges of values: the few sources with relatively high flux densities have led to detected mean values, but the majority of sources have
Results from the model comparison

Flux densities below these mean values. The apparent correlation between the radio luminosity density and optical luminosity density in Figure 4.6 is reflected in the steepening of the power-law indices with the increase in optical magnitude. In light of our results the stacked detection does not mean that faint sources tend to have some mean value, which could be interpreted as a lower envelope. This highlights the caveat of simply using the mean stacked value to infer a characteristic of a population. Stacking, in this case, indeed produced detections in seven of the samples. However, once we probed beyond these mean values and into the actual distributions, the results seen in Figure 4.6 were not supported, e.g. there is no radio-quiet envelope seen in our data. These results are discussed further in Section 5.8.

5.6.1 Goodness of fit tests

The odds ratio model comparison procedure does not return a measure of the goodness-of-fit of the preferred models to the observed data. As an additional reference we have calculated estimates on the goodness-of-fit of the best-fitting models. These are given by the $\chi^2$ statistic probability, the p-value, which is the probability of obtaining a $\chi^2$ statistic at least as extreme as the one calculated. For each of the models we compared to the data, the $\chi^2$ goodness-of-fit p-values are given in Table 5.7. The null hypothesis in each case is that the fit is good, with a significance level of 0.01.

The results of the $\chi^2$ goodness-of-fit test with the null hypothesis that the noise is a good fit to the QSO radio flux density distribution, show the p-values increase as the optical magnitude becomes fainter, which is seen both by eye in Figure 5.2 and from the K-S test results, although they remain well below the significance level of 0.01 throughout all the samples. Therefore the null hypothesis that the noise alone is a good model to the data is rejected at every sample.
Investigating the effect of the model resolution

<table>
<thead>
<tr>
<th>$\langle m_I \rangle$ sample</th>
<th>noise alone</th>
<th>power-law plus noise</th>
<th>Gaussian plus noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.88</td>
<td>0</td>
<td>0.023</td>
<td>0.164</td>
</tr>
<tr>
<td>18.72</td>
<td>0</td>
<td>0.065</td>
<td>0.006</td>
</tr>
<tr>
<td>19.43</td>
<td>0</td>
<td>0.053</td>
<td>0.528</td>
</tr>
<tr>
<td>20.05</td>
<td>0</td>
<td>0.076</td>
<td>0.002</td>
</tr>
<tr>
<td>20.44</td>
<td>0</td>
<td>0.031</td>
<td>0.054</td>
</tr>
<tr>
<td>20.69</td>
<td>$4 \times 10^{-6}$</td>
<td>0.192</td>
<td>0.294</td>
</tr>
<tr>
<td>20.89</td>
<td>$6 \times 10^{-7}$</td>
<td>0.065</td>
<td>0.008</td>
</tr>
<tr>
<td>21.13</td>
<td>$3 \times 10^{-4}$</td>
<td>0.018</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 5.7: The p-values from the $\chi^2$ goodness-of-fit tests for each model. Only the p-values for the power-law plus noise model exceed the significance level of 0.01 in all optical magnitude samples.

The best-fitting power-law plus noise models in all the optical magnitude samples have $\chi^2$ p-values greater than 0.01. Therefore the null hypothesis that the power-law plus noise models are good fits to the data cannot be rejected.

The $\chi^2$ goodness-of-fit p-values for the Gaussian plus noise models in four of the optical magnitude samples do not exceed the assumed significance level of 0.01 and therefore the Gaussian plus noise models for these samples are not accepted as good fits to the observed data. These are the optical magnitude samples with mean magnitudes of $\langle m_I \rangle = \{18.72, 20.05, 20.89, 21.13\}$. The other samples’ goodness-of-fit p-values are all acceptable above the significance level.

5.7 Investigating the effect of the model resolution

We have modelled the distribution of the radio flux densities of the QSO population having fixed the resolution to approximately $3\sigma$ of the optically brightest sample’s rms error. To investigate the dependence the models have on this choice of binwidth, we conduct a study where we vary the resolution between 20 - 70 mJy.
Investigating the effect of the model resolution

The lower resolution corresponds to approximately the 1σ of the rms error for the optically brightest sample. With resolutions finer than this the likelihoods calculated with Equation 5.6 reach the machine limit for small numbers due to the low numbers of observed sources in some of the flux bins.

Varying the resolution does affect the best fit values of the model parameters. This exercise was conducted to see the effect on the power-law in particular, as the resolution dictates the minimum flux value we can calculate the power-law distribution (since the power-law cannot be evaluated at zero flux, we start the from the first flux bin i.e. $x_{\text{res}}$). We notice that the best-fitting power-law indices increase with finer resolution. However they remain below $\alpha = -1.4$, and follow the same trend as when $x_{\text{res}} = 61.25 \mu Jy$ (decreasing power-law index with decreasing optical magnitude). Table 5.8 shows the power-law indices for a selection of flux resolutions and the 95% confidence intervals.

<table>
<thead>
<tr>
<th>$x_{\text{res}}$</th>
<th>$\langle m_I \rangle$</th>
<th>$\alpha$ (95% CI)</th>
<th>$\alpha$ (95% CI)</th>
<th>$\alpha$ (95% CI)</th>
<th>$\alpha$ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 µJy</td>
<td>-1.400 (-1.5:-1.4)</td>
<td>-1.475 (-1.6:-1.4)</td>
<td>-1.550 (-1.7:-1.5)</td>
<td>-1.625 (-1.8:-1.6)</td>
<td></td>
</tr>
<tr>
<td>30 µJy</td>
<td>-1.650 (-1.8:-1.6)</td>
<td>-1.725 (-1.9:-1.7)</td>
<td>-1.825 (-2.0:-1.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40 µJy</td>
<td>-1.950 (-2.1:-1.9)</td>
<td>-2.075 (-2.2:-2.0)</td>
<td>-2.175 (-2.3:-2.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 µJy</td>
<td>-2.000 (-2.1:-2.0)</td>
<td>-2.125 (-2.2:-2.1)</td>
<td>-2.225 (-2.3:-2.2)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.8: The best fitting power-law index parameter in each sample with varying flux resolution. The 95% confidence intervals are included to illustrate the effect of constructing the power-law model with a finer resolution. To be compared to the best fitting $\alpha$ values using the resolution corresponding to 3σ of the rms in the brightest optical sample (see Table 5.1).

The best fitting power-law indices are affected by the resolution chosen and this reflects the differing minimum value of flux that each power-law has. This is, in effect, forcing the power-law to turn over at the very first bin of width $x_{\text{res}}$, which may be thought of as an additional variable in the power-law model. The smaller the
binwidths, the closer this turnover is to zero, and this acts to flatten the power-law as sources which would previously have been within the sensitivity of our models are now in the first few bins. In the previous modelling, we were treating the resolution $x_{\text{res}}$ as a delta function where $p(x_{\text{res}}) = 1$ when $x_{\text{res}} = 61.25 \, \mu \text{Jy}$ and $p(x_{\text{res}}) = 0$ at all other values of $x_{\text{res}}$.

We are keeping the resolution the same with all models, the noise, power-law and noise, and Gaussian and noise, to ensure equal treatment. If we were to vary the resolution for one, we should vary the resolution for all three models we are fitting the data with. This would add one more variable parameter to each model, thereby increasing the number of dimensions each model is integrated over to obtain the evidence. Therefore we would not expect the results from the model comparisons in Section 5.4.1 to be significantly different. We reiterate here that our initial choice for $x_{\text{res}}$ was motivated by the aim to reduce the lateral uncertainty in each bin’s population.

5.8 Conclusions from the modelling

The QSO radio flux density distributions constructed for each sample allowed us to probe behind the results from the stacking in Chapter 4 shown in Figure 4.6. The stacked results hinted at the existence of a lower envelope in the radio luminosity density. The stacked means are only meaningful when the shape of the measured flux density distribution is one in which the mean is representative of all the flux densities. For example, if the underlying distribution was a Gaussian, the mean is simultaneously the median and the mode of the distribution, hence it provides a good representation of the characteristic value of the population (sample) being investigated. If, however, the stacked mean value was drawn from a distribution that was not well represented by the mean value, e.g. a power-law distribution, the
mean instead of representing the characteristic value of the sample, is influenced by
the extremes because the distribution is spread over a large range of values. In order
to determine whether the apparent lower envelope in Figure 4.6 is a real feature of
the underlying QSO flux density distributions we investigated the shape of these
distributions further in this current chapter.

The observed distributions were compared to the noise distribution, constructed
from random positions in the NVSS sky, and fit with three different models, the
empirical noise distribution alone, a power-law combined with the noise and a single
Gaussian combined with the noise. The noise distribution is shown in Figure 5.2; it resembles a positively skewed Gaussian function. Models of the underlying QSO
flux density distribution were added with the noise distribution and compared to
the observed flux density distributions of the eight optical magnitude samples. The
best-fitting parameters were determined by maximising the likelihoods and assuming
flat priors on the parameters. The best-fitting models were then compared with each
other using the odds ratio in Section 5.6.

The strong preference given by the odds ratio for the power-law models with
negative indices suggest an increasing number of QSOs with fainter radio fluxes.
Therefore we can conclude the idea of a minimum radio flux density for a given
accretion rate is not supported up to our sensitivity limit (given by $x_{\text{res}} = 61.25 \mu\text{Jy}$).
Furthermore, the shape of the power-law models for the underlying QSO flux density
distributions do indicate there may be many more sources with flux densities fainter
than the 61.25 $\mu\text{Jy}$ limit. Through our extensive modelling it has become apparent
that the mean stacked flux densities we detect in Chapter 4 are not typical of the
sources within each stacked sample, i.e. such as one would expect if the distribution
were characterised by the mean. Our modelling of the distributions of QSO flux
densities lead to the conclusion that there is no lower envelope and in fact there are
more and more sources with lower values of radio flux density.
Conclusions from the modelling

If we had only the results from the stacking shown in Figure 4.6, which found detections in all but the faintest $\langle m_I \rangle$ sample and showed increasing radio luminosity density with increasing optical luminosity density, it would have led to incorrect conclusions. However, we have shown from the modelling of the distributions behind the mean stacked results that the models do not support the lower envelope scenario. This is an important illustration of the possibility of misinterpreting the results obtained from stacking. Relying on the stacked mean (or median) measurement alone and physically interpreting these results assumes, inherently, that the underlying distribution is characteristically represented by the mean. If the mean does not provide a good measure of the characteristic value in a distribution, the actual shape of the distribution must be taken into account to avoid misinterpreting the stacked results.

The resolution of the modelling is fixed for all models which affects the power-law in particular as it is evaluated from the sensitivity limit to the maximum model flux (i.e. not from zero). We investigated how varying the resolution of the modelling, e.g. increasing the sensitivity limit, affects the power-laws’ best fit indices. Decreasing the bin size increases the best fit indices for every sample. However the indices remain below -1.4 and with decreasing optical magnitude, the power-law index decreases also. Table 5.8 shows a sample of different binwidths and the corresponding best fitting power-law models for each sample. The 95% confidence intervals for the power-law index parameters are displayed, and show that once the resolution reaches $x_{\text{res}} = 40 \mu$Jy the confidence intervals agree with the original choice for the binwidth ($x_{\text{res}} = 61.25 \mu$Jy).

Interestingly, as the apparent optical magnitude increases, the slope of the power-law becomes steeper (regardless of the resolution of the models, see Tables 5.1 and 5.8). This suggests that although the distribution of radio flux densities within each sample can be described by a single power-law plus the noise distribution, across the
samples the index of the power-law varies systematically. This trend is witnessed in the relation between the radio and optical luminosity densities in Figure 4.6. This warrants further investigation and we can use the power-law models of the underlying QSO radio flux density to investigate the evolution of the radio loudness with optical luminosity for our radio-undetected QSO sample. The next chapter describes this aspect of our investigation, and evaluates our results in the context of theoretical models of jet production in AGN.
Chapter 6

Radio loudness parameter: distribution and modelling

In this chapter we use the best-fitting power-law models from Chapter 5 for the radio flux density distributions of our SDSS QSOs and determine the distribution of the radio loudness parameter. In each power-law model the same radio flux density range is used, which allows us to investigate evolution of the radio loudness parameter with respect to each optical luminosity density range. We model the radio loudness parameter distribution with three different models: a single power-law, a Schechter function and a double power-law. Similar to Chapter 5, we use the odds ratio to determine the most preferred model. We convert the radio loudness to the jet efficiency assuming a radiative efficiency and estimate the minimum jet efficiency we observe is \( \sim 5 \times 10^{-4} \). With this estimate we can constrain recent simulated models of the jet efficiency.

6.1 Motivation

Modelling the distribution of the radio flux densities of the radio-undetected SDSS QSOs in Chapter 5 has revealed that the underlying QSO emitted radio flux densities are best modelled by a single power-law with negative indices. The implication of this result is that despite the detections in the stacked samples, which alone
may have been interpreted as pointing towards the existence of a lower envelope, the modelling showed the vast majority of QSOs have decreasing radio luminosities.

The relation of the radio luminosity density with optical luminosity density shown in Figure 4.6 reflects the power-law models’ indices steepening as the sample optical luminosity densities decrease.

The NVSS detected SDSS QSOs in the redshift range $0.3 \leq z_{\text{phot}} \leq 0.5$ are also shown in Figure 4.6. The radio luminosity densities span several dex for a given optical luminosity density, which agrees with the well established observation that sources of a given bolometric luminosity (accretion rate) have radio luminosities spanning $\sim 3$ dex (e.g. Kellermann et al. 1989; Sikora, Stawarz & Lasota 2007; McNamara, Rohanizadegan & Nulsen 2011). The process producing the radio emission in QSOs must be able to account for both the radio quiet and the loudest radio sources which infer the efficiency converting the accretion energy into radio jets is of the order of unity (Fernandes et al. 2011; McNamara, Rohanizadegan & Nulsen 2011).

A jet production mechanism that can produce the large range of efficiencies and explain the observations is needed. The spin of the black hole (represented as $\hat{a}$ where $0 \leq |\hat{a}| \leq 1$) provides a mechanism which can explain the variation in jet powers. The revised spin paradigm is a hybrid model in which the jet energy is extracted from the accretion disc as well as a rotating black hole via the Blandford-Znajek mechanism (e.g. Blandford & Znajek 1977; Punsly & Coroniti 1990; Wilson & Colbert, 1995; Meier, Koide & Uchida 2001; Sikora, Stawarz & Lasota 2007). With the advance in computing capabilities, there have been a number of numerical simulations incorporating $\Lambda$CDM simulations and semi-analytical models which have found the spin distribution to be bimodal, with a population at $\hat{a} \sim 1$ and one at $\hat{a} \sim 0.1$ (Lagos, Padilla & Cora 2009; Fanidakis et al. 2011; Martinez-Sansigre & Rawlings 2011). In addition to the final spin distributions, Lagos, Padilla & Cora
Motivation

(2009) and Fanidakis et al. (2011) have investigated how the processes growing the black hole affect the final spin and the radio power.

Fanidakis et al. (2011) found the accretion modes were responsible for the two spin distributions. The lower spins fit chaotic accretion where episodes of accretion onto the black hole with random angular momentum orientations act to spin down the central black hole (i.e. King et al., 2005; King, Pringle & Hofmann, 2008). While the higher spins were found to fit either the prolonged (where sustained accretion with angular momentum aligned with the black hole’s angular momentum acts to spin up the black hole), or chaotic accretion scenarios coupled with mergers of black holes producing the high spin merged black hole.

Lagos, Padilla & Cora (2009) and Fanidakis et al. (2011) found that cold accretion, following an episode of hot accretion at a higher Eddington ratio, produced the most powerful jets. In these simulations the production of radio power is dependent on a combination of the spin of the black hole and the accretion (including the geometry of the disk and how the mass is accreted onto the black hole). Observations and simulations show it is the highest mass black holes that are capable of the highest radio powers. The simulations were able to fit the radio loud ends of the RLFs well, while the fit to the radio quiet end of the RLF had less success. Radio quiet emission remains hard to constrain due to the difficulty in accumulating a large population of detected RQQ. The low radio emission of each source requires much lower sensitivities, therefore longer observation times, limiting the success of wide-field surveys with these objects.

Martinez-Sansigre & Rawlings (2011) investigated the distribution of the black hole spin given six different simulations of jet efficiency from the literature. They determined the best model of spin distribution to be a double Gaussian; the bimodality in the spins has been found by comparing the simulated RLFs produced by the different efficiency distributions to the observed RLF from Best & Heckman
They found that a bimodal distribution of spin parameters was preferred, where the locus of one population was located at or near maximal spin, and the other population had much lower spins, e.g. close to zero. It is important to note however, the observations they used to constrain their models were more sensitive to the higher spin population.

Here is where studies of optically detected QSOs matched with radio surveys can contribute to our understanding of this vast, yet elusive population of RQQs. The detections we obtained from stacking illustrate that there are at least some of these sources emitting in the radio wavelength observed. Therefore, the distribution of the radio flux densities of these sources (which were significantly different from the noise of the NVSS) provides information on the radio emission from these objects. Modelling this radio emission we have determined that they are best described by a negative indexed power-law function, and it is these model distributions we can use to try and model the distribution of the radio loudness parameter in these radio undetected QSOs.

**6.1.1 Radio loudness parameter**

The radio loudness parameter $R$ is often used to characterise the radio emission from quasars. It is given by the ratio of the radio and optical luminosities (e.g. Kellermann et al. [1989]),

$$R = \frac{L_{\text{rad}}}{L_{\text{opt}}}.$$  \hspace{1cm} (6.1)

The $R$ parameter has been investigated for different populations of quasars with the distributions showing alternatively bimodality between radio-loud and radio-quiet or a smoother transition between the two (see Kellermann et al. [1989], Ivezić et al. [2002], Cirasuolo et al. [2003b], Sikora, Stawarz & Lasota [2007], for example).
6.2 Investigating the distribution of the radio-loudness parameter

Ivezić et al. (2002) outlines the definition of radio loudness and the case where the $\mathcal{R}$ parameter is meaningful in their Appendix C; $\mathcal{R}$ is only meaningful if the optical and radio emissions are correlated. The existence of an intrinsic bimodality in the radio loudness parameter of quasars remains a contentious issue with studies recognising the importance of selection effects on the shape of the distribution (e.g. Sikora, Stawarz & Lasota 2007).

We will use the distributions of radio flux density we determined in Chapter 5 and convert the radio flux densities we modelled into radio loudness values $\mathcal{R}$ to determine the probability distributions of $\mathcal{R}$. The models of the distribution of the radio loudness parameter can then be used in the context of the jet efficiency, $\eta$, by converting $\mathcal{R}$ to $\eta$ assuming a radiative efficiency ($\epsilon$).

6.2 Investigating the distribution of the radio-loudness parameter

We will model the distribution of the radio loudness parameter from our measured flux densities in all eight optical magnitude samples. We will use the power-law models determined in Chapter 5 collectively to give the distributions of the radio loudness. Modelling the radio loudness parameter distribution may be able to help constrain theoretical jet production models at the faint end of the radio luminosity. We need to convert our radio flux densities to $\mathcal{R}$ first, to build the distribution given by our data.

We can calculate the radio loudness $\mathcal{R}$ via Equation 6.1 from the radio flux densities and optical luminosity densities. Historically, the B-band optical luminosity was used to determine $\mathcal{R}$ (Kellermann et al. 1989), while there have been a range of radio frequencies used, for example 5 GHz, 1.4 GHz and below (Kellermann et al.)
Investigating the distribution of the radio-loudness parameter

The SDSS I-band at a redshift of 0.4 approximately corresponds to B-band in the rest-frame. The flux densities are converted to luminosity densities using Equation 4.3. This is the B-band luminosity in the rest-frame at the redshifts we are probing.

Since $R$ has been determined with a variety of radio frequencies, and more recently the $\nu = 1.4$ GHz observations from the large scale surveys being used, we have kept our 1.4 GHz radio flux densities. This ensures no added systematic error in our estimate of $R$ as we would assume that the radio flux density follows the power-law relation with frequency $S \propto \nu^{-\alpha_R}$ with $\alpha_R = 0.8$ and we are in no position to test whether this assumption holds for our sample of QSOs since they are undetected in radio.

6.2.1 The $R$ parameter of the mean stacked radio flux densities

The stacked radio flux density detections and the mean B-band optical luminosity densities give the $R$ values shown in Table 6.1. Figure 6.1 shows the distribution with mean optical B-band luminosity density and this shows a scattered distribution of $R$. Note that since we are defining $R$ as the ratio of $L_{2.5 \text{ GHz}}/L_B$ the empirical division between radio-quiet and radio-loud ($R \geq 10$ defined by Kellermann et al. (1989) for 5 GHz data) is shifted to 30-40 assuming a spectral index of $\alpha_R = 0.8$.

Table 6.1 gives the $R$ values for the optical magnitude samples. None of the samples satisfy the radio loud criterion set out in Kellermann et al. (1989) ($R \geq 30 \sim 40$ for $\nu_{\text{rad}} = 1.4$ GHz). Other studies have found that $R$’s uncorrelated with optical luminosity density (Kellermann et al. 1989; Cirasuolo et al. 2003b). In our stacked results, there appears to be a break in the trend of $R$ versus optical luminosity density at $\log_{10}(L_B) \sim 22$, which corresponds to $\log_{10}(R) = 0.8$ (Figure
6.2 Investigating the distribution of the radio-loudness parameter

\[
\langle m_I \rangle \text{ sample} \quad \log_{10}(L_{2.5 \, \text{GHz}}/[\text{WHz}^{-1}\text{sr}^{-1}]) \quad \log_{10}(L_B/[\text{WHz}^{-1}\text{sr}^{-1}]) \quad R
\]

<table>
<thead>
<tr>
<th>\langle m_I \rangle</th>
<th>\log_{10}(L_{2.5 , \text{GHz}}/[\text{WHz}^{-1}\text{sr}^{-1}])</th>
<th>\log_{10}(L_B/[\text{WHz}^{-1}\text{sr}^{-1}])</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.88</td>
<td>22.98</td>
<td>23.01</td>
<td>0.92</td>
</tr>
<tr>
<td>18.72</td>
<td>22.82</td>
<td>22.61</td>
<td>1.62</td>
</tr>
<tr>
<td>19.43</td>
<td>22.60</td>
<td>22.33</td>
<td>1.84</td>
</tr>
<tr>
<td>20.05</td>
<td>22.54</td>
<td>22.08</td>
<td>2.91</td>
</tr>
<tr>
<td>20.44</td>
<td>22.39</td>
<td>21.92</td>
<td>2.98</td>
</tr>
<tr>
<td>20.69</td>
<td>22.16</td>
<td>21.82</td>
<td>2.21</td>
</tr>
<tr>
<td>20.89</td>
<td>22.10</td>
<td>21.74</td>
<td>2.31</td>
</tr>
<tr>
<td>21.13</td>
<td>21.90</td>
<td>21.64</td>
<td>1.83</td>
</tr>
</tbody>
</table>

Table 6.1: The stacked mean radio and optical luminosity densities for the eight optical magnitude samples. The \( R \) parameter is the ratio of the radio to optical luminosity densities (Equation 6.1).

6.1. It should be noted that the cutoff in the radio flux density of the sources in our sample will set upper limits on the mean \( R \) values, especially in the higher optical luminosity samples. This is due to the exclusion of sources with radio flux densities above \( 3 \sigma_{1\text{rms}} \) in our stacked samples.

Using the photometric quasar catalogue, we are able to collate the largest sample of undetected, radio-quiet QSOs. Using the method of stacking, we have detected the mean radio flux densities of these QSOs in samples of increasing optical luminosity densities. Therefore our mean \( R \) parameter values are the result of the largest population of radio faint QSOs gathered in this way to date.

6.2.2 The distribution of the \( R \) parameter given the underlying QSO flux density models

The stacked mean detections in radio and optical luminosity densities give \( R \) parameters that appear to have a break at the optical luminosity density of \( \log_{10}(L_B) \sim 22 \). However, as we ascertained in Chapter 5, the mean stacked values do not give an accurate portrayal of the underlying radio flux density distribution within the population of the stacked samples. Therefore the models of the underlying QSO flux
Investigating the distribution of the radio-loudness parameter

Figure 6.1: The $R$ parameter values for the mean stacked radio luminosity densities and the mean B-band optical luminosity densities. The $R$ parameter is plotted against the B-band luminosity densities. The $R$ parameter distribution appears to “break” at around $\log_{10}(L_B) \sim 22$. The horizontal error-bars represent the range of B-band optical luminosity densities. The vertical error-bars are the uncertainty of the $R$ parameter incorporating the spread in optical luminosities, the photometric redshift uncertainty and the radio flux density measurement uncertainty.

which were fit in Chapter 5 are used to investigate the distributions of $R$. The range of flux densities our underlying QSO flux density models covered was from 0 $\mu$Jy to $5\sigma_{I_{\text{rms}}}$ (where $\sigma_{I_{\text{rms}}} = 0.45$ mJy). In order to model the distribution of the radio loudness we must convert the range of flux densities used in the modelling, to $R$ using Equation 6.1 and the mean B-band optical luminosity density for each sample. The mean optical luminosity has associated uncertainties including the uncertainty in the photometric redshift and the variance; we note here that accounting for these uncertainties extends the range in radio loudness that each optical magnitude sample covers.

Figure 6.2 displays the distributions of the radio loudness parameters while using the underlying QSO flux density models. Each optical magnitude sample is a
6.2 Investigating the distribution of the radio-loudness parameter

different colour for identification. The slopes are given by the best-fitting power-law indices given in Table 5.1 ranging from -1.7 to -3.375. They have been re-normalised such that \( \int p(R) \Delta R = 1 \) where \( \Delta R \) covers some subspace of \( R \), in this case \( \Delta R = 1 \).

This has the effect of bringing each slope into the same range and lining up with the others. The essential feature of the distributions of the radio loudness is the negative indexed power-law, i.e. there are more sources in the lower \( R \) end of each optical sample. Note the much wider span in \( R \) when using the range in flux density we modelled in Chapter 5 compared to the mean values from the stacked images.

![Figure 6.2:](image)

**Figure 6.2:** The power-law model distributions of the QSO underlying radio flux density, renormalised to \( \int p(R) \Delta R = 1 \) where \( \Delta R = 1 \). The optically brightest \( \langle m_I \rangle = 17.88 \) is the top, black line; the blue line is for \( \langle m_I \rangle = 18.72 \) sample; the brown line is the \( \langle m_I \rangle = 19.43 \) sample; the yellow line is for the \( \langle m_I \rangle = 20.04 \) sample; the forest green line corresponds to the \( \langle m_I \rangle = 20.44 \) sample; the cyan line corresponds to the \( \langle m_I \rangle = 20.69 \) sample; the purple line corresponds to the \( \langle m_I \rangle = 20.89 \) sample; and the optically faintest sample \( \langle m_I \rangle = 21.13 \) is shown by the magenta line. The lines represent the slopes of the power-law best-fitting models given in Table 5.1

The shape of the whole population’s distribution in radio loudness, incorporating all optical magnitude samples, shows a break in the distribution at around \( \log_{10}(R) = 1 \). This is very close to the hint of a break in Figure 6.1 at \( \log_{10}(L_B) \sim 22 \)
Investigating the distribution of the radio-loudness parameter

which has $\log_{10}(R) = 0.8$. The shape of the distribution resembles that of a luminosity function. Therefore, to parameterise the distribution of $R$ we fit three models to it that are motivated by the shape shown in Figure 6.2. The three models are: a single power-law, a Schechter function, and a double power-law distribution.

6.2.3 Modelling the distribution of $R$

The $R$ parameter distribution is shown in Figure 6.2. We are using the best models from Chapter 5, power-laws with negative indices so there are a greater number of sources with low radio loudness parameters. However, during the construction of the $R$ values we kept the radio flux luminosity densities constant in each sample thus the shift in $R$ is solely due to the varying optical luminosity density. In essence, we are modelling the evolution of the distribution of the radio loudness with optical luminosity. At a cursory glance, the shape of the distribution appears to be very similar to the optical luminosity function (OLF) of quasars at $z \sim 0.4$ (e.g. Croom et al., 2009). From this initial observation, one would expect a model of the same form as the OLF to fit the distribution well.

We fit three models to the $R$ distributions. These are: a single power-law, a Schechter function and a double power-law function. Given the shape of the distribution in Figure 6.2, a single power-law should not provide a good fit to the data since there appears to be a break in $p(R)$ around $\log_{10}(R) = 1$. However, rather than make this decision from the figure, we allow the data to tell us whether a single power-law model provides a reasonable description of the data. The Schechter and double power-law functions are used when modelling luminosity functions and are characterised by power-law indices and break values of the independent parameter where the different components of the functions come into importance. Figure 6.2 shows the distribution of $R$ appears to follow approximately this form, two different functions separated by a break $R^*$. However, these two models have more variable
Investigating the distribution of the radio-loudness parameter

parameters for which they will be penalised when we compare the three models using the odds ratio (Section 5.4.1)

The first model we fit is the power-law (model A) with a single varying parameter, the index which we denote $S_A$ (slope $A$). The prior probability distribution for this parameter is flat in the range of $-5 \leq S_A \leq 0$. The power-law is constrained to negative slopes (see Chapter 5). Thus model A is given by

$$p(R) \propto R^{S_A}, \quad (6.2)$$

where the constant of proportionality is simply a normalisation term.

The second model we fit is a Schechter function (model B) which takes the form [Schechter, 1976]:

$$p(R) \propto \left(\frac{R}{R^*}\right)^{S_B} e^{-R/R^*}, \quad (6.3)$$

where $R^*$ is the break value of the radio loudness parameter. The range of values for $S_B$ (slope $B$) is the same as for model A ($-5 \leq S_B \leq 0$) and for $R^*$ the range of allowed values is $-0.4 \leq \log_{10}(R) \leq 2.4$. Note this is not the whole range of radio loudness values modelled over since when $R^*$ approaches the minimum $R$ value the transition to the exponential component of the Schechter function happens earlier, and $p(R)$ goes to zero at high values of $R$. This becomes a problem when we are taking the logarithm of the function in order to compare it to our distribution in Figure 6.2. Initial fitting of the best parameters for this model were done allowing $R^* = R$, and the best-fitting value for $R^*$ was not near the minimum $R$ value we modelled over.

The double power-law model (model C) has three variables; the slope of the first power-law $S_{C1}$, the slope of the second power-law $S_{C2}$, and the break value $R^*$. 
The function takes the form:

\[ p(R) \propto \left\{ \left( \frac{R}{R^*} \right)^{-\text{SC1}} + \left( \frac{R}{R^*} \right)^{-\text{SC2}} \right\}^{-1}. \] (6.4)

The range of values for SC1 and SC2 are from \{-5, 0\}. The first power-law is significant when \( R > R^* \), while the second power-law is significant when \( R < R^* \). The range of \( R^* \) is the same as for the Schechter function (model B).

### 6.2.4 Parameter estimation

In order to determine the best-fitting model parameters, we construct the three models A, B and C for each combination of their variable parameters and then compare these to the data (where the data in this case are the slopes from the modelling of the radio flux density distributions in Chapter 5). Each optical magnitude sample spans a wide range \( R \) values, which is due to the range of radio flux densities we used in the modelling of the underlying QSO flux density in Chapter 5.

To compare the three models to the data, we determined the average slope each model gives within the range of \( R \) values for each optical magnitude sample. The average slopes are then compared to indices of the power-law models for the underlying QSO flux in each optical magnitude sample (\( \alpha \)). Figure 6.3 shows the indices from the power-law models determined in Chapter 5. The error bars on the indices are from the 95% confidence intervals calculated in Section 5.5.1 while the horizontal bars show the range of \( R \) values for each optical magnitude sample. The 95% confidence intervals are not symmetric about the best-fitting values for \( \alpha \) and as such, we have to account for this when determining the likelihood of our \( R \) model slopes, e.g. for model A: \( p(\alpha|\text{model A}) \). To do do this we follow the empirical parameterisation of the likelihood with asymmetric errors given in Barlow (2004). The log likelihood takes the form of a variable Gaussian:
6.2 Investigating the distribution of the radio-loudness parameter

Figure 6.3: The values of $\alpha$, the power-law indices from the modelling of the underlying QSO radio flux density distributions, vs the radio loudness parameter determined by the ratio of the radio to optical luminosity densities. The vertical error bars are the 95% confidence intervals from the parameter estimation in Section 5.4. The horizontal bars indicate the range of $R$ each optical magnitude sample covers. The red line is given by the best-fitting slopes in each optical magnitude sample of model A determined by finding the average slope of model A in each optical magnitude sample’s range of $R$. The green line is given by the best-fitting slopes of model B, and the blue line is the best-fitting slopes from model C. The steepest $\alpha$ corresponds to the faintest optical magnitude sample which was not detected in the stacking procedure in Chapter 4 (i.e. 2$\sigma$).

$$\ln(p(\alpha|x)) = \sum_n \frac{-1}{2} \left( \frac{(x_n - \alpha)^2}{\sigma(\alpha)} \right),$$

where $x$ is the slope from the model, and $\alpha$ is the data to be fitted. The standard deviation depends on the value of $\alpha$ and is given by:

$$\sigma(\alpha) = \sigma + \sigma'(\alpha - x)$$

and

$$\sigma = \frac{2\sigma_u\sigma_l}{\sigma_u + \sigma_l} \quad \sigma' = \frac{\sigma_u - \sigma_l}{\sigma_u - \sigma_l}$$

and $\sigma_u, \sigma_l$ are the upper and lower confidence intervals around $\alpha$. The best fitting
parameters for each model are determined by maximising the likelihood.

The best-fitting index for model A (single power-law) was $SA = -2.3$. Figure 6.4 shows the posterior PDF for model A.

![Figure 6.4: The posterior probability density function for the single power-law model, A.](image)

Figure 6.5 shows the PDF for model B, the Schechter function, with the best-fitting slope $SB = -1.7$ and the break radio loudness value $\log_{10}(R^*) = 1.35$ shown on the plot as the red asterisk.

The double power-law model (C) has the best-fitting first index as $SC1 = -4.1$, second index $SC2 = -1.7$ and break radio loudness $\log_{10}(R^*) = 1.225$. Figure 6.6 shows the marginalised PDF for the indices and the break radio loudness. The marginalised PDFs for the slopes SC1 and SC2 have come out being the same, and neither of them are peaked near the best-fit values of the maximum likelihood.

The best-fitting models are shown in Figure 6.7; the top figure shows the models by themselves, while the bottom figure shows the models and the slopes from the
Figure 6.5: The PDF of the Schecter function model, B. The best-fitting index and break radio loudness are shown by the red asterisk. The contour lines show the 10%, 30%, 50%, 70%, and 90% contours from the maximum.

QSO flux distribution models.
Investigating the distribution of the radio-loudness parameter

Figure 6.6: The marginalised PDFs of the parameters of model C.
(c) Marginalised PDF of $R^*$

Figure 6.6: (continued).
Investigating the distribution of the radio-loudness parameter

(a) The best-fitting models of the $\mathcal{R}$ distribution

(b) The best-fitting models and the slopes from Figure 6.2

**Figure 6.7:** The three models of the radio loudness parameter distribution. The red, dash-dot-dot line shows the simplest of the three models, the single power-law. While the green, dashed line corresponds to the Schecter function model. The blue, dash-dot line is the double power-law model. The black lines in the right hand figure are the slopes from the underlying QSO flux density distribution models shown in Figure 6.2.
6.2.5 Model selection

The marginal likelihood for each model is calculated by integrating over the likelihood and prior as described in Section 5.4.1. When multiple dimensions need to be marginalised over we follow the method outlined in Section 5.6. Table 6.2 gives the natural logarithm of the odds ratios for each of the model comparisons. An odds ratio value over 5 is considered to be a decisive preference. The results from the odds ratio model selection show that the single power-law (model A) is decisively disfavoured over the models B and C. Neither model B or model C are decisively favoured over the other, although the double power-law function (model C) is slightly favoured over the Schechter function (model B) with an odds ratio of 1.8.

<table>
<thead>
<tr>
<th>model comparison</th>
<th>(\ln B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>model B/model A</td>
<td>5.6</td>
</tr>
<tr>
<td>model C/model A</td>
<td>7.3</td>
</tr>
<tr>
<td>model C/model B</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Table 6.2: The odds ratio comparing the models for the distribution of the radio loudness. Model A is the single power-law, model B is the Schecter function and model C represents the double power-law model. The preference for model 1 over model 2 is decisive when \(\ln(B_{1/2}) > 5\). The power-law (model A) is strongly ruled out compared to the other two models. Of the remaining models (B and C) neither are strongly preferred over the other.

It is interesting to note that the double power-law model indices are similar to the QSO optical luminosity function values (e.g. Boyle et al. 2000; Croom et al. 2009). This is not entirely unexpected given that only the optical luminosity is changing when we model the radio flux density distribution of the QSOs. The radio flux density range was the same for each of the optical magnitude samples. We kept the \(p(R)\) equal to the probability distribution of the radio flux density since \(R\) is the ratio of the radio and optical luminosity densities and we assume that each optical magnitude sample is represented by their average optical luminosity...
6.2 Investigating the distribution of the radio-loudness parameter

densities. The Schechter function power-law index is also similar to QSO optical luminosity functions.

The break radio loudness is more difficult to compare directly to the luminosity break since the radio loudness is the ratio of the radio to optical luminosity densities. However, a crude estimate can be made comparing the break radio loudness with the break magnitude in the quasar OLF $M_G^* \sim -22.2$ (see Table 2 in Croom et al. [2009]). From this we can roughly estimate the optical luminosity density in the B-band (assuming $\alpha_\nu = -0.44$ as from Figure 1.2 from Vanden Berk et al. [2001]), of $\log_{10}(L_B) \approx 22.5$. We can estimate the optical luminosity density that $R^*$ corresponds to using the radio luminosity densities in Table 6.1. If we assume the typical radio luminosity density is $L_{1.4\,\text{GHz}} \sim 22.8$, and $\log_{10}(R^*) = 1.48$ (since the $\log_{10}R^*$ values from the Schechter function and double power-law models are 1.575 and 1.375 respectively), we estimate the break optical luminosity density to be $\log_{10}(L_B) = 21.3$. This is an order of magnitude below the break luminosity from Croom et al. [2009]. However, given the rough estimation of both the break optical luminosity of the quasar OLF from the G-band absolute magnitude and the estimate on the characteristic radio luminosity density, the comparison between the break optical luminosities suffers some uncertainty. If the radio luminosity were larger, this would lead to a larger optical luminosity given the same $R^*$.

The shape of the preferred models (B and C) indicate there are an increasing number of radio-quiet objects with decreasing $R$ values. There was no hint of a turn over in the radio flux density distributions up to our sensitivity limit (61.25 µJy in Chapter 5), and this is also reflected in the models of the radio loudness distribution. The minimum mean $R$ value we observe is $\log_{10}(R) = 0.1$, i.e. from the stacked average radio flux density, which although it does not portray the underlying distribution of the sources within the stacked sample, it is the observed radio flux density. Our models span $R$ values as low as $\log_{10}(R) \sim -0.7$. 
6.2 Investigating the distribution of the radio-loudness parameter

Cirasuolo et al. (2003a) and Baloković et al. (2012) have done Monte Carlo simulations of simulated radio properties of different QSO populations, and both probe $R$ values less than our minimum. They incorporated models that had the radio and optical emission independent and models where there was a relationship between the two (i.e. $R$). In both studies, they found the best fitting models were the ones where the optical and radio emission was related and the observed $R$ distribution is best modelled with a double Gaussian distribution with the majority of sources being radio quiet. The peak of the radio quiet Gaussian is centred on $\log_{10}(R) = -0.5$ in the Cirasuolo et al. (2003a) model with a turnover occurring at $\log_{10}(R) = -1$ (see their Figure 4). In the best-fitting model from Baloković et al. (2012) the peak of the radio quiet Gaussian lies at $\log_{10}(R) = -0.11$ and their model presented in Figure 6 shows a turnover at $\log_{10}(R) \sim -0.2$. Our models do not allow for a turn over at lower $R$s since they are motivated by the shape of the distribution of the underlying QSO flux density that had a sensitivity limit of $\sim 60 \mu$Jy. Our models of the distribution of radio loudness parameters in our radio-undetected QSO sample are not incompatible with the simulations from Cirasuolo et al. (2003a).

Perhaps the most interesting result is the shape of the distribution of the $R$ parameter. We have held the range of the radio luminosity densities fixed for each sample (although the mean radio luminosity density changes), and varied only the mean stacked optical luminosity density in each. In addition, the mean radio luminosity densities vary less than the mean optical luminosity densities, meaning the varying optical luminosity densities have a greater effect. The resulting distribution resembles the optical luminosity function. If we were to keep the optical luminosity fixed and vary only the radio luminosity, we would expect recover a shape of the distribution of $R$ similar to the radio luminosity function under the premise that the radio and optical luminosities are independent. For example, the probability of having a radio loudness value $p(R)$ is simply the product of the probability of
having a certain optical luminosity with the probability of having a certain radio luminosity. In our data, the distribution of the radio loudness parameter is exclusively a function of optical luminosity density since we are essentially keeping the radio luminosity density fixed. \textbf{There seems to be nothing fundamentally special about the radio loudness; the more luminous objects in the optical are rarer, as are the more luminous radio objects.} Hence objects that are very luminous in both are particularly rare (e.g. RLQs).

In the following section, we convert the radio loudness to jet efficiencies to relate this work to the recent simulations of the radio jet efficiencies and spin parameters of quasars.

\section{Discussion}

\subsection{Jet efficiencies}

In order to relate the models of the radio loudness distributions to the spin distributions from \cite{Martinez-Sansigre & Rawlings(2011)}, we convert $R$ to jet efficiencies. To obtain the jet efficiencies we relate the parameters of the radio loudness to parameters which determine the jet efficiency $\eta$. Firstly we consider the bolometric luminosity which we assume can be given by a fraction of the energy available from the accretion rate:

\begin{equation}
L_{\text{bol}} = \epsilon \dot{m} c^2, \tag{6.8}
\end{equation}

where $\dot{m}$ is the accretion rate, $c$ is the speed of light and $\epsilon$ represents the radiative efficiency. The radiative efficiency is a function of the black hole spin due to its dependency on the binding energy of the innermost stable circular orbit in the pre-
Discussion

The bolometric luminosity can be obtained from the observed optical luminosity density.

The other component of the radio loudness parameter is the radio luminosity density which we have assumed is due to synchrotron radio emission from a jet (Section 4.3.3). We make the approximation that the jet power \( Q_{\text{jet}} \) can be modelled by,

\[
Q_{\text{jet}} = \eta \dot{m} c^2, \tag{6.9}
\]

where \( \eta \) is the jet efficiency. We assume that \( \eta \) is a function of spin only, and additionally is a monotonic function of spin.

Equations 6.8 and 6.9 both relate the energy available by the accretion to the observables. Balancing these equations we can isolate \( \eta \):

\[
\eta = \frac{Q_{\text{jet}} \epsilon}{L_{\text{bol}}}. \tag{6.10}
\]

In the standard thin disc model (i.e. [Shakura & Sunyaev 1973]), \( \epsilon \) lies within the range of \( 0.057 - 0.43 \). We adopt \( \epsilon = 0.057 \) since we are dealing with radio-quiet QSOs. We are assuming our sources have minimal spin and thus assign them a spin parameter value \( \dot{a} = 0.1 \) (the assumed value of \( \epsilon \) is justified further by Figure 1 of [Martinez-Sansigre & Rawlings 2011], which shows that at spins of less than \( \dot{a} \sim 0.6 \) the radiative efficiency changes very little).

To convert the observed optical luminosity density \( L_B \) to \( L_{\text{bol}} \), we use the bolometric correction from [Hopkins, Richards & Hernquist 2007] \( C_{\nu_B} \), and

\[
L_{\text{bol}} = C_{\nu_B} \nu_B L_B, \tag{6.11}
\]
where $\nu_B$ is the frequency of the B-band. The converted bolometric luminosities from the mean B-band optical luminosity densities are given in Table 6.3.

<table>
<thead>
<tr>
<th>$\langle m_i \rangle$ sample</th>
<th>$L_{bol}$ (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.88</td>
<td>38.88</td>
</tr>
<tr>
<td>18.72</td>
<td>38.51</td>
</tr>
<tr>
<td>19.43</td>
<td>38.27</td>
</tr>
<tr>
<td>20.05</td>
<td>38.04</td>
</tr>
<tr>
<td>20.44</td>
<td>37.91</td>
</tr>
<tr>
<td>20.69</td>
<td>37.82</td>
</tr>
<tr>
<td>20.89</td>
<td>37.76</td>
</tr>
<tr>
<td>21.13</td>
<td>37.67</td>
</tr>
</tbody>
</table>

Table 6.3: The bolometric luminosities for the eight optical magnitude samples. The bolometric luminosities are converted from the mean B-band luminosity densities via Equation 6.11.

The jet power can be related to the observed radio luminosity density using the conversion from Miller, Rawlings & Saunders (1993) and Willott et al. (1999),

$$Q_{jet}[W] = 3 \times 10^{38} f^{3/2} \left( \frac{L_{\nu_{151}}}{10^{28} \text{WHz}^{-1} \text{sr}^{-1}} \right)^{6/7}.$$

The term $f$ represents several sources of uncertainties in converting from $L_{\nu_{151}}$ to $Q_{jet}$ and we adopt $f = 20$ (e.g. Cavagnolo et al. 2010), though Willott et al. (1999) found that $f$ could take values between $1 \leq f \leq 20$. Cavagnolo et al. (2010) compared the jet powers they obtained using $f = 20$ to the jet powers using $f = 1$ and found their jet powers were two orders of magnitude lower. Cavagnolo et al. (2010) determined $f = 20$ using observations of low-luminosity radio galaxies with powers comparable to RQQs while Willott et al. (1999) had observations of RLQs which led to $f$ values as low as 1. Since our data are RQQs we use $f = 20$.

We are using 1.4 GHz radio luminosity densities which we convert to 151 MHz assuming $S_\nu \propto \nu^{-\alpha_R}$ and $\alpha_R = 0.8$, to determine the jet power we modelled over using Equation 6.12.
The jet efficiencies are calculated using Equation 6.10 and assuming $\epsilon = 0.057$. Since $\eta$ depends on the optical luminosity density through the bolometric luminosity and the radio luminosity density through the jet power, we assume that the probability density function for the radio loudness parameter can be used as the probability density function for the jet efficiency.

In addition to the slopes from the modelling of the underlying QSO radio flux density, the models we constructed to model the distribution of $\mathcal{R}$ are used to model the distribution of jet efficiency $\eta$. The two models that are strongly preferred over model A (the single power-law, see Section 6.2), are shown in Figures 6.8 and 6.9.

![Figure 6.8](image)

**Figure 6.8:** The Schechter function model B (green) modelling of the $\mathcal{R}$ distribution: a Schecter function with index -1.7 and break value $\mathcal{R}^* = 22$. This model has been used as a proxy for the probability density functions for the jet efficiencies which have been calculated using Equation 6.10 and the observed optical luminosity densities and the model radio flux densities from Chapter 5. The optical luminosity densities have been converted to the bolometric luminosity and radio jet power via Equations 6.11 and 6.12 respectively (black lines). The dashed vertical line shows the minimum jet efficiency from the observed mean radio flux and optical magnitude from our stacked samples ($\eta = 4 \times 10^{-4}$).

The minimum jet efficiency from our radio flux density models, which is given
by the mean radio luminosity density and the mean optical luminosity density, is \( \eta_{\text{min}} = 4 \times 10^{-4} \). There is no hint of a turnover at the low \( \eta \) end of our models. Therefore we expect more objects with \( \eta < \eta_{\text{min}} \). Bearing in mind this jet efficiency has been calculated assuming a low spin value of \( \dot{a} \leq 0.1 \), we compare the minimum jet efficiency value \( \eta_{\text{min}} \) to the different theoretical models of jet efficiencies shown in Figure 2 of Martínez-Sansigre \\Rawlings (2011). All of the jet efficiency models reach spins of 0.0, although due to a lack of space and dynamic range the plots may not show the jet efficiencies at these low spins (i.e. the model by McKinney, 2005 shown in Figure 2 of Martínez-Sanisgre \\Rawlings, 2011). However, most models have high jet efficiencies at this spin value. The models from McKinney (2005)
and Tchekhovskoy, Narayan & McKinney (2010) probe as low as our minimum jet efficiency at $\hat{a} = 0.1$ although their models drop off sharply at low $\hat{a}$. The other models that Martinez-Sansigre & Rawlings (2011) show have higher jet efficiencies at the spin value $\hat{a} = 0.1$, but each of their model slopes flatten towards a minimum $\eta$ at spins $\hat{a} = 0$ (Hawley & Krolik 2006; Nemmen et al. 2007; Benson & Babul 2009). We reiterate our assumption that the star formation contribution to the radio flux density is minor. If star formation was to make up a significant amount of the radio flux density in the objects we have stacked, the estimate on the minimum jet efficiency may not be reliable.

6.4 Conclusion

We have modelled the radio emission from undetected SDSS QSOs and found the distribution of the underlying QSO radio flux density in each optical magnitude sample to be a power-law with a negative index, that steepens with decreasing optical luminosity. Note that we are using radio undetected QSOs and so we do not look at the radio dichotomy. Our aim was initially to investigate the possibility of a radio lower envelope that was hinted at in previous stacking experiments. As deep as our data goes, it suggests there is no lower envelope in terms of the radio power as was shown in Chapter 5.

In the current chapter, the distribution of the radio flux densities in each optical magnitude sample is converted to distribution of the radio loudness parameter $\mathcal{R}$. The minimum $\mathcal{R}$ parameter value and the shape of the distribution in $\mathcal{R}$ shown in Figure 6.2 does not show any hint of a turnover. Therefore we expect $\mathcal{R} < \mathcal{R}_{\text{min}}$, and our preferred model of this distribution supports this (model B the Schechter function). The lack of a minimum radio loudness translates into no minimum jet efficiency in as far as the data allows us to infer. Data that goes deeper, beyond
6.4 Conclusion

our sensitivity limit of \( \sim 60 \mu\text{Jy} \), may reveal new insights but as far as our data is concerned, there seems to be no minimum jet efficiency.

For a narrow range of radio luminosity densities, the distribution of jet efficiencies resembles a luminosity function which we suspect is actually the optical luminosity function for quasars. The two slopes from the double power-law in model C are very close to the QSO optical luminosity function of Croom et al. (2009), illustrating the likeness. This leads to the idea that the radio loudness is not a fundamental property of the QSO but rather the ratio of two independent properties, the radio and optical luminosities. Physically this suggests that the optical and radio luminosities may be driven by processes that are approximately independent.

We are sensitive to jet efficiencies down to \( \eta_{\text{min}} = 4 \times 10^{-4} \) to which value no turnover is observed, hence we expect jet efficiencies below \( \eta_{\text{min}} \) to exist. Therefore models should be able to produce jet efficiencies below this value.

We assumed a spin value of \( \dot{a} = 0.1 \). If we assumed a lower spin instead, e.g. \( \dot{a} = 0.01 \), the radiative efficiency would remain essentially unchanged (approaching 0.057; Novikov & Thorne, 1973), and hence our jet efficiencies would also be similar. However, if we consider the models used in Martinez-Sansigre & Rawlings (2011), only three of the six models probe down to these low spins. The models from Tchekhovskoy, Narayan & McKinney (2010) and McKinney (2005) do reach efficiencies below our limit. The three other models from Hawley & Krolik (2006), Nemmen et al. (2007) and Benson & Babul (2009) have minimum jet efficiencies above our limit, and hence do not meet our inferred constraint. In these models they assume thick advection dominated disks with magnetic fields whose strength is set initially, so the magnetic pressure is some assumed fraction of the total gas pressure. Nemmen et al. (2007) produce models for two different viscosity parameters of the disk where the viscosity parameter characterises the flow within the disk. Their hybrid models, with the jet power dependent on the black hole spin, accretion flow
and accretion rate, can produce jet efficiencies down to $\sim 10^{-4}$ at spins of 0 with a lower viscosity parameter (see Figure 1 Nemmen et al. (2007)) than the model which is used in Martinez-Sansigre & Rawlings (2011).

The assumptions we have made when converting $R$ to the jet efficiency introduce many sources of uncertainty. Firstly, converting the observed B-band optical luminosity to the bolometric luminosity was done via a correction term. We used the correction term from Hopkins, Richards & Hernquist (2007) where the bolometric correction was given in the form of a double power-law function of the luminosity. The uncertainty in this correction term at B-band luminosities was realised to follow Equation 3 in Hopkins, Richards & Hernquist (2007). A typical uncertainty in the $\log_{10}(L_{\text{bol}})$ for our data determined from the B-band luminosity density is $\sigma_{\log_{10}(L_{\text{bol}})} = 0.08$.

At low spins the jet efficiency is not dominated by the effects related to the spin of the black hole, e.g. frame-dragging and the Blandford-Znajek mechanism. The Blandford-Znajek mechanism causes the jet power to be dependent on the spin, hence at spins of $\hat{a} = 0$ the jet power goes to zero. However, there may still be radio emission via disc processes in these models (see the wind efficiencies in Figure 4 of Tchekhovskoy, McKinney & Narayan, 2012 for example). Again, we restate here our assumption that the radio emission from the radio-undetected sources is due to the AGN and not from star formation in the host (see Section 4.3.3 for a discussion on this). However we acknowledge that it is possible, probable even, that there is at least some contribution from star formation to the radio emission in these objects.

The recovery of a function that is very similar to the optical luminosity when we modelled the $Rs$ distribution holding the radio luminosity fixed, leads to our conclusion that for our data the optical and the radio luminosities are independent. This has repercussions on the models of jet production, especially those that attempt to explain jet production via accretion since in those cases the optical and radio
6.4 Conclusion

Luminosities would not be independent. In any case the models need to be able to produce jet efficiencies at least as low as our minimum jet efficiency ($\sim 4 \times 10^{-4}$) which is an observational constraint on the lower efficiencies required of a jet production model. The conclusion postulated here assumes that the AGN jet contribution dominates the radio flux density in all objects. If, however, there is significant star formation and AGN that do not produce jets in our stacked samples, this results in a blending of the jet and SF contributions to the radio flux density which limits the conclusions we can draw.

6.4.1 Future work

This study focused on the SDSS photometrically selected QSOs that are undetected in the NVSS sample. With the future spectroscopic QSO catalogues that give an increase in the numbers of spectroscopically confirmed QSOs we can avoid the uncertainties in the photometric redshifts, and any contamination from stellar emission in the classification of the QSO.

The very near future will see the rise of the radio regime, in particular in the lower frequencies with the Square Kilometre Array (SKA) interferometer under planning and commissioning phases. The 1.4 GHz Evolutionary Map of the Universe (EMU)\(^\text{1}\) which will use the Australian Square Kilometer Array Pathfinder (ASKAP)\(^\text{2}\) telescope, aims to reach an rms sensitivity of $\sim 10 \, \mu$Jy. This survey has not yet been started, although it is likely to begin in the next few years.

Currently undergoing observations is the LOw Frequency ARray (LOFAR) Million Source Sky Survey in low frequencies (30 - 240 MHz). Using the full international capabilities of LOFAR the sensitivities in the higher frequency components reach sub mJy\(^\text{3}\).

\(^{1}\)http://www.atnf.csiro.au/people/rnorris/emu/index.html
\(^{3}\)http://www.astron.nl/radio-observatory/astronomers/lofar-imaging-capabilities-
These optical and radio surveys will be able to lower the sensitivity limit, therefore we will be able to detect more objects. In addition to survey data, a campaign of deep observations of radio quiet QSOs at a constant redshift could also provide valuable information on the nature of the radio emission in these objects. Non-detections are expected by our model of the radio flux density distribution which indicates that there are more sources with progressively fainter radio flux densities.
Chapter 7

Conclusion

This thesis is split into two individual parts: observations and analysis of the molecular gas in an obscured quasar, and a study into a population of radio quiet quasars, optically selected and matched with a radio survey. The following sections are split into concluding remarks for each of the two projects.

7.1 Molecular gas in an obscured quasar

In Chapter 2 observations of the molecular gas towards the unlensed, obscured, $z = 2.8$ quasar AMS12 that revealed strong detections of high-excitation CO and both fine structure lines of [CI] were analysed. AMS12 is the first unlensed, high-redshift source to have both [CI] lines detected. The high CO lines (3-2), (5-4), and (7-6) we used to construct the CO SED (e.g. Figure 3.3). The highly-excited molecular gas probed by CO(3-2), (5-4) and (7-6), is modelled with large velocity gradient (LVG) models in Chapter 3. The gas kinetic temperature $T_G$, density $n(H_2)$, and the characteristic size $r_0$, are determined assuming no prior information on any of the parameters and then again using the dust temperature from the FIR SED as a prior for the gas temperature (see Appendix A for the fitting of the FIR.
The LVG models are fit finding parameters which give the maximum likelihood. The best fitting parameters assuming a flat prior on all the parameters including the temperature gives $T_G = 11.6$ K, $n(H_2) = 10^{6.1}$ cm$^{-3}$ and $r_0 = 6.2$ kpc.

When we assume the temperature has the prior given by the dust temperature PDF the best fitting parameters are $T_G = 89.6 \pm 8$ K, $n(H_2) = 10^{3.9} \pm 10^{0.06}$ cm$^{-3}$ and $r_0 = 0.8 \pm 0.001$ kpc. Figures 3.4 and 3.5 show the contours and the best-fitting LVG model.

The ratio of the [CI] lines can give the [CI] excitation temperature under the assumption that the [CI] lines are optically thin. The excitation temperature given by the [CI] lines in AMS12 is $T_{ex} = 43 \pm 10$ K.

The difference between the [CI] excitation temperature and the gas kinetic temperature inferred from the high-excitation CO indicate they are not in thermal equilibrium. The [CI] excitation temperature is lower than the dust temperature and the gas kinetic temperature of the high-excitation CO, which may indicate that the [CI] lies at a larger radius. Since [CI] is closely related to the low-excitation CO, namely CO(1-0) there may be a large reservoir of cool gas possibly detectable through the CO(1-0) line.

In Section 3.5.1 we estimate the gas mass using four different methods; one method uses the [CI]($^3P_1 - ^3P_0$) to estimate the strength of the CO(1-0) line while the other three assume that the molecular gas represented by a single highly-excited component. The CO gas mass from the [CI]($^3P_1 - ^3P_0$) line estimate is higher than the CO gas mass estimated using the high excitation CO. There may be ($\sim 30\%$) of the molecular gas that is missed from the high-excitation line analysis, giving a gas
mass higher than that inferred from the assumption that the high-excitation gas is a good tracer of the low-excitation gas.

The inferred parameters from the observations and modelling suggest that the gas and dust in the host of AMS12 may be heated by the AGN as well as from star formation. The SFR inferred from the FIR dust temperature is $\sim 5300 \ M_\odot \ \text{yr}^{-1}$ which is seen only in the most extreme starburst galaxies (Chapman et al., 2005; Solomon & Vanden Bout, 2005; Tacconi et al., 2006; Coppin et al., 2008). The quasar bolometric luminosity of AMS12 suggest that the AGN can heat up to the dust temperature of 88 K out to scales of $\sim 3 \ \text{kpc}$ assuming that the UV photons travel unhindered. The characteristic scale of dust is found to be $\sim 2 \ \text{kpc}$ in many objects (Greve et al., 2005; Tacconi et al., 2006; Younger et al., 2008), so the dust temperature of 88 K for the FIR dust in AMS12 is achievable through AGN heating. Given the high dust and gas kinetic temperatures that may be heated by the AGN, the $L_{\text{FIR}}$ in this object may not portray an accurate SFR, which would be severely over-estimated if the dust was significantly heated by the AGN.

The detection of both [CI] lines provides an independent estimate on the temperature of the gas given by the ratio of the lines' brightness temperatures. The temperature given by the [CI] is significantly lower than both the dust and high-excitation gas temperatures, suggesting there may be an additional cooler, diffuse component of the gas. The abundance of [CI] in this object is found to be $5-8 \times 10^{-5}$ which indicates the molecular gas is already enriched which supports the findings from Walter et al. (2011).

AMS12 has a stellar mass of $\sim 3 \times 10^{11} \ M_\odot$ which would only increase by $\sim 15\%$ if all the molecular gas was converted into stars at 100% efficiency. The Eddington limited black hole mass was estimated to be $\gtrsim 1.5 \times 10^{9} \ M_\odot$ giving $M_*/M_{\text{bulge}} \gtrsim 0.005$. This is greater than the local $M_*/M_{\text{bulge}}$ relationship determine by Häring & Rix (2004) $M_*/M_{\text{bulge}} \approx 0.002$. For AMS12 to evolve to the local
relationship, the bulge would have to grow \( \sim 3 \) times as much as the central black hole from \( z = 2.8 \) to \( z = 0 \). Given that the molecular gas mass we have inferred is at most \( \sim 15\% \) of the current stellar mass, with secular evolution AMS12 would not evolve to the local relationship.

AMS12 is host to a massive black hole, has enriched molecular gas and has already amassed most of its stellar mass at \( z = 2.8 \). These properties indicate AMS12 is a mature system. It requires a significant increase in the stellar bulge mass compared to the black hole mass in order to evolve to the local relationship. However, this object was selected for its high \( L_{\text{bol}} \) and there is the possibility it is accreting at a super-Eddington rate, and so the departure from the local relationship could be due to selection bias.

## 7.2 Stacking of radio undetected QSOs

In the second half of this thesis we investigate the existence of a lower radio envelope in the “radio loudness” property of quasi-stellar objects (QSOs). Our population is comprised of NVSS undetected QSOs from a volume-limited sample of SDSS QSOs of the photometric catalogue of \cite{Richards2009}. The optically selected QSOs in our sample were selected based on the following criteria:

- optical photometric redshifts within a narrow range \( 0.3 < z_{\text{phot}} < 0.5 \)
- NVSS beam integrated flux below \( 3\sigma_{\text{rms}} \)
- no bright outliers
- each postage stamp NVSS cut-out the same size.

Following this selection procedure 63424 radio image cut-outs of the QSOs remained. These were split into eight samples according to their optical magnitudes.
In order to characterise undetected sources in the NVSS survey, methods of minimising the effect of the noise, or adding the signals from the undetected sources must be applied. Stacking the undetected sources lowers the noise in the mean image so it is possible to detect fainter radio flux densities.

The radio images that were not detected in the NVSS survey around each SDSS QSO position in our sample were stacked and the mean radio flux density in each sample was measured. The stacked images were detected in all but the faintest optical magnitude samples.

Figure 4.6 shows the radio luminosity density versus the optical luminosity density for our samples, as well as the NVSS detected QSOs in the SDSS photometric quasar catalogue within our redshift range. The stacked mean radio flux densities suggest that there may be a lower radio envelope. The lower envelope physically may be interpreted as a minimum radio power for a given optical luminosity. However, there are several selection effects which may be creating an artificial correlation between the radio and optical samples. Notably, there are more faint objects than bright ones, therefore when matching independent surveys in different wavelengths, there is a ‘clustering’ of detected sources at each of the surveys’ detection limits.

To determine whether the mean statistic does represent the underlying distribution of the QSO radio flux density we have modelled the distribution of sources within each stacked sample. The models were made from two components; a distribution of the noise in the NVSS survey is combined with a theoretical model of the underlying QSO radio flux density (either a single Gaussian or a single power-law).

The measured QSO flux density distributions (shown in Figure 5.1) were fitted with models of the underlying QSO radio flux density with the noise distribution added. The best-fitting parameters were found for each model by maximising the likelihood, and the different models were compared using the odds ratio in Section 5.6. The power-law model is the preferred model for the underlying QSO radio
7.2 Stacking of radio undetected QSOs

flux density distribution. The best-fitting indices range from -1.7 in the optically brightest sample getting progressively steeper as the optical magnitude becomes fainter to -3.375 in the optically faintest sample.

The power-law model with negative indices suggests that instead of some characteristic minimum flux density which is correlated with the optical luminosity density, there are actually more QSOs that are progressively fainter and we expect even more objects at flux densities fainter than our sensitivity limit of 61.25 $\mu$Jy. In other words, there is no lower envelope; the appearance of Figure 4.6 gives an artificial lower envelope, but this is due to the fact that the distribution of the NVSS noise which is approximately Gaussian in shape is hiding the true underlying distribution of the QSO radio flux density.

Stacking only provides a meaningful result if the distribution of the property of interest can be characterised by the mean statistic (or median).

The radio loudness parameter is often used to quantify the contribution of the radio emission (i.e. from jets) to the thermal, optical continuum. The distribution of the radio loudness may contribute to determining whether there is a dichotomy in the radio loudness between RQQ and RLQs. Constraining the distribution of the radio loudness is plagued by selection effects similar to those affecting the determination of the existence of a radio lower envelope. Samples of optically selected QSOs are often incomplete in the radio surveys due to the sensitivity limits of these surveys.

We have modelled the radio flux distribution of the NVSS undetected QSOs revealing an underlying power-law distribution to the radio flux densities. We have used our results from the modelling of the radio flux density distributions to model the distribution of the radio loudness parameter $R$. Three models were chosen based on the shape of the distribution in Figure 6.2. The three models are a single power-
law, a Schechter function and a double-power law function. The single power-law was strong disfavoured compared to the other two models (via the odds ratio). Neither the Schechter function nor the double power-law function were strongly favoured over the other. These two models have a very similar shape to the QSO optical luminosity function \cite{Croom2009}, which may suggest the radio loudness is not a fundamental property of the QSO but rather the ratio of two independent properties, the radio and optical luminosities. Thus, the probability of having a radio loudness value is essentially given by $p(R) \approx p(L_{\text{rad}})p(L_{\text{opt}})$ summed over all $p(L_{\text{rad}})$, i.e. the product of the probabilities of having certain radio and optical luminosities. In each wavelength band, higher luminosity objects are rarer therefore this is reflected in the $R$.

Assuming jet power is related to the accretion rate via the jet efficiency (which we have assumed is a function of black hole spin), and a value for the radiative efficiency converting the accretion rate into bolometric luminosity, we relate the radio and optical luminosity densities to jet efficiencies. Our models of the radio-loudness distribution are recycled since they depend upon the radio and optical luminosity densities also. We are sensitive to jet efficiencies down to $\eta_{\text{min}} = 5 \times 10^{-4}$ to which value no turnover is observed. This $\eta_{\text{min}}$ is due to our sensitivity limit and we expect there to be jet efficiencies lower than our minimum value $\eta < \eta_{\text{min}}$. We have provided an observational constraint on the minimum jet efficiency, hence jet production models need to be able to produce jet efficiencies at least as low as our minimum jet efficiency ($\eta_{\text{min}} = 5 \times 10^{-4}$).

### 7.3 Future work

For the molecular gas in the obscured quasar AMS12, future observations of this object would provide useful information about the distribution of the molecular gas
and constrain the heating mechanisms of the gas and dust in this object:

- mapping the low-excitation CO would reveal whether this system consists of multiple components; a diffuse, low-excitation gas component with a more concentrated, high-excitation component from which arises the strong high-excitation CO lines

- observing even higher CO transitions, such as CO(8-7), CO(9-8) and above could constrain the high-excitation end of the CO ladder.

With the detection of higher CO transitions we could investigate in detail whether the heating mechanism of these lines involves only photons or has an X-ray component such as has been seen in lower redshift objects. Recent SPIRE spectra of lower redshift objects have revealed a plethora of molecular lines previously unattainable by ground-based observations. These lines have been successfully modelled by a combination of photon dominated regions (PDRs) and X-ray dominated regions (XDRs) (e.g. Meijerink & Spaans [2005] [van der Werf et al., 2010]), suggesting that the excitation to higher molecular transitions is directly linked to the AGN.

The investigation into the radio-quiet QSOs can be expanded by repeating the investigation in another redshift band (i.e. \( z \sim 1 \) where the number density of the SDSS photometric QSO catalogue peaks; [Richards et al., 2009]), to determine whether the modelling of the radio flux densities recovers the same shape (power-law distributions). Determining the radio flux density distribution at different redshifts could also provide information on whether there is any evolution in the distribution.

In the near future, deeper observations in both optical spectroscopy and radio surveys could be used to, for example:

- using spectroscopic QSO catalogues can avoid the uncertainties in the photometric redshifts, and any contamination from stellar emission in the classifi-
7.3 Future work

cation of the QSO

- developments in future radio sky surveys such as LOFAR, and SKA will observe down to $\mu$Jy sensitivities.

The $\mu$Jy sensitivities reached by these radio surveys would mean more faint radio sources would be detected, reducing the need for stacking and providing further observational constraints on the distribution of radio flux densities.
Appendix A

The far infrared spectral energy distribution

The far-infrared spectral energy distribution was constructed from existing observations and the continuum measurements from the Plataeu de Bure Interferometer made when observing the CO lines. The Herschel data was provided by Mark Lacy and the SED fitting was completed by Alejo Martínez-Sansigre and appears in Schumacher et al. (2012).

A.1 Continuum measurements

A compilation of data between 70 $\mu$m and 3.0 mm is used to infer the best-fitting parameters for the FIR SED of AMS12. The fluxes are given in Table A.1.

A.1.1 Existing data: MAMBO and Spitzer

The Max-Planck Millimetre Bolometer Array (MAMBO) observations were done at a wavelength of 1.2 mm. The source was observed with other sources in the AMS
Continuum measurements

A sample in blocks of typically 20 minutes. The data were reduced using the MOPSIC pipeline. The rms noise achieved was \( \sim 0.55 \) mJy beam\(^{-1} \) (see\footnote[1]{http://hermes.sussex.ac.uk} Martínez-Sansigre et al., 2009 for details).

Archival Spitzer measurements of AMS12 at 160 and 70 \( \mu \)m are used. These were made as part of the Spitzer extragalactic First Look Survey (FLS). See Frayer et al. (2006) for details on reduction and handling.

A.1.2 New data: Herschel and PdBI

The FLS field was observed as part of the Herschel Multi-tiered Extragalactic Survey (HerMES\footnote[1]{http://hermes.sussex.ac.uk}). Fluxes of the AMS objects at 250, 350 and 500 \( \mu \)m were measured off the level 2 SPIRE mosaics distributed by the Herschel Science Archive. Aperture photometry was carried out in apertures of 13, 17 and 23\" at 250, 350 and 500 \( \mu \)m, respectively, with background annuli 23-60\", 30-100\" and 40-140\". Aperture corrections were derived from the beams associated with version 1.0 of the SPIRE beam release note\footnote[2]{Sibthorpe et al., 2010}, sampled with 1" pixels, and the corresponding beam areas applied. The beam areas assumed were 426, 771 and 1626 square arcseconds at 250, 350 and 500 \( \mu \)m, respectively, and aperture corrections were 1.51, 1.52 and 1.60, respectively.

The continuum towards AMS12 in the PdBI 3 mm band was found over the line free bandwidth 684.4 MHz or 2053.8 kms\(^{-1} \), leading to an rms noise of 0.09 mJy. For the 2 mm band, 3.35 GHz of the line free spectrum was fit yielding an rms noise of 0.06 mJy. The 1.3 mm band yielded the continuum measurement over a spectral line free region spanning 2.572 GHz, giving an rms noise of 0.19 mJy.
### A.2 Modelling the FIR emission

At FIR wavelengths dust is not optically thick, and the SED of the radiation can be described by a modified black body. If the absorption coefficient of the dust, $\kappa(\nu)$, is assumed to follow a law $\propto \nu^\beta$, the emission will be given by:

$$L_\nu = \frac{A \nu^{3+\beta}}{(e^{h\nu/kT_D} - 1)} \tag{A.1}$$

where $A$ is a normalisation term given by:

$$A = \frac{L_{\text{FIR}}}{\zeta(\beta + 4)\Gamma(\beta + 4)} \frac{h}{kT_D}. \tag{A.2}$$

The three variables are: the dust temperature $T_D$, the emissivity index $\beta$ and the FIR-luminosity $L_{\text{FIR}}$. Here $h$ and $k$ are Planck’s and Boltzmann’s constants, respectively, while $\zeta$ and $\Gamma$ are the Riemann zeta function and the Gamma function, respectively.

In order to determine the parameters of the fit to the FIR SED, we use the parameter estimation methods outlined in Section 3.3. Figure A.1a shows the contours

<table>
<thead>
<tr>
<th>$\lambda_{\text{obs}} , (\mu\text{m})$</th>
<th>Flux (mJy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>$15.0 \pm 5$</td>
</tr>
<tr>
<td>158</td>
<td>$74.0 \pm 20$</td>
</tr>
<tr>
<td>250</td>
<td>$65.0 \pm 6$</td>
</tr>
<tr>
<td>350</td>
<td>$49.4 \pm 5$</td>
</tr>
<tr>
<td>500</td>
<td>$42.5 \pm 12$</td>
</tr>
<tr>
<td>1200</td>
<td>$3.7 \pm 0.6$</td>
</tr>
<tr>
<td>1400</td>
<td>$2.16 \pm 0.19$</td>
</tr>
<tr>
<td>1960</td>
<td>$1.01 \pm 0.06$</td>
</tr>
<tr>
<td>3260</td>
<td>$0.75 \pm 0.09$</td>
</tr>
</tbody>
</table>

Table A.1: The flux measurements of the FIR infrared measurements from observations by Spitzer/MIPS, Herschel/Spire, MAMBO and PdBI.
of dust temperature and emissivity index, marginalised over the FIR luminosity, $P(T_D, \beta | S_\nu)$. Figures A.1b, A.1c and A.1d show the posterior distribution function for the dust temperature, emissivity index and $L_{\text{FIR}}$ marginalised over the other two parameters respectively.

### A.2.1 Far-infrared luminosity

Figure A.2 shows the FIR SED for AMS12. The parameters that give the maximum likelihood are a temperature $T_D = 88$ K and emissivity index of $\beta = 0.6$. There are two points which are not well fit by the model; the 3 mm PdBI measurement, and the 500 $\mu$m Herschel point. The 3 mm point could possibly suffer from contamination from the radio continuum from the AGN, AMS12 has a steep extended radio spectrum and a flatter, inverted spectrum within a compact 150 pc region. Extrapolating the radio continuum out to $\sim 90$ GHz using the flatter compact spectral index ($\alpha = -0.22$ where $S_{\text{radio}} \propto \nu^{-\alpha}$), (Klöckner et al., 2009), shows a significant amount of the measured flux could be due to radio emission from the AGN.

The 500 $\mu$m measurement could be boosted by FIR emission lines. Smail et al. (2011) have estimated the FIR emission lines may contribute $\gtrsim 20 - 40\%$ to the broad-band flux, thus there could be possible contamination from the [CII] ($\nu_{\text{rest}} = 158\mu$m) line which at $z = 2.7672$ is at 595 $\mu$m, (the Herschel/Spire 500 $\mu$m band has a width of $\lambda/\Delta\lambda = 3$, Griffin et al., 2006).

Figure A.2 illustrates the importance of obtaining points at both the longer and shorter wavelengths: it is necessary to probe wavelengths shorter than those where the peak of the emission appears, to constrain the location of this peak and hence the temperature. In order to constrain the emissivity index, long wavelength data are critical, in this case the data at 1.2 mm from MAMBO, and the additional longer...
wavelength data from PdBI.

The best fitting $L_{\text{FIR}}$ determined from the single graybody model fitting of the FIR SED is $\log_{10}(L_{\text{FIR}}/L_\odot) = 13.5$. 
A.2 Modelling the FIR emission

Figure A.1: (a): The contours of the dust temperature versus the emissivity index for AMS12 marginalised over $L_{\text{FIR}}$. The best fits to the temperature and emissivity index correspond to 88 K and 0.6. (b): The posterior PDF for the dust temperature, marginalised over $\beta$ and $L_{\text{FIR}}$, with best fit 89 ± 8 K. (c): The posterior PDF for the $\beta$ parameter, marginalised over $T_D$ and $L_{\text{FIR}}$, with best fit value 0.6 ± 0.1. (d): The posterior PDF of the $L_{\text{FIR}}$, marginalised over $T_D$ and $\beta$, the best fit is $\log_{10}(L_{\text{FIR}}/L_\odot)=13.5$. 
Figure A.1: (continued).
Figure A.2: The graybody dust model fit to AMS12’s FIR SED. The points are from observations from PdBI, MAMBO, Herschel/Spire and Spitzer/MIPS. The best fitting temperature and spectral index are shown on the figure.
References


Begelman M. C., Blandford R. D., Rees M. J., 1984, Reviews of Modern Physics, 56, 255


References


References

Fabian A. C., 1994, ARA&A, 32, 277
—, 2012, ARA&A, 50, 455


Ivezić Ž. et al., 2002, AJ, 124, 2364


References


Krolik J. H., 1999, Active galactic nuclei: from the central black hole to the galactic environment


References


Osterbrock D. E., Ferland G. J., 2006, Astrophysics of gaseous nebulae and active galactic nuclei


Peterson B. M., 1997, An Introduction to Active Galactic Nuclei


Schneider D. P. et al., 2010, AJ, 139, 2360


Sibthorpe B. et al., 2010, A&A, 518, L130+


Smolčić V. et al., 2008, ApJS, 177, 14


References


van der Werf P. P. et al., 2010, A&A, 518, L42+


